

# **Shifting foreign policies and punctuated diplomatic behavior**

**Eric Dunford**

*McCourt School of Public Policy*

*Georgetown University, 3700 O St NW, Washington, DC 20057*

*email: edunford@umd.edu*

**Michael F. Joseph**

*Political Science*

*Yale University, 31 Hillhouse Avenue, New Haven, CT 06520*

*email: michael.frederick.joseph@gmail.com*

*Word Count: 13,916*

---

Thanks to Brenton Kenkel, Robert Trager, James Morrow, Yon Lupu, Michael Miller, Jeffery Friedman for comments that improved the paper

# Shifting foreign policies and punctuated diplomatic behavior

*Word Count: 13,916*

## Abstract

We develop a method to identify when states change their foreign policies based on an observable indicator: patterns of elite diplomatic meetings. We argue that elites choose diplomatic partners to advance a specific foreign policy agenda. When that agenda changes so do the incentives to choose diplomatic partners. To locate these breaks we apply non-parametric structural break tests to time series models that predict a state's diplomatic behavior. We argue that where these tests identify breaks in diplomatic behavior, a state has changed its foreign policy. We validate our theory using expert foreign policy analysis and quantitative cases. We first collect new daily diplomatic events data for Russia, Iran, the USA and Australia. We then compare structural breaks in these time series to expert assessments about foreign policy change. Consistent with expert reports, we locate structural breaks in Iran's diplomatic behavior when Rouhani first comes to power, then in the months after sanctions are lifted; and in Russia's diplomatic behavior 6 months before the Ukraine crisis, and then again when Russia extends its military into Syria. No break occurs in Australia's or America's diplomatic behavior as expected. We contribute to the empirical literature on conflict by providing new diplomatic data and a method to measure foreign policy change, and to theories of diplomacy by linking aggregated patterns of diplomatic behavior to foreign policy choices—not underlying intentions.

# 1 Introduction

Measuring changes in states' foreign policies is crucial to scholars and policy-makers alike. In the United States context, policy-makers want to know if Russia is intending to re-ignite the Cold War; if Iran really abandoned its nuclear aspirations in favor of global economic engagement; and if China's investments in Africa and Latin America signs of a more expansive foreign policy? Policy-makers want to anticipate shifts in their rival's foreign policies. However, it is difficult to identify when a state embarks on a new foreign policy because certain rivals want to conceal their policies. The incentives to misrepresent are highly contextualized and aggregated measures of observable behaviors often miss finer dynamics that are apparent in qualitative evidence ([Arena and Joyce 2015](#); [Minhas et al. 2016](#)). We offer a new framework that explains how changes in a state's foreign policy manifest into measurable behaviors and how those indicators of change elucidate the onset of crises, rivalries, and cooperative relationships in critical cases. To do so, we turn to an untapped source: patterns of diplomatic meetings.

A recent wave of scholarship shows that leaders use high-level diplomacy to effectively signal resolve ([Sartori 2005](#); [Kurizaki 2007](#); [Avenhaus and Zartman 2007](#); [Trager 2010](#); [McManus 2018](#)), communicate preferences ([Trager 2011](#); [Gibler 2008](#)), negotiate cooperative agreements ([Fang and Stone 2012](#); [Ramsay 2011](#)) and mediate disputes ([Kydd 2003](#); [Debs and Weiss 2016](#)). These theories all explore what leaders can accomplish during a specific meeting. They assume, however, that leaders bare no cost from choosing to meet with their counterparts. Thus, they cannot explain why leaders choose who they meet with and who they ignore. In fact, the conventional wisdom is that aggregated patterns of meetings correspond with a state's security and commercial interests ([JÖNSSON 2002](#)). This has led many to assume that leaders meet with the same partners year after year ([Lebovic and Saunders 2016](#)). We assume that leaders do focus their diplomatic efforts on their security and commercial partners, but within this set, they choose the partners that most closely align with

their specific foreign policy objectives. When their foreign policies change, they will choose different partners within those that serve their security or commercial interests.

We argue that shifts in aggregated patterns of diplomatic meetings—face-to-face meetings, scheduled in advance, between the heads of state or cabinet officials responsible for foreign policy of two countries—can indicate a state has embarked on a foreign policy. So long as leaders pursue the same foreign policy they are likely to meet with the same counterparts consistently. However, when leaders pursue a new foreign policy they must lay groundwork. They are forced to develop new status of force agreements, consolidate trade partners, build new intuitions, and reassure allies. Thus, at times of foreign policy change, leaders will stop meeting with old partners who served the old agenda and focus on new relationships. This generates a dramatic change in what motivates a leader to choose who to meet with.

We argue that new foreign policies produce a sudden shift in choices of diplomatic partners, rather than an expansion of the number of meetings or a gradual shift for two reasons. First, leaders are busy ([Lindsey and Hobbs 2015](#)). If they want their new policy to succeed they must focus their limited time on achieving the objectives they care about the most. Second, meeting with many partners dilutes the signaling effects of any specific meeting. Thus, we expect a stark difference in the predictors of a state’s diplomatic choices under different foreign policy regimes.

Using these insights, we develop a statistical technique that identifies breaks in the patterns of a state’s diplomatic behavior. We start with the diplomatic choice model suggested by [Lebovic and Saunders \(2016\)](#), who show that diplomatic visits are predicted by commercial and security ties, region, as well as the strategic importance of the target state at a given point in history. Our insight is that while these variables will consistently predict diplomatic engagements, their relative importance will change depending on a state’s foreign policy. To locate these breaks we employ a structural break test that looks for time-partitions in the data that improve the model fit ([Bai and Perron 2003; Zeileis 2005](#)). The test returns a specific time period (or periods) if they exist that improve the model fit if the

data is separated into these periods. We argue that each partition corresponds with a break in a state's foreign policy.

We validate the relationship between diplomatic breaks and new foreign policies using quantitative case studies of Australia, Iran, Russia and the United States. We report new daily diplomatic meetings event data for the leaders and foreign ministers of the United States, Iran, and Russia and the foreign minister of Australia using a web-scraping and text recognition algorithm we designed. Our data includes all diplomatic meetings (not just visits) and therefore is an expansion over existing collection efforts. We estimate structural break models for each state. We show that structural breaks appear at times that are consistent with qualitative reporting on foreign policy shifts in these states. Yet breaks do not appear at other plausible periods such as changes in leaders or Secretary of State, elections, or financial shocks.

In this article, we make three contributions. First, we develop a new theory that links patterns of diplomatic engagements to foreign policies not underlying preferences. In doing so, we provide a framework that aggregates the insights from studies that show diplomacy serves communication and negotiation functions in specific contexts ([Trager 2010](#); [Sartori 2005](#); [Kurizaki 2007](#); [Avenhaus and Zartman 2007](#); [Ramsay 2011](#); [Fang and Stone 2012](#); [Kydd 2003](#); [Fey and Ramsay 2010](#)) or signal to third-parties ([Bull 2002](#); [Evans et al. 1993](#); [Whitehead 2006](#); [Howell and Pevehouse 2007](#); [Rathbun 2014](#)) to explain the overall pattern of a state's diplomatic behavior. Second, we find that changes in a state's agenda are observable in partnerships, and that changes in partnerships can help understand and anticipate the onset of crises, rivalries and cooperative relationships ([Goldstone et al. 2010](#); [Cederman and Weidmann 2017](#); [Weidmann and Ward 2010](#)). Finally, we demonstrate how a combination of micro-level meeting event data and non-parametric statistics can serve as a strong basis for quantitative case analysis. The combination of methods helps us make sense of puzzling findings from prior research and illuminate the precise relationships between diplomatic choices state interests and foreign policy.

### 1.1 Foreign policy, interests and incentives to misrepresent

We define a state's interests as its value from achieving objectives.<sup>1</sup> From existing studies, we know that different assumptions about interests yield different predictions about behavior in international politics. Holding other inputs constant, researchers get different results if they assume states value either security (Waltz 1979), prosperity (Gilpin 1983), status (Renshon 2016), or political survival (Mesquita et al. 2005). Given how sensitive theoretical predictions are (Jackson and Morelli 2011), we permit states to hold complex underlying interests.

We define a foreign policy as how a state configures its diplomatic, military and commercial capabilities to best achieve its interests. There are often several plausible foreign policy options that leaders can choose from and states are uncertain about the consequences of their choices. We do not assume decision-makers know the best foreign policy. Rather, we assume that decision-makers choose a foreign policy based on their personal beliefs about the effects of different options and their understanding of the constraints they face (Chiozza and Choi 2003).

For example, in 1989 American policy-makers argued about America's best foreign policy for the post-Cold War world. Some argued that the United States should expand its military presence in Europe and retain Cold War era levels of military spending. These experts believed that if the United States withdrew, smaller European powers would fight for dominance in Europe and ultimately challenge the United States. Therefore, the United States should maintained a strong, forward deployed force that prevented the rise of competitors (Mearsheimer 1990). Others argued that the United States should largely withdraw from intense meddling in the affairs of others and instead dedicate its resources to domestic production and trade relations. These experts believed that extensive bilateral trade created incentives for states to cooperate with the United States. Further, the defensive advantage of nuclear weapons meant that the United States could deter critical threats. As a result,

---

<sup>1</sup>Interests are specific to each state and not relative levels of contestation between two states (Bailey et al. 2017).

an expensive forward deployed conventional force was a waste of valuable resources (Everb 1990). A third group argued that the United States should develop global institutions to promote liberal values world-wide. They believed that institutions could set standards of behavior to preserve American values. In their view, many states, especially emerging economies in Asia, who may face incentives to compete with the United States could be persuaded to support an American-led world order once they saw the advantages of joining these institutions (Keohane and Martin 1995).

All these experts shared a common goal: maximize American prosperity and security. Yet they each argued for different foreign policies because they held different assumptions about the efficacy of different foreign policies. Our theory accepts that a rational president could have plausibly adopted any of the these foreign policies. But the policy they chose would depend on the assumptions they held about causes and effects in international relations. We assume that governments choose their foreign policies based on their underlying set of beliefs about how the world works and that guides their general approach to foreign policy (Saunders 2011).

A leader's chosen foreign policy dictates its rivals and allies, and which counterparts are of little consequence. Leaders want to identify which adversaries adopt policies that carry a high risk of competition. For example, the United States and China had several competing interests in 1990. Yet the extent of competition between them depended on whether the United States focused on Asia or Europe, and whether the United States first turned to coercive threats, interdependence or institutions to exert influence. The Chinese government monitored American foreign policy choices closely in an effort to understand how much competition they would expect. Ultimately, their own policy hinged on their beliefs about what the Americans would do (Chen 1993).

## 2 Why it's hard to identify foreign policy shifts

Policy-makers find it difficult to identify changes in other states' foreign policies because states make foreign policy choices in dynamic, highly complex environments. This presents two challenges for analysts that try to identify when others change their foreign policies. First, depending on the context, foreign policy changes sometimes invite more cooperation and other times more competition. As a result, a state's incentives to misrepresent varies from case to case. In some cases, states cannot convince their rivals that they have changed a long-standing policy, but other times they try to deny that their policy has changed at all. There is no simple rule to apply across cases about how to understand a state's incentives to misrepresent, and how those incentives may be overcome. As a result, it is difficult to not only trust governments that promise their foreign policy has changed, but also develop a standard way to evaluate the potential truth based on their claims.

Second, most observable behaviors are highly contextualized. It is difficult to draw consistent inferences across cases from observing the same type of behavior. Two states may deploy entirely different capabilities to achieve the same objective. For example, the United States and Russia would deploy different economic sanctions and military coercive capabilities to convince Iran to end its nuclear weapons program. Thus, we cannot rely on the same pattern of military posturing or sanctions legislation in Russia and the United States to understand if they are serious about ending Iran's nuclear program. Similarly, two states may develop the same capabilities to achieve very different objectives. For example, increased military investment sometimes signals a commitment to defense, and other times signals a more aggressive foreign policy. As a result, it is difficult for analysts to draw the same inferences from similar patterns of a states' military spending, alliance choices, institutional affiliation or other observable behaviors.

One reason it is so difficult to rely on observable indicators is that states that want to conceal their foreign policies also conceal behaviors that signal their policy choice. Through

history, states amassed offensive weapons systems, developed nuclear weapons, or signed aggression pacts in secret.<sup>2</sup> They also have used covert action to meddle in the affairs of others without detection (O'Rourke 2018). States will often forgo the most obvious signals of a nefarious policy choice until the last moment or offer a pretext for their behaviors. But secrecy comes at the cost of efficiency (Joseph and Poznansky 2018). States face a trade-off between advancing their preferred policy choice and risking exposure. At least early on, states will try to avoid exposure while they lay the foundations for what is to come.

Some argue that sudden leadership changes brought about by regime change are a strong signal of foreign policy change (Goemans et al. 2009; Saunders 2011; Byman and Pollack 2001). As a result, leadership change might well predict foreign policy change. But, as we shall see, states often have large bureaucracies that determine their foreign policy. Even leaders that want to pursue a new foreign policy face barriers from their military and diplomatic advisers. Furthermore, foreign policies often change in the middle of an administration. As such, relying on leadership change would not adequately capture these changes.

In these complex settings a state's incentives to misrepresent varies from case to case. In some cases, states cannot convince their rivals that they have changed a long-standing policy, but other times they try to deny that their policy has changed at all. States often pursue a variety of different foreign policy behaviors that are difficult to interpret and masked by attempts to conceal their true foreign policy.

The most consequential foreign policy choices in the last decade have been so difficult to resolve because experts could not evaluate when their rivals had changed their foreign policies. In this section we describe four different countries that each experienced critical junctures in their foreign policy choices over the last ten years. These cases demonstrate the enormous challenge analysts face in understanding the foreign policy choices of their rivals.

---

<sup>2</sup>Examples include India's nuclear program (Thomas 1986) and Hitler's conventional arms build-up (Wark 1985).

The cases illustrate enormous variation in each state's incentives to misrepresent, the types of indicators that analysts pointed to post-hoc also vary across cases. Later we will apply our quantitative analysis of diplomatic events to these cases to demonstrate that our model identifies changes that are consistent with expert accounts of these cases.

## **2.1 Did Iran Reform Following the Joint Comprehensive Plan of Action?**

In 2011, the United Nations imposed tough sanctions against Iran for its nuclear weapons program devastating Iran's economy. In 2012, the newly elected Iranian president, Hassan Rouhani, committed to exchange Iran's nuclear weapons program for easing sanctions. Over the next three years Rouhani and foreign minister Mohammad Zarif negotiated the terms of this exchange with members of the European Union, China, Russia and the United States (known as the P5+1). In April 2015, the parties agreed on the Joint Comprehensive Plan of Action (JCPOA). Iran would reduce its centrifuge capacity, cede its enriched uranium stockpile, close several enrichment facilities, and accept aggressive IAEA monitors. The P5+1 would end nuclear related sanctions against Iran once Iran met some benchmarks.

For the United States, the decisions to both engage in negotiations for sanctions relief; and sign the JCPOA hinged on one question: was Iran really going to end their nuclear weapons program ([Katzman 2015](#))? It was possible that years of sanctions had convinced Iran that the pursuit of these weapons was not worth the cost. But it was also possible that Iran would restart its nuclear program in secret as soon as sanctions were lifted ([Hadley 2012](#)). Several foreign policy experts warned against engaging Iran in the first place. In 2012, they noted that Rouhani had long been involved in Iran's nuclear program and even served as its chief nuclear negotiator for 6 years. During that time, Rouhani had helped Iran evade international sanction while secretly developing its nuclear program. Given his history, many worried that he would exploit negotiations to gather legitimacy. As a result, the Obama administration wanted to know: was Rouhani different from his predecessors ([Parsi 2012](#))? Or was this just another ploy to survive under sanction? These concerns led

Secretary of State John Kerry to engage Iran slowly. Yet after frequent diplomatic meetings with the Iranian leadership between 2012 and 2014, Kerry eventually developed sufficient trust to discuss, and ultimately negotiate the JCPOA in April 2015.

Several foreign policy experts urged the United States Congress to block the deal.<sup>3</sup> They warned Iran could pursue its nuclear weapons program in secret or re-start it when Iran's economy had re-developed. Iran had strong incentives to pretend to change its foreign policy to end sanctions even if it never intended to abide by its promises. Furthermore, it was difficult to secure Chinese, Russian and European support for sanctions in 2011. These states were unlikely to agree to sanctions again. Thus, Iran could pretend to end their nuclear program until sanctions were lifted then re-start that program shortly after with impunity.

Those that supported the JCPOA noted that ending Iran's nuclear program before Iran acquired weapons would dramatically serve American interests ([Allison 2015a; Pillar 2015](#)). If Iran followed the JCPOA it would prevent nuclear proliferation across the Middle East ([Allison 2015b](#)). Further, by signing the JCPOA, Iran and the United States could build trust, this trust would ease Iranian concerns about threats from the United States. As a result, the JCPOA would pave the way for cooperation with Iran on other issues ([Gordon and Singer 2015](#)).

Despite fierce criticism, the JCPOA passed Congressional review without disapproval in September 2015 leading to a confirmed agreement in October 2015. In January 2016, the IAEA confirmed that Iran met the preliminary terms of the JCPOA and sanctions were lifted. For the next two years, the world watched to see if Iran had altered its foreign policy or if Iran would cheat on the JCPOA.

Several events between 2016 and 2017 signaled that Iran had not changed its foreign policy. Iran did not reduce its heave water stocks to the agreed upon levels in either July or September 2016.<sup>4</sup> In March 2016 Iran tested two medium range ballistic missiles. Iran

---

<sup>3</sup>See: [McCarthy \(2015\)](#); [Royce \(2015\)](#); [Rubin \(2015\)](#)

<sup>4</sup>It did reduce its stocks, but not fast enough leading some to argue Iran was dragging its heels ([Albright](#)

correctly observed that these tests were technically not prohibited by the JCPOA. Yet they prompted outrage from American and European policy-makers (including United States Ambassador Powers). In January 2017 Iran tested different mid-range missiles in clear violation of UN Resolutions. These tests prompted the National Security Advisor to put Iran ‘on notice’ ([DeYoung 2017](#)). Outside of the nuclear realm, Iranian foreign policy was even more concerning. Iran re-started conventional weapons transfers to Hezbollah, the Assad regime in Syria and Shiite militias in Iraq and Yemen. Iran increased its stranglehold over Iraq and Syria and used covert action to expand its influence throughout the Middle East ([Einhorn 2016](#)). As a result of these events, some American policy-makers called on the United States to scuttle the deal and re-start sanctions against Iran.

By mid 2017, concrete evidence emerged that Iran had actually abide by the terms of the JCPOA. The IAEA reported that Iran met all their requests for inspection and met almost all of their benchmarks ([Morello 2017](#)).<sup>5</sup> The IAEA confirmed that Iran dismantled major nuclear reactors and sold off many of the key materials required for nuclear weapons development. Experts now agree that Iran did change its foreign policy following the JCPOA: Iran gave up its nuclear ambitions and sought to re-integrate into the global economy. However, they also note that Iran has retained other aspects of its foreign policy that work against American interests ([Gordon and Nephew 2017](#)).<sup>6</sup>

The question is: were there signs of these changes before IAEA inspections in mid-2017? Even though the United States avoided conflict with Iran in 2016 and 2017, it could have easily gone the other way because Iran’s behavior did not clearly signal one way or the other.<sup>7</sup>

---

and Stricker 2016).

<sup>5</sup>It reported Iran barely missed those it did not make.

<sup>6</sup>2017 debates about new sanctions against Iran point to Iran’s behavior outside issues.

<sup>7</sup>For a survey of different expert opinions one year after the Iran deal see: [Einhorn et al. \(16\)](#).

## 2.2 Could NATO Allies Have Foreseen Russia’s Invasion of the Ukraine

With the fall of the Soviet Union Russian-European relations began to improve. By 2012, Europe accounted for 50% of Russian exports, Europe had reduced immigration controls on Russian citizens and awarded Russia observer status in NATO. After signing a modernization treaty that would ensure technology transfers to Russia in 2010, the EU assessed that, “The European Union and Russia, as long-standing strategic partners in a changing multi polar world, are committed to working together to address common challenges with a balanced and result-oriented approach, based on democracy and the rule of law, both at the national and international level” ([EU 2010](#)).

This is not to say that Russian and European interests were completely aligned. Russia had long declared that it would impose economic and political influence over many former Soviet states along its borders. Russia supported pro-Russian parties in this region and sometimes backed those parties with force. In 2008, Russia intervened in a conflict between the Georgian government and pro-Russian separatists To protect the separatists Russia occupied Ossetia for four months. But even in this case, the Russian intervention responded to the advancing Georgian military within the bounds of Russia’s declared interests. Further, Russia withdrew forces once the separatists had been secured. Despite these incidents the Europeans assessed that Russia would usually pursue its interests in a manner consistent with international law with only restrained uses of force in response to instability along its borders ([EU 2010, 2009, 2005](#)).

The Russian intervention in the Ukraine “challenged the European security order at its core” ([Mogherini 2016](#)) because it represented a fundamental shift in Russian foreign policy ([Hill 2016](#)). Once Putin annexed Crimea, it became clear that Russia was now willing to use force to take territory. In response, the Europeans and Americans imposed sanctions on Russia and bolstered their security ties. But this did not deter Putin. In 2015, Russia “surpris[ed] the world by entering the fray in Syria” ([Slim 2016](#)). The Russian intervention in Syria marked a second change in Russian foreign policy because it was Russia’s first major

military deployment outside the former Soviet Union ([Pezard et al. 2017](#)).

Putin did not invade the Ukraine on a whim. The NATO allies have “been on a collision course with Russia for a very long time, and have failed to acknowledge it until Putin forced us to do so in 2014 when Moscow annexed Crimea” ([Hill 2016](#)). Europeans now ‘widely’ believe that the “cause of Russia’s recent assertive international behavior is to be found in the need for Vladimir Putin to restore the legitimacy of the regime after the demonstrations that took place in Russia following the 2011 elections and to hide the government’s inability to reform the country’s economy and institutions” ([EU 2017](#)).

Of course, Putin had strong incentives to conceal his invasion plans for as long as possible. He knew that if his plans were discovered the Europeans would have forward deployed forces and imposed sanctions. But a foreign intervention is an enormous undertaking and Russia needed to prepare for it. In the aftermath, analysts uncovered several indicators that they could have used to identify Putin’s shifting strategy. For example, in 2012, Russia started a program to train urban Guerrilla units appropriate for taking and holding city centers ([Sherr 2014](#)). These units, deployed into Ukraine, were not appropriate for defensive operations In 2012 and 2013, Finland reported unauthorized submarine exercises in their waters that they strongly suspected were Russian. In 2012, Russian military planes conducted open air exercises in Swedish airspace following a minor diplomatic dispute. At the time, Russia justified these fly-overs as a response to a diplomatic incident, but in hindsight, he may have been testing Russian capabilities, or using the exercises to gauge Russian domestic support for foreign military activities ([Kofman et al. 2017; Cozad 2018](#)).

Russia has an extraordinarily complex foreign policy agenda. It was difficult for analysts in 2012 to understand the implications of a few indicators given Putin’s efforts to hide his shifting foreign policy ([Cozad 2018](#)). But Putin couldn’t hide everything. Ideally, there would be some trend in the data that analysts could have used to forewarn the invasion of the Ukraine. The question is: could analysts have foreseen the Russian invasion of the Ukraine in 2012, or the extension into the Middle East in 2015?

### 2.3 Did president Obama Really Pivot to Asia?

In 2011 Secretary of State Clinton (2011) declared that “the United States stands at a pivot point... One of the most important tasks of American statecraft over the next decade will therefore be to lock in a substantially increased investment—diplomatic, economic, strategic, and otherwise—in the Asia-Pacific region.” President Obama envisioned that the Pivot to Asia would involve “strengthening bilateral security alliances; deepening our working relationships with emerging powers, including with China; engaging with regional multilateral institutions; expanding trade and investment; forging a broad-based military presence; and advancing democracy and human rights” (Clinton 2011). To implement this policy, Obama needed Asian governments to accede to new institutions, facilitate America’s regional military presence through the construction of bases, promote American values and engage the American economy. While the United States had some long-standing allies in East Asia, it lacked key partnerships in South Asia, South East Asia and the Pacific. Obama correctly believed that governments across Asia would only embrace the Pivot to Asia if they believed that the United States was committed to their long-term security and prosperity (The White House 2011; Bush 2012).

However, Asian experts were skeptical that Obama’s foreign policy was any different from American policy in the half-century prior (Lieberthal 2011). The United States was still deeply committed in the Middle East and Europe. Many wondered how the United States could focus on Asia, Europe and the Middle East simultaneously. Even if Obama did want to pivot to Asia, it was unclear if he could overcome a well-established bureaucracy focused on the Middle East and Europe. Further, the last decade of United States policy in Asia was focused on the rise of China specifically rather than an interest in engaging Asia broadly. Skeptics worried that the Pivot to Asia was designed to contain China but otherwise ignore engaging Asia (James 2012). If true, the United States would not seek to grow Asian prosperity, but rather stifle China’s growth.

Thus, the critical question for Asian states in 2012 was: had Obama really Pivoted to

Asia? Obama may have held a genuine interest in re-focusing United States foreign policy in Asia and the Pacific. But it was just as likely that Obama really wanted Asian states to support a United States policy designed to contain China and otherwise ignore key Asian issues.

During 2011 and 2012, Obama took several steps to convince the world that American foreign policy would turn sharply towards Asia. By the middle of 2012, the United States had joined the East Asia Summit and secured progress in negotiations to form the Trans-Pacific Economic Partnership (TPP) ([Green 2016](#)). The United States strengthened its security commitment to Japan, South Korea and Australia by seeking new naval deployments ([Donilon 2011](#)). The United States also joined Vietnam and the Philippines in military exercises for the first time and increased investments in PACOM even as the financial crisis forced cut-backs in other areas of defense spending ([Dale and Towell 2012](#)). These behaviors carry all the hallmarks of costly signals in international relations theory. Obama used considerable American resources to verify American security and commercial commitments.

Furthermore, the president repeatedly promised that he had Pivoted to Asia, even claiming to be the first Pacific president ([The White House 2011](#)). During these speeches, Obama argued that his diplomacy verified that he was serious about the Pivot. In 2012, Obama argued that the best evidence that the Pivot was sincere was that he had visited Asia more times than any other president in the first two years in office, and visited Asia earlier than any other president. He also pointed to Secretary Clinton's extensive travel to the region. To the extent that audience costs can increase credibility, these repeated claims should have impacted beliefs about Obama's policy. At least in 2011 and 2012, analysts that relied on new military and commercial agreements, or even the president's account of his diplomatic efforts could have genuinely believed that Obama was Pivoting to Asia.

Taking stock of Obama's presidency analysts generally agree that the United States did not pivot to Asia.<sup>8</sup> "In reality, the policy was less of a sharp turn than the jazzy

---

<sup>8</sup>For a bi-partisan analysis written by former Obama and Bush Officials see [Green et al. \(2016\)](#).

word pivot implies and more of an astute shifts to follow global winds” (Putz and Tiezzi 2016). Foreign policy experts critical of the Administration argued that the president did not successfully implement a single major initiative in Asia. The Trans-Pacific Partnership (TPP) and many high-profile military were never finalized and ultimately fell through (Cha 2016). The president was unable to resolve repeated North Korean Crises (Ford 2017). Supporters of the president argued that while the president opened up new relations in Asia (Green 2016), the Administration main accomplishments were in Europe and the Middle East (Putz and Tiezzi 2016). They point to the JCPOA—a deal with Iran—as the largest successes of Obama’s foreign policy (O’Hanlon 2013). In 2015, even the president changed how he described his foreign policy. He no longer described it as the Pivot to Asia. Instead, he used the phrase a re-balance towards East Asia. This change in his rhetoric reflects that Asia was not the focus of his policy. Rather, his policy balanced interests in Asia, Europe and the Middle East.

Given so many new initiatives and high profile events in Asia it is unclear if even the best analyst could have foreseen the outcome. Obama did dramatically increase United States engagements in Asia across diplomacy, military and economic portfolios. However, he never reduced these efforts in other regions (James 2012; LaFranchi 2013). He simply had a more active foreign policy than prior presidents.

#### **2.4 Did irregular regime change in Australia drive foreign policy change?**

Over the last three decades, Australian foreign policy has walked a fine line between China and the United States. Rather than take sides, Australia’s policy has been to demonstrate their worth to both major powers by striking deals with both. However, “above all, Australians do not want to be put in the position where they have to choose between the U.S. as security guarantor and China as the crucial contributor to national prosperity” (Oliver 2015).<sup>9</sup>

---

<sup>9</sup>China makes up 27.9% of Australia’s total trade volumes.

In 2013 two debates among Australian policy-makers risked unsettling this delicate balance. Australians started to seriously debate a Chinese-Australian Free Trade Agreement (ChAfta). They also debated whether to end support for United States operations in Iraq and Afghanistan. Taken together, these events might have signaled a move towards China. Given Australia's long history of standing between the United States and China, these issues would have gone unnoticed by international audiences. Yet a series of irregular leadership changes, combined with the personal views of prime minister Turnbull, called into question broader changes in Australia's foreign policy.

Between 2010 and 2015 Australia had five different prime ministers but only one election. These changes arose from political infighting within the two major Australian political parties.<sup>10</sup> In 2010, the Labor party voted to oust the sitting prime minister Kevin Rudd and replaced him with Julia Gillard as head of party. Thereby making Gillard the prime minister. Three years later, Rudd convinced the party to oust Gillard and return him to power. Shocked by these events, which had only happened twice before in Australia's 200 year political history, Rudd lost the 2013 election to the Liberal Party leader Tony Abbott. Tony Abbott was then ousted by his own party and replaced with Malcolm Turnbull in 2015. These events exposed deep divides within Australian politics and "reflected the increasing tension between domestic and foreign policy interests. This trend was already evident, yet it was exacerbated by the challenge to his leadership of the Liberal Party" ([Simpson 2015](#)).

Under Malcolm Turnbull's leadership Australia-watchers were much more uncertain about Australia's foreign policy course. Analysts wanted to know: would Turnbull tilt Australia's foreign policy towards China at the expense of their relations with the United States? Turnbull claimed he would continue with Australia's long-standing policy of balancing relations with China and the United States. However, analysts were skeptical. Before taking office, Turnbull had made several pro-China comments ([McDonald 2015](#)). During his first

---

<sup>10</sup>In Australian politics, the leader of the party with the most congressional seats serves as the prime minister.

year in office, he passed a version of ChAfta that included controversial investor-state arbitration provisions that previous prime ministers had resisted ([Nottage 2016](#)). He also made appeals to president Obama to push the TPP through a lame duck congress (consistent with China’s interests). Given so many pro-China policies in his first year, analysts noted that “Turnbull risks unsettling some of Australia’s most important relationships and unnerving his domestic constituency. If he is too pro-China, this will antagonize the United States as Australia’s alliance partner” ([Oliver 2015](#)).

Later in his tenure, Turnbull took considerable steps to strengthen Australia’s ties to the United States ([Scott 2017](#)). In his second year in office, Turnbull retained Australian support for United States operations in the Middle East. He facilitated new basing arrangements for the United States air force in Australia and sent Australian forces to participate in military exercises in the South China Sea. Turnbull has fiercely criticized China’s human rights policy and supported United States’ efforts to prevent China’s expansion in the South China Sea.

Following the release of Australia’s Defense White Paper (2017), analysts agreed that Turnbull had in fact stayed the course ([Scott 2017](#)). But it was not obvious in his first few years in office. The turbulent domestic situation combined with Turnbull’s unique personal style caused some concern. These concerns were exacerbated by Turnbull’s personal preferences for pro-China economic policy.

## **2.5 Summary of the problem**

These cases demonstrate the difficulty analysts face in locating shifts in foreign policies as they are happening. As Table 1 shows, there is enormous variation in incentives that different governments face to represent their policies. Iran and the United States faced incentives to convince others that they had changed their policies, while Russia faced incentives to claim their policies had not changed. With hindsight, analysts pointed to specific indicators of policy change in each case. However, the indicators across cases were markedly different.

Table 1: What they said and what they did

	Foreign policy changed	Foreign policy did not change
Promised change	Iran: Exchanged nuclear weapons for economic integration (2015)	USA: Did not pivot sharply to Asia
Promised continuity	Russia: Started militarization (2012) in prep for Ukraine. Then expanded to Middle Eat (2015).	Australia: Remained balanced between USA and China

For example, Iran sent subtle signals that it had complied with the JCPOA. It sold off critical technologies and was open to inspectors, yet Iran also tested missiles on contravention of other UN Resolutions, and supported non-state violent actors to expand its influence through Iraq, Syria, and Lebanon. Evidence of Russia's shifting policy came in the form of newly formed military united and a few crises that could easily be explained in the pretext of security. However, these signals were masked by extensive commercial and military cooperation with European partners. It is little wonder that analysts did not foresee the Russian invasion of the Ukraine. The United States too sent a variety of mixed signals. President Obama tried hard to signal his strong commitment to Asia through scheduling summits and a number of public speeches declaring his commitment to the policy. Many analysts believed that Obama had refocused American foreign policy on Asia.

In what follows, we develop a method to locate shifts in these state's policies as they are happening. We argue that this single method accurately reports changes in state foreign policies across all these cases even though they each contain incredibly different contexts. The central insight is that no matter the context, leaders that embark on new foreign policies must achieve new objectives. To do that, they need to establish new agreements with new counterparts.

### 3 Diplomacy and Foreign Policy

We argue that patterns in a state's diplomatic choices vary with its foreign policy. We detect changes in a state's foreign policy by searching for sudden breaks in the pattern of its

diplomatic choices. First, we explain why breaks in the pattern of diplomatic events covary with changes in foreign policy. Second, we argue that structural break tests are the best way to detect these changes and rule out alternative modeling strategies.

### **3.1 Why are breaks in diplomacy linked to new foreign policies?**

To develop a theory about the overall pattern of diplomatic meetings a leader selects we start with existing research on the micro-foundations of diplomacy. This research has shown how bilateral meetings can help a leader use specific diplomatic encounters to achieve specific objectives. Scholars have shown that leaders use private diplomacy to overcome coordination problems and facilitate efficient commercial and security interactions between them. Leaders can communicate their resolve (Trager 2010; Sartori 2005; Kurizaki 2007) reducing the hazard of conflict, negotiate new agreements that lead to the efficient division of contested goods (Avenhaus and Zartman 2007; Ramsay 2011; Putnam 1988) and mediate disputes (Kydd 2003). In fact, private, high-level diplomacy is more effective at coordinating than either public statements, neutral third parties or interlocutors (Fey and Ramsay 2010). Foreign visits also signal to third parties an intent to honor an alliance commitment (McManus 2018) or a leader's conformity to norms (Whitehead 2006). High-level meetings can help leaders signal their type to domestic audiences (Howell and Pevehouse 2007) or develop status as a world power (Rathbun 2014).

To be clear, we do not think leaders leave every meeting with a newly signed commercial or security agreement. However, high level meetings are the first step to negotiating cooperative agreements between two states (Plouffe and van der Sterren 2016). Leaders use personal interactions to identify their most pressing issues and resolve the main barriers to reaching agreements. Over the following weeks and months their staff negotiate the specific details and put new policies into place. In these studies, high-level diplomacy facilitates broader cooperation between two states.

These studies emphasize the effects of specific diplomatic meetings on the countries

that attend the meetings. Obviously, leaders can only coordinate with those who they speak to. How much a state can benefit depends on the gains they can achieve from cooperating with each specific partner. When two states have few contested issues (or compatible factors of production) the gains from increased security (or commercial) coordination are small.

Furthermore, the decision to engage in diplomacy costs leaders the opportunity to pursue other objectives. A meeting between heads of state takes weeks out of a leader's schedule to prepare for, travel to and attend. In that time, a leader could have solved a different coordination problem with a different counterpart or championed a domestic issue. As a result, leaders are unlikely to meet with everybody. Instead, they will focus their attention where diplomacy can overcome coordination problems leading to the largest gains.

Leaders will weigh the costs and benefits of diplomatic meetings and focus their attention where they get the most gain. When a leader wants to forward deployed her military, she is likely to meet with states where she wants to establish military bases. When a leader wants to promote norms, she is likely to seek out partners with whom she can establish these norms, and violators to convince them to change their behavior.

**Hypothesis 1** *Leaders will select diplomatic partners that help them achieve their specific foreign policies.*

These predictions clarify the results of a recent wave of diplomatic events data that researchers to better understand the predictors of leaders' diplomatic choices. We now know that foreign visits are well predicted by commercial and security ties, as well as the significance of partners (measured in terms of GDP, military capabilities and status) (Lebovic and Saunders 2016; Kinne 2014; Plouffe and van der Sterren 2016; Neumayer 2008). From these trends researchers infer that patterns of diplomatic behavior correspond broadly with states' underlying interests.

However, two features of these studies are often overlooked. First, variables that measure policy divergence also predict diplomatic activity. For example, states recently

engaged in conflict are more likely to visit each other or establish embassies. We intuit that states want to avoid war and so face strong incentives to coordinate with their rivals. One way to interpret this evidence is that states do not meet with others based on the alignment of their interests. Rather, states have specific policy objectives—e.g. avoid war, foster commercial agreements, strengthen alliance partnerships—that they want to achieve. They choose diplomatic partners to serve these specific objectives.

Second, the predictors of diplomatic behavior are not constant across cases or within one case across time. In some cases, measures of commercial ties are the strongest predictors of diplomatic visits, in other cases variables that measure security relationships are stronger. Different still, some states seem to focus their diplomatic efforts on a single region, whereas others spread their efforts globally. We argue that this variation is based on a state's foreign policy choice. When foreign policies change, the relative importance of different predictors will also change.

New foreign policies should trigger a sudden shift in how the leader chooses its partners, rather than a gradual shift, for two reasons. First, the decision to meet costs leaders the opportunity to pursue other objectives. A meeting between heads of state takes weeks out of a leader's schedule to prepare for, travel to and attend. In that time, a leader could have solved a different coordination problem with a different counterpart or championed a domestic issue. As a result, leaders are unlikely to meet with everybody. Instead, they will focus their attention where diplomacy can overcome coordination problems leading to the largest gains. New foreign policy priorities imply that the benefit from meeting partners that served the old agenda is dramatically reduced and the gains from new partners is enhanced.

Second, diplomacy has the largest impact when states are most uncertain about each other and share few existing agreements. Two states with a sparse history of interactions face larger barriers to coordination. Diplomacy can facilitate trust and identify common goals between them. Thus, early encounters can have a large impact as there are more areas to coordinate on. When states embark on a new foreign policy, they will identify counterparts

that are all of a sudden relevant. In the past, the benefit from formalized agreements with these partners was small to the point that the benefits of coordinating failed to outweigh the cost. However, once the gains are sufficiently high, coordination is valuable. When this is the case, leaders should prioritize counterparts where the coordination problem is largest as their meeting can have the largest impact.

**Hypothesis 2** *When leaders change their foreign policies, they will change the basis for which they select diplomatic partners.*

Since some leaders face incentives to conceal their foreign policies they also face incentives to alter their diplomatic behavior because it provides information about their foreign policy choices. In particular, leaders will avoid controversial visits that cause third-parties to worry they have embarked on a nefarious foreign policy. These leaders will avoid meetings to avoid backlash from the international community.

But it is difficult for states to avoid diplomatic encounters that serve their foreign policies because these meetings are important. In particular, third parties know that leaders choose partners to facilitate their foreign policies. So when one leader meets with a partner that likely serves a specific foreign policy choice, others infer that this leader is pursuing that policy.

Clever leaders will establish pretexts to meet counterparts. For example, two countries that are members of the same trade institution can use their shared affiliation as a pretext to meet. But once they are in private they can discuss military planning, clandestine technology transfers or other policies. While pretexts are useful for hiding the content of a specific meeting, they cannot conceal the overall pattern of meetings. There are only so many trade meetings that Iran can pretend to hold with Libya, North Korea and Russia before the rest of the world grows suspicious. Over a long enough time period, a complete analysis of all attended meetings should signal some information about a state's policy choice.<sup>11</sup>

---

<sup>11</sup>It is incredibly difficult in the modern world for a Cabinet Official to hold secret meetings with a foreign counterpart. Typically their whereabouts are tracked in great detail. Even if they could keep it secret from

As a result, leaders that want to conceal their foreign policies face a trade-off. They can attend meetings that help them achieve their chosen foreign policy and risk backlash. Alternatively, they can avoid these meetings making it more difficult that others will discover their aims and more difficult to achieve their chosen policy. As a result of these trade-offs leaders never completely mask their diplomatic behavior. Instead they mask it just enough such that third parties are sufficiently uncertain about their policy choices.

In practice leaders often forgo meetings that would advance their foreign policies. However, rarely do they pursue meetings that do not serve their interests. To completely mask the pattern of their diplomatic behavior, leaders would need to both avoid meetings that benefit them and attend meeting that did not benefit them. We believe that leaders may do the former, but almost never the latter. Analysts that focus on the significance of individual meetings are unable to detect these subtle changes. However, a statistical model equally weights changes towards meetings attended and meetings no longer attended. Thus, we believe that we can locate these changes even if leaders hope to conceal their policies.

### **3.2 An empirical method for identifying changes in diplomatic choices**

We develop an empirical approach to diplomatic events data by building on the work of [Lebovic and Saunders \(2016\)](#). We rely on a decision model where one state (for example the United States) chooses its diplomatic partners. We'll refer to that state as the target state. Each observation in the dataset represents a potential partner that the target state's elites could visit in a given year. We construct a time-series cross sectional data-structure with one observation for every country in every quarter of the year in the international system (omitting the target state).

The outcome variable is a count of the number of times the target state's leader met with that partner in a given quarter. The predictor variables capture all the reasons that the target may choose one particular partner state. We chose predictors that capture the

---

the media, it is likely that a foreign intelligence service would identify them.

two dimensions of foreign policy: regional focus; and preference between military, diplomatic and commercial foreign policy instruments. The basic model takes the form:

$$Meeting_{cq} = \beta_0 + \beta X_{cq} + Meeting_{cq-1} + \epsilon_{cq} \quad (1)$$

where subscripts refer to country quarters. The model accounts for data clustered across time with a lagged dependent variable. The coefficients report how important a specific factor was for the target in selecting partners. High coefficients do not relate to any single partner. Rather, they report the importance of properties of partners overall.

If our first hypothesis is correct, the coefficients in the model will remain constant over the periods that the target pursues the same foreign policy. For example, so long as Iran is trying to acquire nuclear weapons, it should meet with rogue states who will sell it the technologies and trade with it under sanction. The variable that measures rogue states should therefore remain positive and significant for that whole period. Furthermore, these coefficients should match expert assessments of that state's foreign policy. That is, we expect that the United States should meet with Middle Eastern and European partners more than African or South American partners. Further, Iran should meet more with potential trade partners once sanctions are lifted but not before. As part of our analysis we will qualitatively assess the coefficients of our models to make sure they match expert assessments.

If our second hypothesis is correct, the coefficients under one foreign policy will be different from another foreign policy. When one state no longer believes that supporting institutions and international law are good ways to achieve its objectives, then that state will no longer meet with states to bolster institutions. Instead it will seek new partners based on its new beliefs about how best to achieve its interests.

To locate these different foreign policy regimes, we employ a statistical technique known as structural break analysis (Durbin et al. 1975). The technique describes a class of non-parametric methods that identify points in a time series where the pattern of sequential events substantively differ from those that came before (see Bai and Perron 2003; Zeileis

2005, for review). Structural break tests were popularized in applied research by finance researchers who wanted to predict and understand financial shocks (Narayan et al. 2013; Hatemi-J 2008; Andreou and Ghysels 2002). These tools are increasingly used in political science (Greene 2008) to explain judicial regimes (Pang et al. 2012), human rights violations and treaty negotiations (Spirling 2012).

We use two versions of the test to observe changes in model fit with time series cross-sectional data. Both tests are sensitivity tests applied post-estimation to an underlying statistical model. In both cases, we use the model defined by equation 1 as the underlying model. First we use an M-Fluctuation test (Zeileis and Hornik 2007). The test yields a p-value score that can tell the researcher if there are significant fluctuations in the model fit across different periods.<sup>12</sup> The Null hypothesis is that autocorrelation exists between consecutive periods in the dataset. Rejecting the null implies that there is at least one time period where auto-correlation is inappropriate. The advantage of this test is that it does not require the user to specify a model parameter (such as the residuals, or R-squared value, or any specific coefficient) to evaluate fit. Rather, the estimate is a global test. The disadvantage of this test is that it does not explain if the model fit would be improved by partitioning the data (rather than identifying highly erratic data) or where one would partition the data to improve the model fit.<sup>13</sup>

We address these concerns with a second test. We group the residuals from our underlying statistical model by quarter and estimate the mean of these residuals. We then apply the test recommended by Bai and Perron (2003) to estimate the empirical fluctuation process in the mean of these residuals across time. This model considers every conceivable partition of the data and provides two pieces of information. First, it identifies the number of breaks that best fit the data (should you partition into 1, 2, or 3 periods). Then it identifies

---

<sup>12</sup>Please refer the Appendix A for detailed outline.

<sup>13</sup>Zeileis and Hornik (2007) recommends using the visual plots — which we provide in the Appendix. But there is much room for interpretation.

where the optimal break points are given a fixed number of break points.<sup>14</sup> It relies on two different model fit scores to determine the most appropriate number of break points, given the optimal partitions. Researchers can select the number of breaks that minimize the Bayesian Information Criterion (BIC) or they can observe the largest shift (termed the elbow) in the reduction of the residual sum of squares. We use this test to identify where and how many breaks occur. The test has two disadvantages. First, the researcher must specify the model parameter to test. We chose the mean residuals because it is most commonly chosen. Second, the test does not provide a confidence interval. Instead, it relies on information criterion and brute force to identify the best fitting partitions. Fortunately, both of these weaknesses are strengths of the M-Fluctuation test. Thus, we rely on the combination of break models to overcome the weaknesses of each.

Taken together, these tests identify whether the target state alters the types of diplomatic partners it chooses across time. If it does, the tests identify a specific break point (or many points) that demarcate the different periods. When the model identifies a break point, it implies that one set of coefficients well predict diplomatic visits in all the quarter-years before the break, yet a different set of coefficients predict diplomatic visits in all the quarter-years after the break.

Structural break tests help us avoid many challenges that the other measurement techniques must confront. We could have used other variables that researchers have theoretically linked to the causes of foreign policy change: leadership change, technological developments, changes in polarity or domestic political variables. But these theories are still developing and available measures imperfectly measure changes in foreign policy. In many cases, one leader will change their policy in the middle of their tenure, and at other times new leaders do not change their nation's course at all. We tried to develop such a model but could not find one that well predicted foreign policy changes accurately.

Similarly, we could have used changes in tools of foreign policy other than diplomatic

---

<sup>14</sup>This is a clear advantage over a Chow Test, which requires users to theorize about when breaks occur.

meetings such as military spending, trade ties or alliance choices. These measures suffer from two shortcomings. First they are slow moving. Second, it is unclear what the changes mean. Does increased military spending mean a state is doubling down on an existing policy or embarking on a new one? Since some foreign policies rely on certain instruments but not others, it was hard to find a model and data-source to identify change points. If our theory is correct, diplomacy plays a facilitative role in all foreign policy decisions. States use diplomacy to coordinate on military, institutional and commercial activities. Further, diplomatic events occur frequently enough to analyze with statistical techniques in small time intervals. Thus, the data is well suited to identify granular changes.

One disadvantage of our approach is that it will not identify gradual changes in a target state's foreign policy. In particular, we do not expect to find a change in the United States' diplomatic behavior following the Pivot to Asia. We are comfortable with this limitation for three reasons. First, we suspect that states make large and sudden changes in their foreign policies when they alter their beliefs about how best to achieve their interests. Thus, the emphasis is empirically relevant. Second, when sudden shifts do occur, they produce a heightened risk of conflict that a gradual change would not. New foreign policies drive states to compete over new issues. This leads states into crises over issues they have never contested before, or causes their forces to engage each other in new settings. When states do not have a large history of interactions between them, they are forced to make decisions with little understanding of their consequences. As a result, there is a larger chance of miscalculation that is compounded by the large number of new interactions.

Finally, dramatic changes in a state's foreign policy create incentives for targets to use war or regime change to prevent the policy change from taking place. These costly policies are worthwhile when changes are sudden rather than gradual. Targets have a small window to implement them. If it takes them too long to discover their rival's new policy, their rival would have already configured their capabilities rendering retaliatory policies less effective. Given these factors, we believe decision-makers are more interested in sudden changes than

gradual changes. These changes create a sense of urgency and leaders have limited time to respond.

### 3.3 Data

We analyze diplomatic patterns using more temporally granular information of state diplomatic meetings. To do this, we build a dataset of diplomatic events where the unit of analysis is an elite-meeting-day. Our strategy is to focus on a single elite at a time and construct a time series of meetings that leader attends. We focus on the diplomatic exchanges between heads of state (presidents, prime ministers, chancellors, amirs, kings, etc.) and foreign ministers to reconstruct a state’s ego-centric network of diplomacy over time. We collect data on the four cases outlined above: Australia, Iran, Russia and the United States.

To recover a complete record of elite diplomatic meetings, we exploit data that the elites themselves provide. Our sources are formal diplomatic websites that are updated and maintained by each administration. This offers a record of all self-reported diplomatic exchanges for an administration across its tenure. The websites record formal travel, meetings, and phone conversations between counterparts and other dignitaries, diplomatic gestures (e.g. expressions of condolences or congratulations), and news on domestic events and/or proclamations. In essence, these sites serve as a “news feed” for an administration, which provides relevant information on what a particular head of state does during his or her tenure.

We implement an original dictionary-guided machine coding algorithm to identify meetings that occurred between foreign policy elites. We do this by scanning through the associated text for each reference source to gather information on who, how, and when the reference dignitary meets with another dignitary. The algorithm systematically moves through the scraped text, parsing relevant material from the irrelevant. We define relevant entries as identified meetings or phone calls between two heads of states and/or foreign ministers. Irrelevant entries are defined as any domestic meetings, news and/or proclamations,

the signing of legislation, expressions of condolence or congratulations, and/or references to some *future* meeting. For the analysis at hand, we only consider in-person meetings.<sup>15</sup>

Given the frequency of diplomatic meetings in the countries we observe, we aggregate meetings to quarter-years. Our final dependent variable is a count of the number of times that the target met with a partner in that quarter of the year (Jan-Mar is the first quarter). Given our outcome is a count, we use a poisson link function rather than a logit to account for the count dependent variable. Further, we include a lagged dependent variable, which is the sum of all the meetings between the target and partner  $i$  in the last quarter.<sup>16</sup>

Our model also relies on a different data source for the right hand side variables. There is no reliable data that measures quarterly or annual military spending or GDP over such recent periods. Further, much of the data that covers our period has missing data on key counterparts that Iran and Russia meet with.<sup>17</sup> From a theoretical standpoint, we do not believe that elites think about precise fluctuations in their counterparts military spending or economic output month-to-month when they decide who to meet with. Most of the variation likely comes from selecting different partners based on a general understanding of their trade, security, and/or strategic potential generally or a shared history. Thus, we prioritized keeping the largest number of observations and use crudely measured predictor variables—noting that crudeness in measurement is worth the retention of key interactions.

In addition, two states might meet because they are tethered together by a shared history and culture, commercial and security interests, or geographic proximity. We account for these ties using a dummy variable for shared language, logged exports from the target to the counterpart (Direction of Trade Statistics), a dummy variable for a formal security agreement, and the logged distance between capital cities.

---

<sup>15</sup>For a more detailed description regarding the algorithm's construction and performance see Appendix C.

<sup>16</sup>We tested a number of other lag measures and found this to improve model fit the best.

<sup>17</sup>For example, GDP and military spending data is missing for Iran, North Korea, Cuba, and China. These observations would be lost if we included these covariates.

For each of the four countries we apply the following empirical strategy:

1. Estimate a model of diplomatic choice as outlined in Equation 1
2. Use the M-Fluctuation analysis to test the null hypothesis: are sufficiently consistent across periods to apply time series analysis. When the p-value < .05, we reject the Null hypothesis and move to test if the model fit can be improved by partitioning the data.
  - (a) We compute the mean residual for each time period of the baseline model.
  - (b) We use the [Bai and Perron \(2003\)](#) test to determine the optimal number and placement of breaks. If the test cannot find a break to improve the model fit, we conclude the state has erratic foreign policy, rather than two foreign policy regimes. If the test recommends break-points, we infer that at these point in time, the state's motives for choosing diplomatic partners has changed. We:
    - (c) Evaluate if those break-points match expert assessments of foreign policy change.
    - (d) Partition the data based on the breaks and re-estimate the base-line model.
3. We qualitatively assess if the models' coefficients match expert assessments of foreign policy objectives.

## 4 Quantitative Case Analysis

If our theory is correct, then we should detect breaks in the pattern of diplomatic behavior using a structural break model when states change their foreign policy and not otherwise. To validate our approach we collect and analyze diplomatic events data for four different countries: Australia, Iran, Russia, and the United States. As described below, these four countries face highly complex and different foreign policy settings. As Table 1 showed, they each faced very different incentives to misrepresent and where honest in some cases but not others. In what follows, we conducted quantitative case analyses to test our two hypotheses. First, we search from breaks in the data and then compare them to expert assessments about all four countries' foreign policy changes. Second, we analyze the model's coefficients to determine if the models well capture qualitative accounts of each states' foreign policy.

#### 4.1 Finding Breaks in the Data

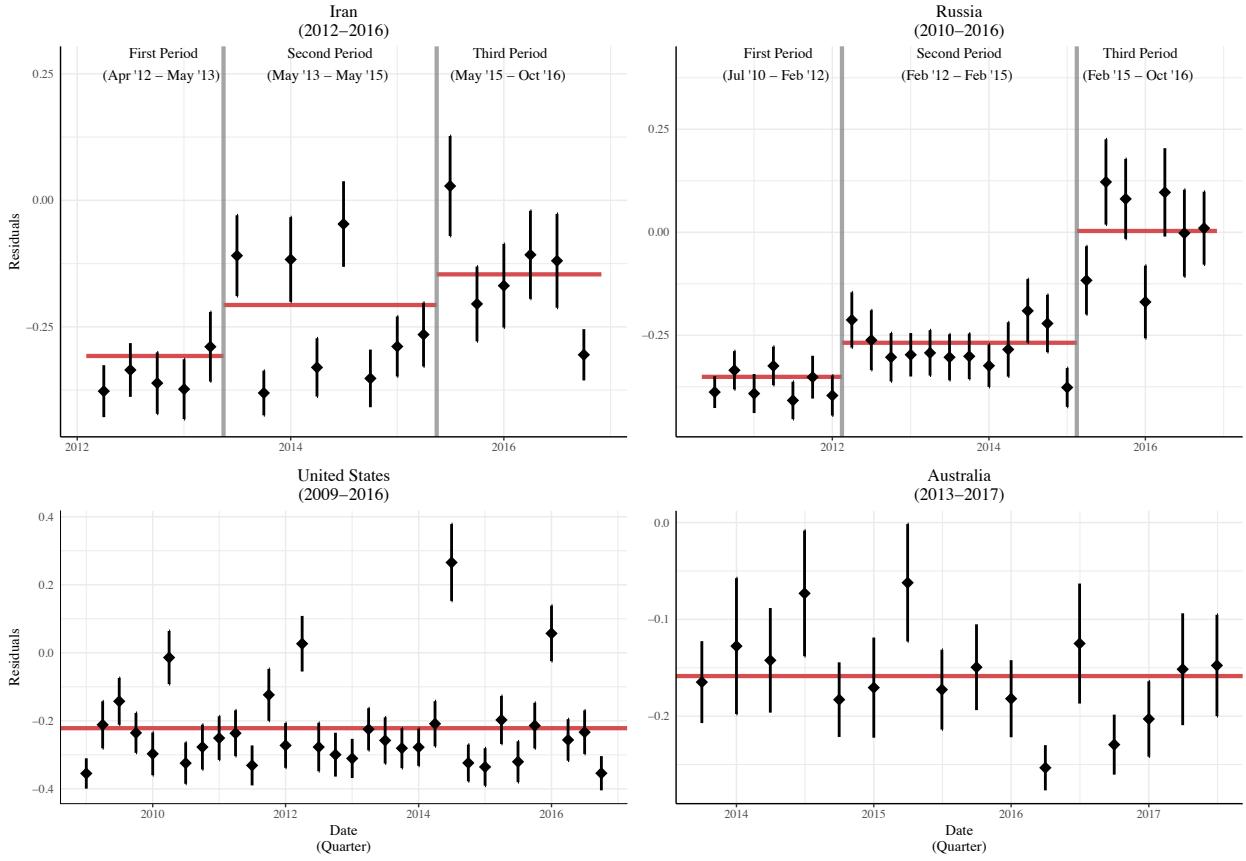
We make a different predictions for each of the four countries we collect data on. Experts assess that Iran and Russia changed their foreign policies. Thus, we expect a break in the model fit for Iranian diplomatic behavior *after* Rouhani comes to power then again *after* sanctions are lifted, in Russian behavior *before* the Ukraine Crisis and then again *before* Russia intervenes in Syria. Experts assess that Australia and the United States did not dramatically alter their foreign policies. As a result, we do not expect structural breaks in the patterns of their diplomatic behaviors even though Australia experienced a series of abnormal leadership changes and the United States tried hard to use diplomacy to convince the world it had altered its policy.

We use structural break models to determine if breaks are appropriate and, if they are, where the breaks are. We summarize the results of our structural break models for all four cases in Figure 1. Each panel represents a different country and includes two pieces of information. First, we plot the mean and confidence intervals of the underlying model's residuals over time. Second, we demarcate the periods of diplomatic change that coincide with the optimal breaks points recommended by the reduction in the residual sum of squares and Bayesian information criterion as vertical lines.<sup>18</sup>

---

<sup>18</sup>More detailed analyses of these tests is presented in Appendix B.

Figure 1: Changes in strategy



The p-values report the confidence level that a structural break is appropriate based on the M-Fluctuation test. The points capture the average residual of Poisson model and the black bars the variance in the residuals. The gray bands denote the location of structural breaks. The horizontal red lines describe the average residual for all quarters located within the relevant period.

Panel A reports the results for Iran. The p-value from the M-fluctuation test suggests that we should reject the Null hypothesis with 99.9% confidence. Thus, we conclude that the use of an autocorrelated linear model is inappropriate. Based on this result, we believe that a linear model is inappropriate. We compute the mean of the residuals from our model and apply our second structural break test to determine if the model fit can be improved by partitioning the data into foreign policy regimes. The residual sum of squares (RSS) test finds that partitioning the time series into two points will improve the model fit: one in the middle of 2012—right after Rouhani comes to power; and a second in the second quarter of 2015 — the period right after sanctions are lifted. Panel (A) plots the mean of the residuals from the underlying model across time. Over the same period, there were several

key domestic events that might cause leaders to divert their attention away from foreign policy—nation-wide protests (February 2011 - May 2012) and a financial crisis (Jan-June 2012). A model that predicts Iran’s diplomatic visits did not change as a result of these events.

Panel B reports the results for Russia. The p-value from the M-fluctuation test leads us to rejects the use of an autocorrelated linear model with 99% confidence. We infer that a linear model is inappropriate. Next we compute the mean of the residuals from our model and apply our second structural break test to determine if the model fit can be improved by partitioning the data into foreign policy regimes. We compute the mean of the residuals from our model and apply our second structural break test. Panel (B) plots the mean of the residuals from the underlying model across time. Based on the results of the residual sum of squares test we identify two breaks: the first quarter of 2012; and the second quarter of 2015.<sup>19</sup>

The first break we identify is consistent with expert assessments. Russia’s pattern of diplomatic behavior shifts about six months before the Ukraine crisis. The second break is more surprising to us. We did not anticipate such a strong shift in the model fit so long after sanctions. On reflection, we note that this corresponds with the implementation of strict sanctions against Russia and the onset of Russia’s intense aerial bombing of Syria. Consistent with expert assessments, this break coincides with Russia’s new global focus for military operations To support a conflict in Syria, Russia required logistic and supply lines through the Middle East. A new regional focus is, indeed, a new foreign policy choice for Russia.

Panel C reports the results for the United States. The p-value from the M-fluctuation test leads us to rejects the use of an autocorrelated linear model with 96% confidence. we computed the mean of the residuals from our model and apply our second structural break

---

<sup>19</sup>The BIC test suggests that one breaks is appropriate. However, the results suggest a negligible difference between one and two breaks. Further, the other tests confirm that a second break is appropriate.

test to determine if the model fit can be improved by partitioning the data into foreign policy regimes. However, both the BIC and RSS statistics recommended that the model fit could not be improved by partitioning the data into different regimes. Taken together, these tests suggest that there is some erratic behavior in Obama’s foreign policy but no strong evidence that he shifted from one coherent strategy to another.<sup>20</sup> These results confirm our expectation: the United States did not Pivot to Asia and instead tried to manage a complex policy focused on many different regions. This is somewhat of a tough test for our method. President Obama specifically pointed to extensive diplomatic visits to Asia as a signal that their policy had changed. A simple count of diplomatic meetings confirms that Obama and his foreign policy staff did visit Asia more than prior Administrations. But Obama also met with European and Middle Eastern partners extensively. Thus, our model did not detect a sharp break in America’s diplomatic behavior because it balanced all regional meetings. Instead, our results support the consensus of experts who look back at the Pivot, it was a slow turn towards Asia and not a sharp shift in American foreign policy.

Panel D reports the results for Australia. The p-value from the M-fluctuation test leads us to reject the use of an autocorrelated linear model with 75% confidence — well below the conventional 95%. The BIC and residual sum of squares tests similarly recommend no breaks. We conclude that there is no break in Australia’s foreign policy behavior over the period. Australia’s turbulent domestic politics during this time provide another tough test for our modeling strategy. Australia experienced several abnormal regime changes during the period we study. Prime minister Turnbull was said to hold a dramatically different ideology from his predecessor. These ideological differences could have driven different diplomatic styles, and thus different patterns of diplomatic behavior. Yet the model did not detect a change in the overall pattern of Australia’s diplomatic activity. This is consistent with the

---

<sup>20</sup>The BIC test suggested that the best place for a break was the second quarter of 2014 although including the break actually reduces model fit. This break coincides with the appointment John Kerry and not with the Pivot to Asia story.

expert assessment that Australia's foreign policy stayed the course under Malcolm Turnbull.

These results uniformly support our second hypothesis. The models locate structural changes in the primary features of a state's diplomatic choices at points where experts claim that these states change their foreign policies and at no other time. The tests successfully identify changes in foreign policy where we expect them to. They also identify erratic patterns in foreign policy across different periods where experts suggest a leader pursues multiple foreign policies.

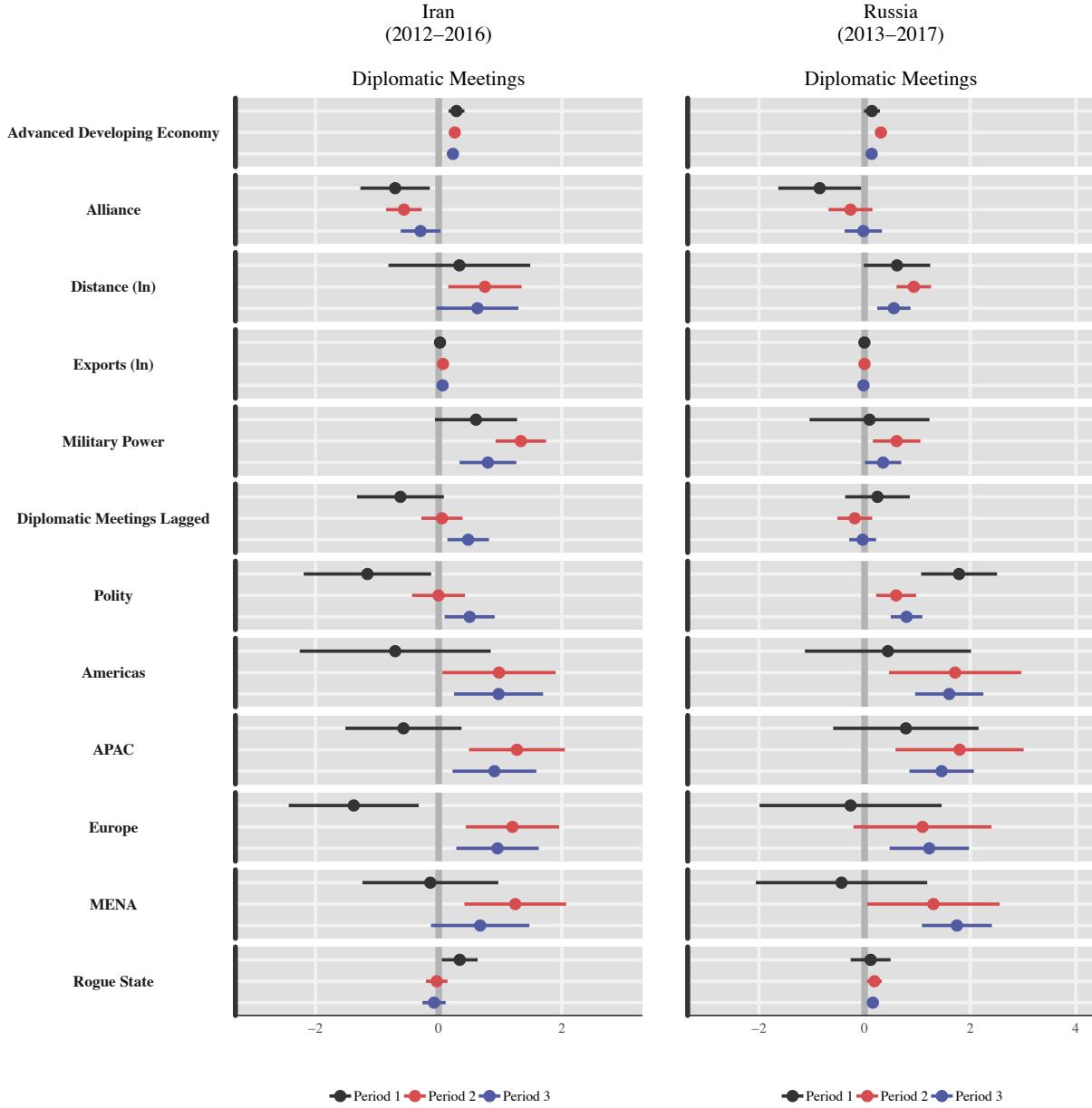
## 4.2 Qualitative Analysis of Coefficients

To determine if these breaks are consequences of foreign policy change, we qualitatively interpret the coefficients of the models. If we are correct, and diplomacy facilitates foreign policy, then the model coefficients should match expert assessments about what states are trying to achieve. For the two cases where we identified foreign policy changes (Iran and Russia) we want to understand what motivated diplomatic meetings each side of the breaks. To do that, we partition the data into periods based on the break-point we identified and re-run our baseline analysis on each partition. Figure 2 plots the coefficients from the model of Iran's behavior before Rouhani engaged the United States (black), during negotiations (red), and after the JCPOA was signed (black). Figure 2 plots the coefficients from the model of Russia's behavior for the three different periods we identified.<sup>21</sup>

---

<sup>21</sup>We use the method suggested by Paternoster et al. (1998) to compare regression coefficients across models. Strictly the method assumes independence of the samples. We think of this as a good rule of thumb to ensure consistency in our qualitative analysis, rather than a strict inferential test of differences.

Figure 2: Qualitative assessment of changing strategies



The plot describes the coefficients and 95% confidence intervals from a Poisson regression model defined in equation 1. The black coefficients plot the results from the first time period, red from the second, and blue from the third. Africa is the baseline category for the Regional Fixed Effects. Regional coefficients are relative to Africa.

We found one break in Iran's diplomatic behavior when Rouhani first came to power, we found another break right after Iran signed the JCPOA. With hindsight, experts believe that these breaks coincide with attempts by the new Iranian regime to re-integrate into the

international community. We expect to see Iran's diplomatic effort to reflect this through changes in these coefficients. As Figure 2 shows, many of the predictors of Iran's diplomatic behavior remained the same across these periods: Iran continued to meet with allies, existing trade partners, proximate states. There are three major changes. First, rogue states were significantly less important to Iran's foreign policy in the latter periods. Second, Iran met increasingly with democratic states over the same period. Second, Iran focused its diplomacy in Europe and the Middle East more in the latter periods.

These differences fit with the expert assessment of changes in Iran's foreign policy. Since Iran no longer pursued nuclear capabilities, it had little reason to meet with rogue states with whom it may trade nuclear technological with, or seek to otherwise subvert sanction. Furthermore, Iran was under an international microscope. It faced strong incentives to cultivate a good impression. Meeting with North Korea, Sudan and others may have caused international concern. Related, democratic countries were much more likely to uphold sanctions legislation against Iran. While these sanctions were fiercely imposed, there was little incentive for Iranian leaders to meet with these states because they had nothing to gain. It was only once Iran believed that sanctions could be lifted that it saw value in rekindling these relationships. Iran's shifting regional focus also fits well with the easing of sanctions. Before sanctions, European and Asian states had long been Iran's major trade partners. But these states also complied with sanctions legislation. Once sanctions eased, Iran had strong incentives to re-kindle these relationships quickly to re-invigorate its economy.

We found two breaks in Russia's diplomatic behavior: before Russia invaded the Ukraine and Syria respectively. As Figure 2 shows, there are two major differences in Russia's meeting choices. First, Russia increasingly focuses on the Middle East.<sup>22</sup> By the third period, the Middle East is the strongest regional coefficient. This is consistent with expert assessments that Russian policy moved towards expanding into the Middle East.

Second, Russia simultaneously increased the number of rogue states it met with and

---

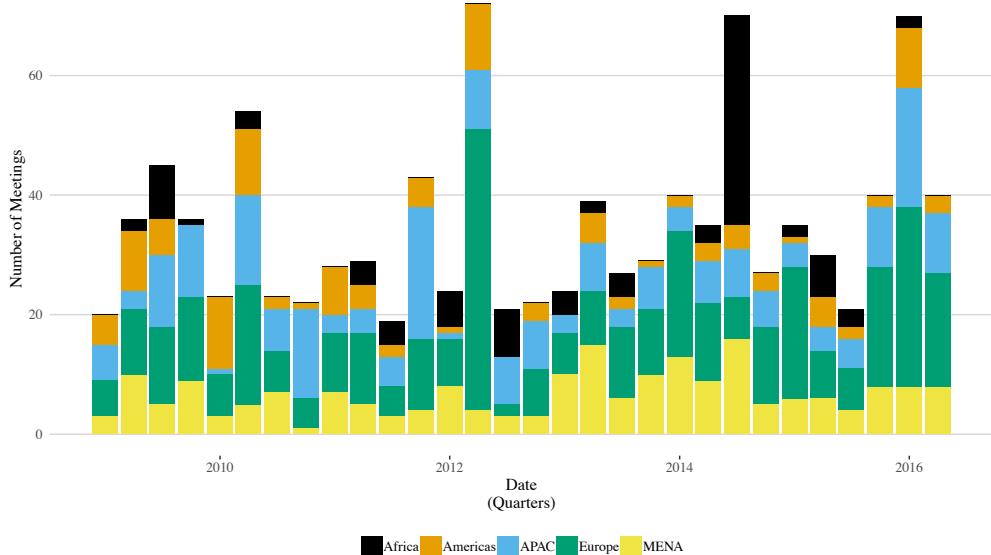
<sup>22</sup>The difference between period 1 and 3 for the Middle East coefficient are statistically different.

decreased the number of democratic states it met with. Third, Russia met more frequently with alliance partners in the second and third periods. This is consistent with expert assessments that Russia increasingly pursued a military-first foreign policy that would put it at odds with major military powers (NATO), and place it as an outcast in the international community.

The changing rogue state and polity score coefficients in Iran and Russia are notably contrast. Experts believe that Iran started as a pariah, but wanted to demonstrate it had rehabilitated. Thus, it met less with rogue states and autocrats across time. In contrast, Russia started off in reasonably good international standing but was increasingly perceived by experts as subverting world order.

We found that considerable fluctuation in American diplomatic behavior. However, we could not find a break in the data that suggested different foreign policy regimes. We reasoned that United States diplomatic behavior was erratic, perhaps caught between different policies. This is somewhat surprising because the president explicitly pointed to his diplomatic record as evidence of his focus on Asia that departed from prior Administrations. Critics of Obama noted that he spent just as much time re-assuring European and Middle East partners that America would honor commitments in those regions as he did cultivating relationships in Asia. Even though the Administration publicized its Asia program, its activities in Europe remained high. If this is true, we expect that the erratic results that the M-fluctuation test found are driven by the Administration switching its focus between different regions in different periods. Summary statistics support this position. Figure 3 plots the number of meetings by region quarter over the full period. As you can see, Obama does intensely focus on Asia in the quarter after he announces the Pivot to Asia. However, just 2 quarters later he surges diplomatic effort in Europe. Comparing these results to the residuals plot in Figure 2, large changes in regional focus coincide with periods with large residuals. Consistent with expert reports, Obama's shifting regional focuses drove erratic fluctuations in his diplomatic behavior.

Figure 3: Count of diplomatic meetings by region during the Obama Administration



This finding emphasizes a feature of our model choice described above: it does not detect gradual changes in a state’s diplomatic activity. In cases where states make only small adjustments to their foreign policies over time, the BIC and RSS will not detect a substantively new foreign policy regime. Yet the M-fluctuation test will identify erratic behavior.

We did not detect a break point in either Australian or American diplomatic behavior and so cannot partition the data and compare coefficients. But to make certain that the diplomatic model is capturing foreign policy choices, we plot the coefficients for the full period in Appendix B. We find that the models’ coefficients are generally consistent with foreign policies of those countries, and prior results (Lebovic and Saunders 2016).

These findings increase our confidence that states do in fact make diplomatic choices to serve their foreign policies. In every case, the magnitude of coefficients reflects the overarching priorities of states. Furthermore, a qualitative analysis of coefficients in different periods supports the idea that the patterns of states’ diplomatic choices change dramatically

with sudden shifts in policy. When experts believe that a state changes their overarching priorities, our structural break test detects a change in diplomatic behavior and the new predictors of diplomatic change we report are consistent with qualitative accounts of how that state's foreign policy changed.

## 5 Conclusion

We argue that aggregated patterns of diplomatic behavior correlate with foreign policies—which can change suddenly—and not a state's underlying preferences. When a state adopts a new foreign policy, their leaders will meetings with different partners to achieve their new objectives. Differences in their diplomatic behavior are observable allowing us to measure when states change their foreign policies across time.

The combination of theory and method has important implications for the future empirical study of conflict. We advance a viable strategy for breaking up a state's international behavior into foreign policy regimes. Rather than thinking that state foreign policy behavior is invariant or that all states shift simultaneously based on a single watershed event (such as the Cold War ending), we argue that a state's diplomatic record can be utilized to identify important shifts in foreign policy behavior that correspond to changes in both the international and domestic domains. We contend that including these trends will improve the accuracy of empirical models that struggle to include intentions and foreign policy in their analysis. The approach also has important implications for forecasting foreign policy events. Using shifts in strategy as the first stage to a broader predictive model of crisis may improve our estimates of international crisis risk.

From a theoretical standpoint, we also recommend that studies of diplomatic behavior unify within a single strategic choice framework. We acknowledge that a leader's specific history will cause him or her to form different beliefs about the role of foreign policy based on psychological and disproportional properties. However, conditional on these beliefs, leaders

behave strategically. We therefore expect that leadership and organizational variables matter as part of a broad strategic process that includes continuity and change in the structure of the international system. Rather than pit rational and psychological approaches against each other, we argue that each will effect diplomatic behavior and foreign policy more or less for different states at different points in history depending on an enormously complex system. Therefore, scholars are simply better off measuring both types of variables and including all of them within a single unified framework.

From an empirical standpoint, our analysis indicates that states might change their agendas over time differently within and across units. These shifts are not necessarily in some linear additive way. Accounting for change-points and variable auto-correlation in cross-national data, may lead to more precise estimates—and ultimately improve our understanding—of conflict and cooperation in the international system.

## References

- Albright, David and Andrea Stricker, 2016. Analysis of the IAEA's Third Iran Deal Report: Filling in Missing Details, Tech. rep., Institute for Science and International Security, Washington D.C.
- Allison, Graham, 2015a. 9 Reasons to Support the Iran Deal, *The Atlantic*.
- Allison, Graham, 2015b. A Point-by-Point Response to the Iran Deal's Critics, *The Atlantic*.
- Andreou, Elena and Eric Ghysels, 2002. Detecting multiple breaks in financial market volatility dynamics, *Journal of Applied Econometrics*, **17**(5), 579–600.
- Arena, Philip and Kyle A. Joyce, 2015. Challenges to Inference in the Study of Crisis Bargaining, *Political Science Research and Methods*, **3**(03), 569–587.
- Avenhaus, Rudolf. and I. William. Zartman, 2007. *Diplomacy Games*. [electronic resource] : Formal Models and International Negotiations., Springer.
- Bai, Jushan and Pierre Perron, 2003. Computation and analysis of multiple structural change models, *Journal of Applied Econometrics*, **18**(1), 1–22.
- Bailey, Michael A., Anton Strezhnev, and Erik Voeten, 2017. Estimating Dynamic State Preferences from United Nations Voting Data, *Journal of Conflict Resolution*, **61**(2), 430–456.
- Bull, Hedley., 2002. *The anarchical society : a study of order in world politics*, Columbia University Press.
- Bush, Richard C., 2012. The Response of China's Neighbors to the U.S. Pivot to Asia, Tech. rep., Brookings Institution, Washington D.C.
- Byman, Daniel L. and Kenneth M. Pollack, 2001. Let Us Now Praise Great Men: Bringing the Statesman Back In, *International Security*, **25**(4), 107–146.

- Cederman, Lars-Erik and Nils B. Weidmann, 2017. Predicting armed conflict: Time to adjust our expectations?, *Science*, **355**(6324).
- Cha, Victor D., 2016. The Unfinished Legacy of Obama's Pivot to Asia, *Foreign Policy*.
- Chen, Qimao, 1993. New Approaches in China's Foreign Policy: The Post-Cold War Era, *Asian Survey*, **33**(3), 237–251.
- Chiozza, G. and A. Choi, 2003. Guess Who Did What: Political Leaders and the Management of Territorial Disputes, 1950-1990, *Journal of Conflict Resolution*, **47**(3), 251–278.
- Clinton, Hillary, 2011. America's Pacific Century, *Foreign Policy*.
- Cozad, Mark, 2018. *Strategic Warning on NATO's Eastern Flank: Pitfalls, Prospects, and Limits*, RAND Corporation.
- Dale, Catherine and Pat Towell, 2012. In Brief: Assessing DOD's New Strategic Guidance, Tech. rep., Congressional Research Services, Washington D.C.
- Debs, Alexandre and Jessica Chen Weiss, 2016. Circumstances, Domestic Audiences, and Reputational Incentives in International Crisis Bargaining, *Journal of Conflict Resolution*, **60**(3), 403–433.
- DeYoung, Karen, 2017. Trump administration says it's putting Iran on notice' following missile test.
- Donilon, Tom, 2011. America is Back in the Pacific and will Uphold the Rules, *Financial Times*.
- Durbin, J., R. Brown, and J. Evans, 1975. Techniques for testing the constancy of regression relationships over time, *Journal of the Royal Statistical Society*, **37**(2), 149–192.
- Einhorn, Robert, 2016. The Iran deal: Off to an encouraging start, but expect challenges, Tech. rep., Brookings Institution, Washington D.C.

Einhorn, Robert, Strobe Talbott, Bruce Riedel, Richard Nephew, Norman Eisen, Tamara Cofman Wittes, Suzanne Maloney, and Kenneth M Pollack, 16. The Iran deal, one year out: What Brookings experts are saying, *Brookings Institution*.

EU, , 2005. EU/Russia: The four Common Spaces MEMO/05/103, Tech. rep., Brussels.

EU, , 2009. EU-Russia Common Spaces Progress Report, Tech. rep., European Union, Brussels.

EU, , 2010. Joint Statement on the Partnership for Modernisation EU-Russia Summit, Tech. rep., European Union, Brussels.

EU, , 2017. Russia's National Security Strategy and Military Doctrine and Their Implications for the EU, Tech. rep., European Parliament, Brussels.

Evans, Peter B., Harold Karan. Jacobson, and Robert D. Putnam, 1993. *Double-edged diplomacy : international bargaining and domestic politics*, University of California Press.

Evera, Stephen Van, 1990. Primed for Peace: Europe after the Cold War, *International Security*, **15**(3), 7.

Fang, Songying and Randall W. Stone, 2012. International Organizations as Policy Advisors, *International Organization*, **66**(04), 537–569.

Fey, Mark and Kristopher W. Ramsay, 2010. When Is Shuttle Diplomacy Worth the Commute? Information Sharing through Mediation.

Ford, John, 2017. The Pivot to Asia Was Obama's Biggest Mistake, *Foreign Policy*.

Gibler, Douglas M., 2008. The Costs of Reneging, *Journal of Conflict Resolution*, **52**(3), 426–454.

Gilpin, Robert, 1983. *War and Change in World Politics*, vol. 1983, Cambridge University Press, Cambridge.

- Goemans, Henk, Kristian Gleditsch, and Giacomo Chiozza, 2009. Introducing Archigos: A Dataset of Political Leaders, *Journal of Peace Research*, **46**(2), 269–283.
- Goldstone, Jack A., Robert H. Bates, David L. Epstein, Ted Robert Gurr, Michael B. Lustik, Monty G. Marshall, Jay Ulfelder, and Mark Woodward, 2010. A Global Model for Forecasting Political Instability, *American Journal of Political Science*, **54**(1), 190–208.
- Gordon, Michael and David E Singer, 2015. Iran Agrees to Detailed Nuclear Outline, First Step Toward a Wider Deal.
- Gordon, Philip and Richard Nephew, 2017. The Worst Deal Ever' That Actually Wasn't, *The Atlantic*.
- Green, Michael J., 2016. The Legacy of Obama's "Pivot" to Asia, *Foreign Policy*.
- Green, Michael J., Kathleen H. Hicks, John Schaus, and Mark F. Cancian, 2016. *Asia-Pacific Rebalance 2025*, Center for Strategic & International Studies, Washington D.C., 1st edn.
- Greene, W., 2008. *Econometric Analysis*, Prentice Hall, Upper Saddle River, N.J, 5th edn.
- Hadley, Stephen, 2012. Eight Ways to Deal With Iran, *Foreign Policy*.
- Hatemi-J, Abdunasser, 2008. Tests for cointegration with two unknown regime shifts with an application to financial market integration, *Empirical Economics*, **35**(3), 497–505.
- Hill, Fiona, 2016. Understanding and deterring Russia: U.S. policies and strategies.
- Howell, William G. and Jon C. Pevehouse, 2007. *While dangers gather : congressional checks on presidential war powers*, Princeton University Press.
- Jackson, Matthew O and Massimo Morelli, 2011. The Reasons for Wars: An Updated Survey, *The Handbook on the Political Economy of War*, pp. 34–57.
- James, Kitfield, 2012. Is Obama's 'Pivot to Asia' Really a Hedge Against China?, *The Atlantic*.

- JÖNSSON, CHRISTER, 2002. Diplomacy, Bargaining and Negotiation, in *Handbook of International Relations*, pp. 212–234, SAGE Publications Ltd, 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom.
- Joseph, Michael F and Michael Poznansky, 2018. Media technology, covert action, and the politics of exposure, *Journal of Peace Research*, **55**(3), 320–335.
- Katzman, Kenneth, 2015. Iran's Foreign Policy, Tech. rep., Congressional Research Services, Washington D.C.
- Keohane, Robert O and Lisa L Martin, 1995. The Promise of Institutionalist Theory, *International Security*, **20**(1), 39–51.
- Kinne, Brandon J, 2014. Dependent Diplomacy: Signaling, Strategy, and Prestige in the Diplomatic Network, *International Studies Quarterly*, **58**(2), 247–259.
- Kofman, Michael, Katya Migacheva, Brian Nichiporuk, Andrew Radin, Olesya Tkacheva, and Jenny Oberholtzer, 2017. *Lessons from Russia's Operations in Crimea and Eastern Ukraine*, RAND Corporation.
- Kurizaki, Shuhei, 2007. Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy, *American Political Science Review*, **101**(03), 543–558.
- Kydd, Andrew, 2003. Which Side Are You On? Bias, Credibility, and Mediation, *American Journal of Political Science*, **47**(4), 597–611.
- LaFranchi, Howard, 2013. Obama cancels Asia trip. Is the US 'pivot' in jeopardy?, *The Christian Science Monitor*.
- Lebovic, James H. and Elizabeth N. Saunders, 2016. The Diplomatic Core: The Determinants of High-Level US Diplomatic Visits, 19462010, *International Studies Quarterly*, **60**(1), 107–123.

- Lieberthal, Kenneth, 2011. The American Pivot to Asia, *Foreign Policy*.
- Lindsey, David and William Hobbs, 2015. Presidential Effort and International Outcomes: Evidence for an Executive Bottleneck, *The Journal of Politics*, **77**(4), 1089–1102.
- McCarthy, Kevin, 2015. 21 Reasons the Iran Deal is a Bad Deal, Tech. rep., Congress, Washington D.C.
- McDonald, Hamish, 2015. Turnbull's foreign policy likely to pivot to China.
- McManus, Roseanne W., 2018. Making It Personal: The Role of Leader-Specific Signals in Extended Deterrence, *The Journal of Politics*, pp. 000–000.
- Mearsheimer, Jj, 1990. Back to the future: instability in Europe after the Cold War, *International security*, **15**(1), 5–56.
- Mesquita, Bruce Bueno De, Alistar Smith, and Randolph M. Siverson, 2005. *The logic of Political Survival*, MIT Press.
- Minhas, Shahryar, Peter D Hoff, and Michael D Ward, 2016. A new approach to analyzing coevolving longitudinal networks in international relations, *Journal of Peace Research*, **53**(3), 491–505.
- Mogherini, Federica, 2016. Shared Vision, Common Action: A Stronger Europe, Tech. rep., European Union, Brussels.
- Morello, Carol, 2017. IAEA confirms Iran is meeting its commitments under nuclear agreement.
- Narayan, Paresh Kumar, Seema Narayan, and Sagarika Mishra, 2013. Has the structural break slowed down growth rates of stock markets?, *Economic Modelling*, **30**, 595–601.
- Neumayer, Eric, 2008. Distance, power and ideology: diplomatic representation in a world of nation-states, *Area*, **40**(2), 228–236.

- Nottage, Luke, 2016. Investor-State Arbitration Policy and Practice in Australia, Tech. rep., Centre for International Governance Innovation., Ottawa.
- O'Hanlon, Michael E., 2013. Missing Hillary Clinton's Hand in the Asia Rebalance, *Brookings Institution*.
- Oliver, Alex, 2015. What to Expect from Malcolm Turnbull, *Foreign Policy*.
- O'Rourke, Lindsey A., 2018. *Covert regime change : America's secret Cold War*.
- Pang, Xun, Barry Friedman, Andrew D. Martin, and Kevin M. Quinn, 2012. Endogenous Jurisprudential Regimes, *Political Analysis*, **20**(04), 417–436.
- Parsi, Trita, 2012. How Obama should talk to Iran.
- Paternoster, RAYMOND, ROBERT Brame, PAUL Mazerolle, and ALEX Piquero, 1998. Using The Correct Statistical Test For The Equality Of Regression Coefficients, *Criminology*, **36**(4), 859–866.
- Pezard, Stephanie, Andrew Radin, Thomas Szayna, and F. Larrabee, 2017. *European Relations with Russia: Threat Perceptions, Responses, and Strategies in the Wake of the Ukrainian Crisis*, RAND Corporation.
- Pillar, Paul, 2015. The importance of the Iran agreement, *The National Interest*.
- Plouffe, Michael and Roos Sterrenvan der , 2016. Trading representation: Diplomacy's influence on preferential trade agreements, *The British Journal of Politics and International Relations*, **18**(4), 889–911.
- Putnam, Robert D., 1988. Diplomacy and domestic politics: the logic of two-level games, *International Organization*, **42**(03), 427.
- Putz, Catherine and Shannon Tiezzi, 2016. Did Hillary Clinton's Pivot To Asia Work?, *The Diplomat*.

- Ramsay, Kristopher W., 2011. Cheap Talk Diplomacy, Voluntary Negotiations, and Variable Bargaining Power, *International Studies Quarterly*, **55**(4), 1003–1023.
- Rathbun, Brian C., 2014. *Diplomacy's value : creating security in 1920s Europe and the contemporary Middle East.*
- Renshon, Jonathan, 2016. Status Deficits and War, *International Organization*, **70**(03), 513–550.
- Royce, Ed, 2015. The Cost of This Bad Iran Deal, Tech. rep., House Committee on Foreign Affairs, Washington DC.
- Rubin, Jennifer, 2015. 10 reasons the Iran deal is ludicrous.
- Sartori, Anne E., 2005. *Deterrence by diplomacy*, Princeton University Press.
- Saunders, Elizabeth, 2011. *Leaders at war : how presidents shape military interventions*, Cornell University Press.
- Scott, Jason, 2017. Australia Fears Asia Power Shift to China If U.S. Withdraws.
- Sherr, James, 2014. A Memo on Russia and Ukraine, Tech. rep., Carnegie Europe, Washington D.C.
- Simpson, Kumuda, 2015. Issues in Australian Foreign Policy, *Australian Journal of Politics and History*, **61**(4), 606–617.
- Slim, Randa, 2016. Putin's Master Plan for Syria, *Foreign Policy*.
- Spirling, Arthur, 2012. U.S. treaty making with American Indians: Institutional change and relative power, 1784-1911, *American Journal of Political Science*, **56**(1), 84–97.
- The White House, , 2011. Remarks By President Obama to the Australian Parliament.

- Thomas, Raju G. C., 1986. India's Nuclear and Space Programs: Defense or Development?;, *World Politics*, **38**(02), 315–342.
- Trager, Robert F., 2010. Diplomatic Calculus in Anarchy: How Communication Matters, *American Political Science Review*, **104**(02), 347–368.
- Trager, Robert F., 2011. Multidimensional Diplomacy, *International Organization*, **65**(03), 469–506.
- Waltz, Kenneth N., 1979. *Theory of international politics*, McGraw-Hill, Reading, Mass.
- Wark, Wesley K., 1985. *The ultimate enemy : British intelligence and Nazi Germany, 1933-1939*, Cornell University Press.
- Weidmann, Nils B. and Michael D. Ward, 2010. Predicting Conflict in Space and Time, *Journal of Conflict Resolution*, **54**(6), 883–901.
- Whitehead, Charles K, 2006. What' s Your Sign? –International Norms, Signals, and Compliance WHAT'S YOUR SIGN?-INTERNATIONAL NORMS, SIGNALS, AND COMPLIANCEt, *Michigan Journal of International Law*, **27**(3).
- Zeileis, Achim, 2005. A Unified Approach to Structural Change Tests Based on ML Scores, F Statistics, and OLS Residuals, *Econometric Reviews*, **24**(4), 445–466.
- Zeileis, Achim and Kurt Hornik, 2007. Generalized M-fluctuation tests for parameter instability, *Statistica Neerlandica*, **61**(4), 488–508.

## A Specification of the M-Fluctuation Test

The M-fluctuation test is performed in three steps. First, we specify a statistical model with a time-ordered component. Second, an empirical partial sum process that is governed by a functional central limit theorem to capture the instabilities in the estimated model. Third, the fluctuation within this process needs to be measured by a scalar functional yielding a test statistic that can be compared with the distribution of the functional as applied to the corresponding limiting process.<sup>23</sup>

We first assume that the outcome of interest,  $Meeting_{cq}$ , is a vector of independent observations distributed according to  $F$  with  $k$ -dimensional parameter  $\theta_i$ :

$$Meeting_{cq} = F(\theta_q). \quad (2)$$

We define  $F$  as the choice model specified in equation 1. We then test the null hypothesis:

$$H_0 : \theta_i = \theta_0 \forall i \quad (3)$$

where  $F(\theta_0)$  represents the distribution of vector values across the full model. The alternative is that at least one  $\theta_i$  deviates conditional on the model, and the ordering of time. To measure the fluctuation within this process we take the scores estimated via maximum likelihood procedure, the minimizing function we select is therefore the partial derivative of the log-likelihood process  $\hat{\theta}$ . The result is a p-value that suggests one (or more) periods deviates from the model mean conditional on the history of events. We plot the shifts to demonstrate how the data fluctuates across time.

To identify where and how many breaks there are we apply the generic structural break test derived by [Bai and Perron \(2003\)](#) to the mean residuals from each time period in equation 1. Again we begin with the ARIMA choice model estimated in equation 1. Next we extract the residuals of each observation and compute the mean and variance of these residuals for each period  $\theta_i$ .

---

<sup>23</sup>For a complete treatment of the M-Fluctuation procedure and its derivation see ([Zeileis and Hornik 2007](#)).

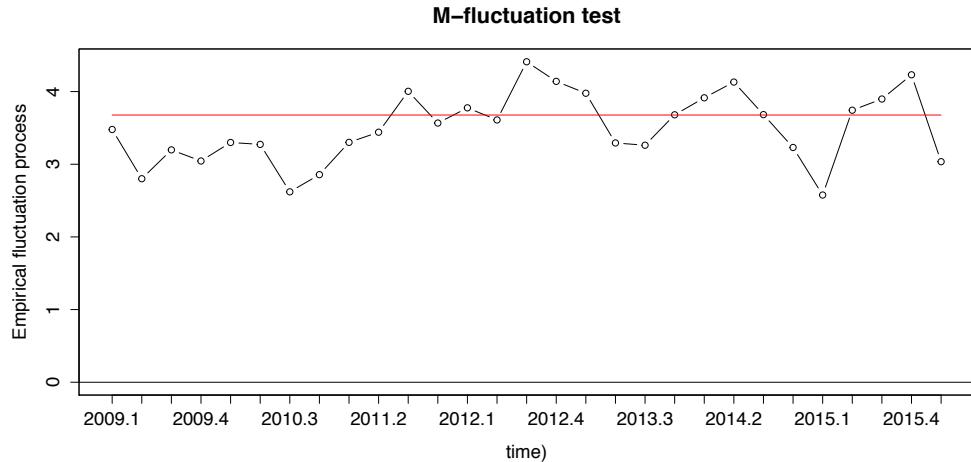
## B Statistical Tests

We report statistical figures of statistical analyses mentioned in the paper.

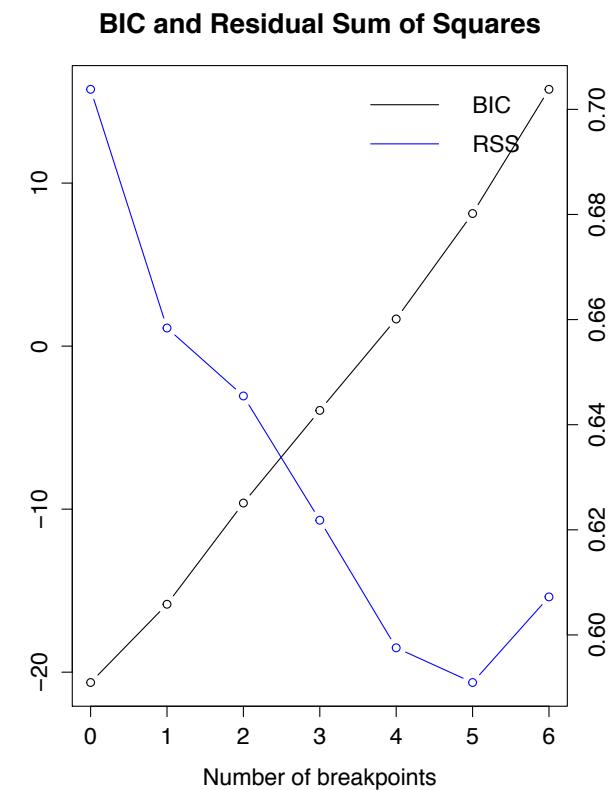
### B.1 Structural Break Tests

We report the visual results for (a) M-fluctuation tests and (b) RSS, BIC tests for each country.

Figure 4: Structural Break Analyses United States

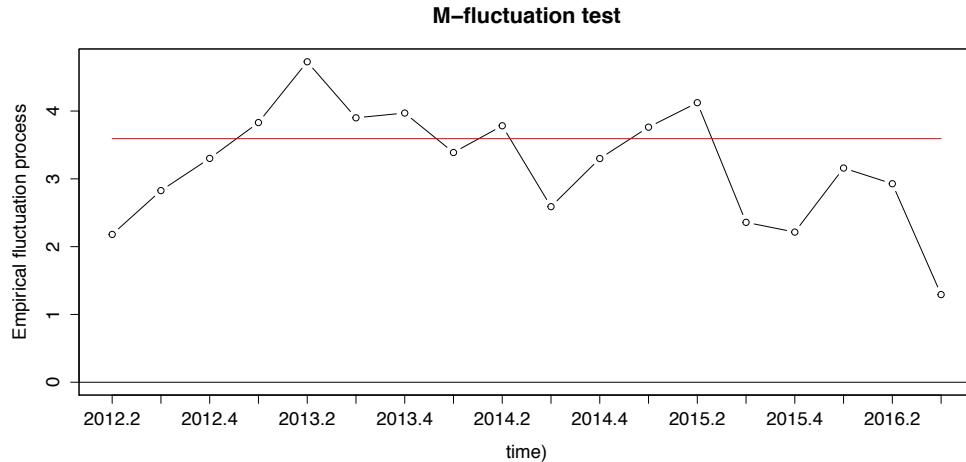


(a)

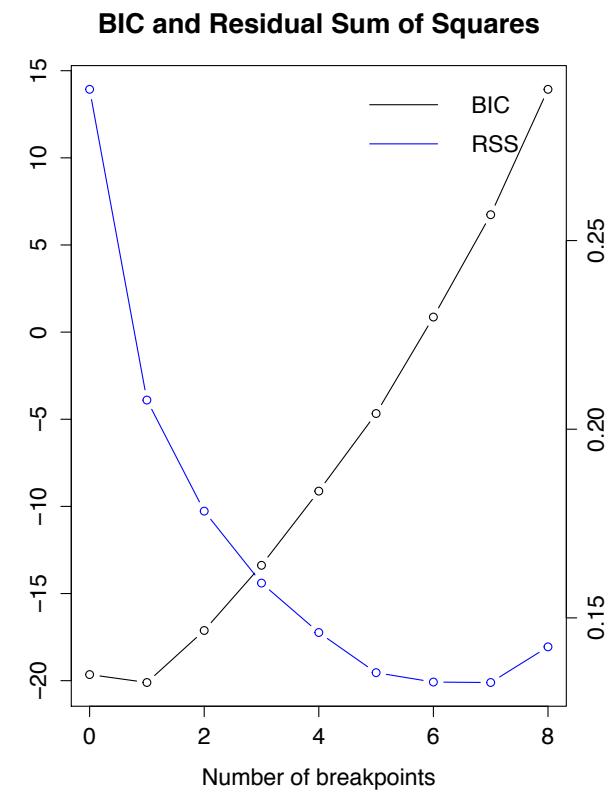


(b)

Figure 5: Structural Break Analyses Iran

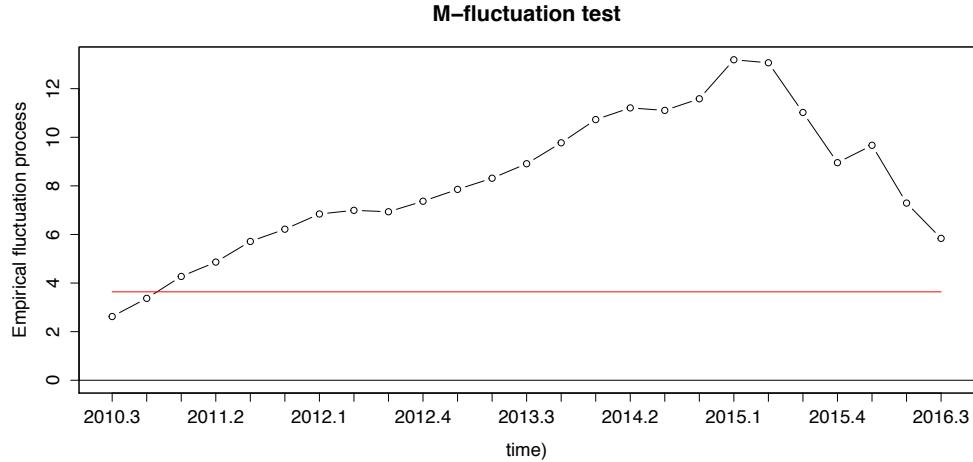


(a)



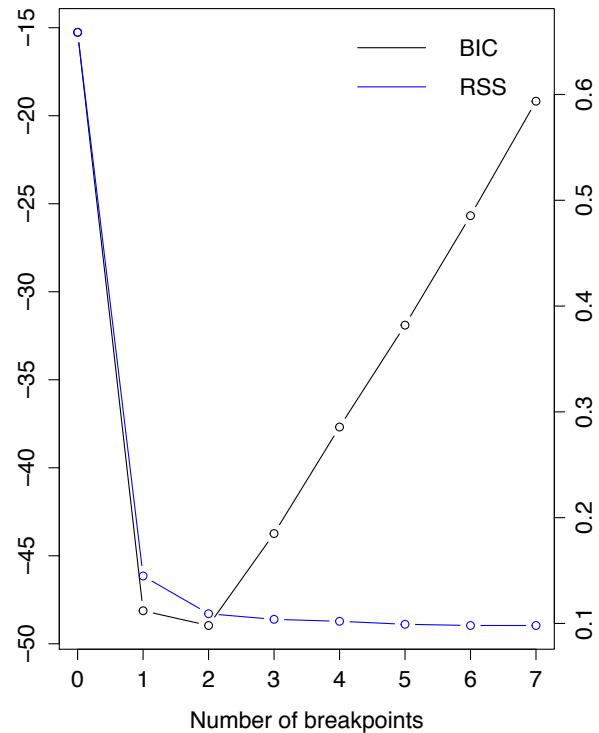
(b)

Figure 6: Structural Break Analyses Russia



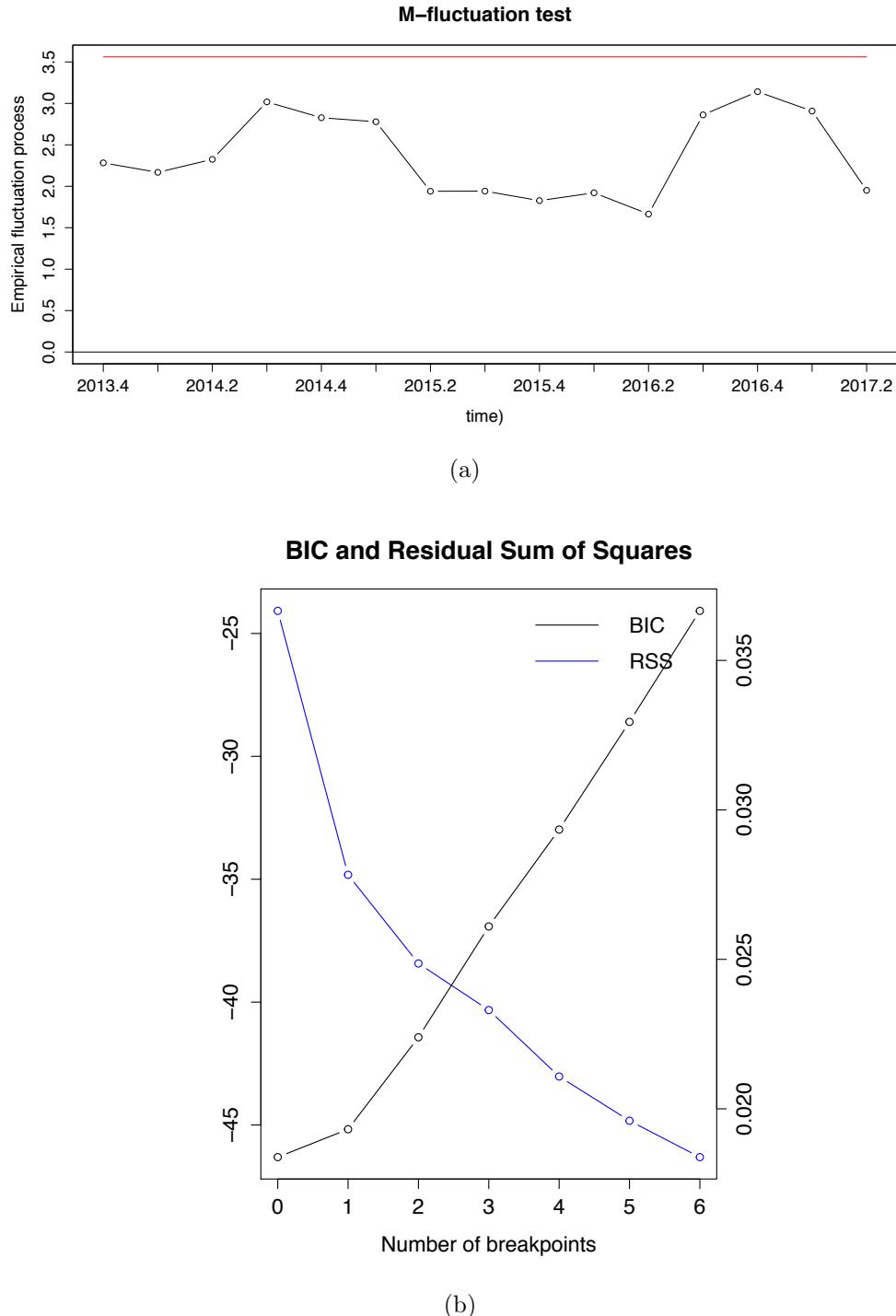
(a)

**BIC and Residual Sum of Squares**



(b)

Figure 7: Structural Break Analyses Australia

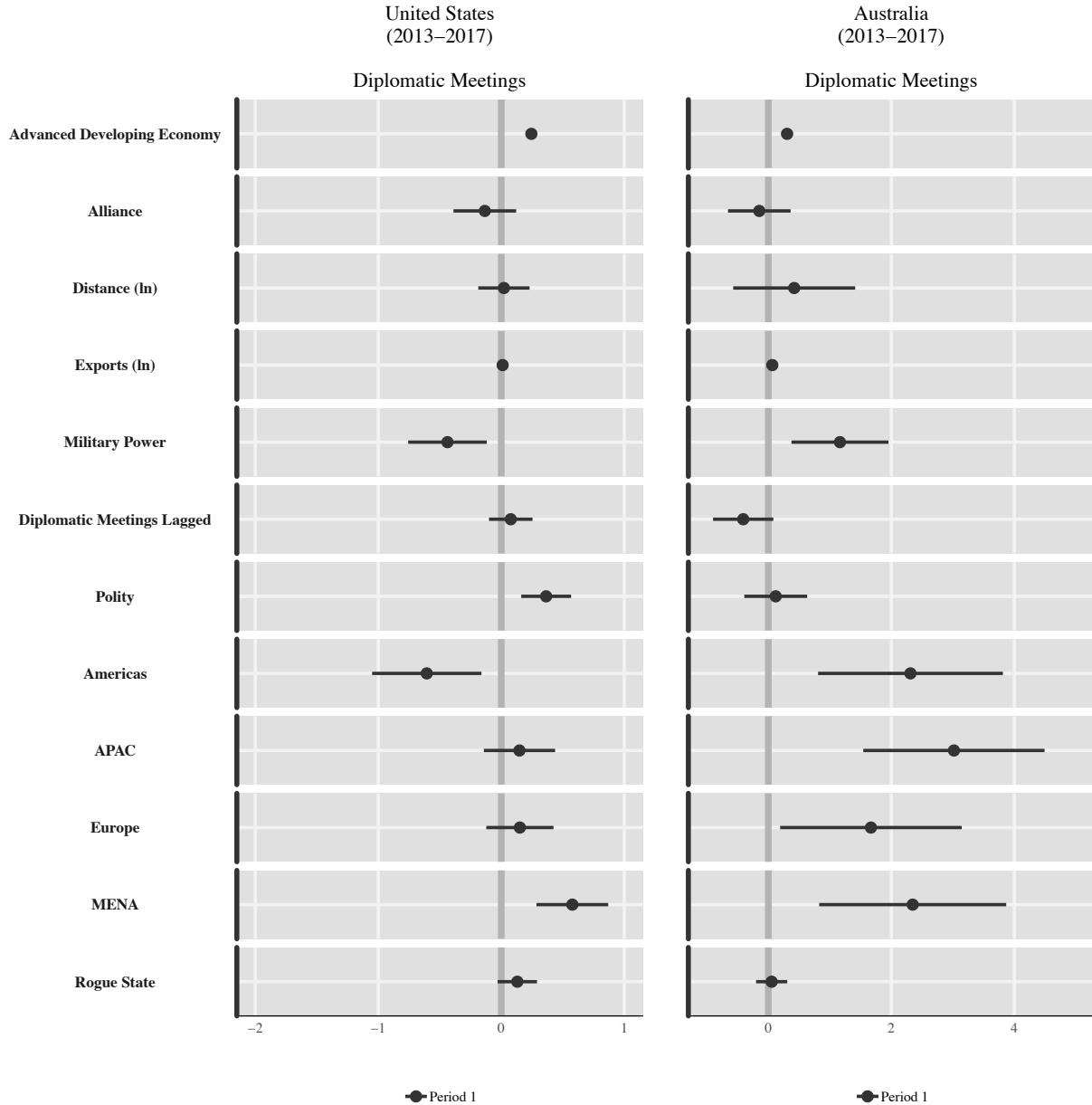


## B.2 Coefficient Plot

We visually report the coefficients for the complete models in Australia and the United States. We did not partition the data because we did not locate break points using our

RSS/BIC tests.

Figure 8: Qualitative assessment of changing strategies



The plot describes the coefficients and 95% confidence intervals from a Poisson regression model defined in equation 1.

## C Machine Coded Events Data

### C.1 Collection Strategy

We focus on the diplomatic exchanges between heads of state (presidents, prime ministers, chancellors, amirs, kings, etc.) and foreign ministers to reconstruct a state’s ego-centric network of diplomacy over time. To recover a complete record of diplomatic meetings between heads of state, we exploit data that the elites themselves provide. Our sources are formal diplomatic websites that are updated and maintained by each respective administration. This offers a record of all self-reported diplomatic exchanges for an administration across its tenure. The websites record formal travel, meetings and phone conversations between counterparts and other dignitaries, diplomatic gestures (e.g. expressions of condolences or congratulations), and news on domestic events and/or proclamations. In essence, these sites serve as a “news feed” for an administration, which provides relevant information on what a particular head of state does during his or her tenure.<sup>24</sup>

We implement an original dictionary-guided machine coding algorithm to identify meetings that occurred between foreign policy elites. We do this by scanning through the associated text for each reference source to gather information on who, how, and when the reference dignitary meets with another dignitary—which we will call the “target”.

To identify the names of elites, we construct a complete cross-national dictionary of all world cabinet positions by month from January 2001 to December 2016. We leverage CIA records—which offer a complete census of world administrations by month—to build these data. The census allows for us to construct an algorithm that flexibly subsets itself by country-month, ensuring no anachronistic entries are considered when identifying named entities in the text.

The machine coding process then proceeds as follows: for a specific text-date (e.g. April, 3, 2015, “President Vladimir Putin met with Ukrainian President Victor Poroshenko”), the dictionary subsets itself to that moment in time. From there, it scans the text for country information and subsets itself again to only include countries that are mentioned. The text then parses all proper nouns and runs that list against the list of all relevant cabinet members (excluding the reference country). When a match is identified, the relevant meta-data for that cabinet member is recorded.

For entries where no match is identified, the algorithm collects textual queues to help infer the target’s identity. For example, from the text “Vladimir Putin meets with his Ukrainian counterpart,” one can infer that Putin is meeting the President of Ukraine, Victor

<sup>24</sup>One potential concern is that leaders may not report all their meetings for strategic reasons. We find this to be unlikely. Leaders are scrutinized intensely by the media and their rivals. Meetings are often pre-approved and take a month at least to plan. Keeping these exchanges a secret is complicated, especially given the high likelihood of a leak on both sides of the exchange. Even in the 1970s, for example, it was difficult for Kissinger to conceal his visit to China. This secret visit took many months of planning and involved a series of complicated plane transfers. Given the time and effort these types of meetings take, we do not believe that there will be a high number of them, but when they occur, they will be observable.

Poroshenko. We account for a range different combinations to assist in the inference process. In addition, we collect additional metadata to identify if the reference is meeting with the target “in person” or “by phone” and whether the meeting is in present or future tense.<sup>25</sup>

The algorithm systematically moves through the scraped text, yielding two end data frames: one containing relevant entries and the other containing irrelevant ones. We define relevant entries as identified meetings or phone calls between two heads of states and/or foreign ministers. Irrelevant entries are defined as any domestic meetings, news and/or proclamations, the signing of legislation, expressions of condolence or congratulations, and/or references to some *future* meeting. For the analysis at hand, we only consider meetings that occur in person.

## C.2 Data Collection: Iran, Russia, Australia and United States

We employ our procedure to systematically collect data for Iran, Russia, Australia, and the United States. For Russia, we report a record of all diplomatic meetings attended by Presidents Vladimir Putin from April 2009 to December 2016, Prime Minister Dmitry Medvedev from May 2012 to December 2016, and Foreign Minister Sergey Lavrov from August 2015 to December 2016. For Iran, we present data on all diplomatic behavior by the Ayotollah Ali Khamenei, Presidents Hassan Rouhani and Mahmoud Ahmadinejad, and Foreign Ministers Ali Akbar Salehi and Mohammad Javad Zarif from February 2012 to October 2016. For Australia, we collected data on Prime Minister Malcolm Turnbull and Foreign Minister Julie Bishop. However, we only utilize the data for Bishop in our analysis due to inconsistency in the Turnbull data.<sup>26</sup> Finally, we collect the archived presidential and state department websites (for Clinton and Kerry) for the entire duration of the Obama administration.

For validation, we randomly sample and review the output for both countries. Review of the relevant material allows us to ensure that the algorithm is capturing what we are intending it to, whereas review of the irrelevant material helps us identify any potential false negatives. Validation through random sampling demonstrates that the machine coding process has a success rate of 97%, which provides considerable confidence in the method. The high accuracy is generally attributed to the consistent underlying structure inherent to meeting data across each case.

To establish that patterns of diplomatic behavior are consistent with common sense expectations about a state’s strategic interests. We present both summary statistics for both Russia and Iran that demonstrate the accuracy and reliability of the collection process and also some face validity checks that diplomatic behavior tracks well with a state’s broad

---

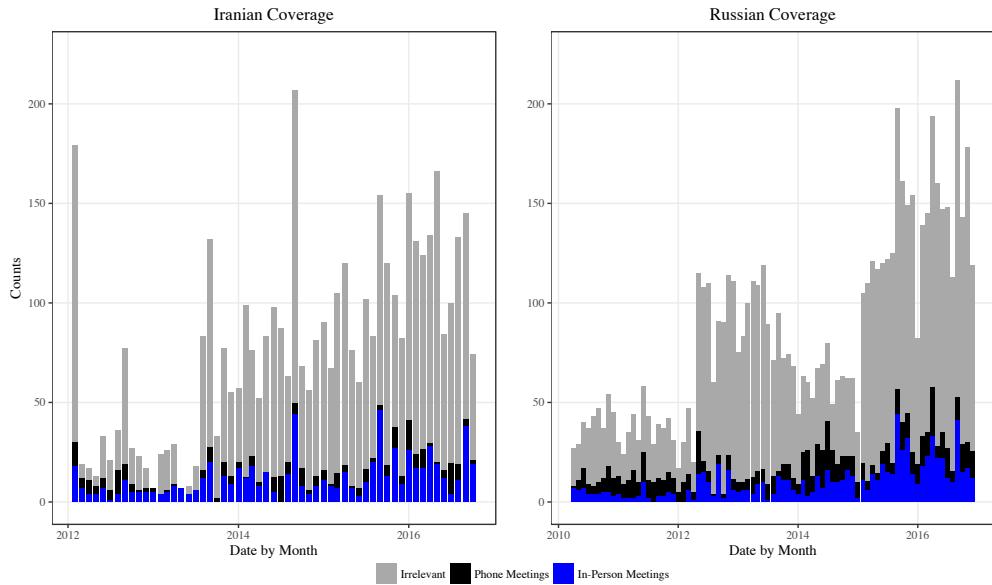
<sup>25</sup>For a small subset of entries, no country information is provided. Thus, the process would miss these entries by design. In our review of the data, we identify alternative structures in the text that help us systematically correct for these occurrences.

<sup>26</sup>Specifically, the data available on Turnbull’s official site are prolonged transcripts of his conversations with reporters. As such, the PM does not consistently reference upcoming or occurring meetings resulting in higher amount of false positives. Given these issues, we opted to rely solely on the Bishop data.

strategic interests as mainstream political science expects.<sup>27</sup>

Figure 9 describes the coverage of meeting information across time. The backward availability of information for each country is not always consistent across time. In addition, different offices tend to update their websites more regularly, e.g. Ayotollah's website for Iran has 80% less content than its Foreign Ministers. These differences are to be expected. As noted above, not all activity that is reported refers to diplomatic meetings. Rather, this information only makes up a small portion of the content on these sites. Figure 9 pools all content by month for the three positions collected for each country and then delineates between in-person meetings and phone conversations with dignitaries from other countries, and irrelevant content. When pooling different elite positions, we are able to recover a robust picture of foreign interactions across time. As the time series shows, there are sometimes surges in meeting activities but that meeting behavior is relatively consistent across time.

Figure 9: Pooled Temporal Distribution of Meeting Coverage



Bar plots present the pooled coverage of Ayatollah Ali Khamenei, Presidents Hassan Rouhani and Mahmoud Ahmadinejad, and Foreign Ministers Ali Akbar Salehi and Mohammad Javad Zarif for Iran; and President/Prime Minister Dmitry Medvedev, Prime Minister/President Vladimir Putin, and Foreign Minister Sergey Lavrov for Russia. Coverage is broken up into irrelevant, telephone, and in-person meetings.

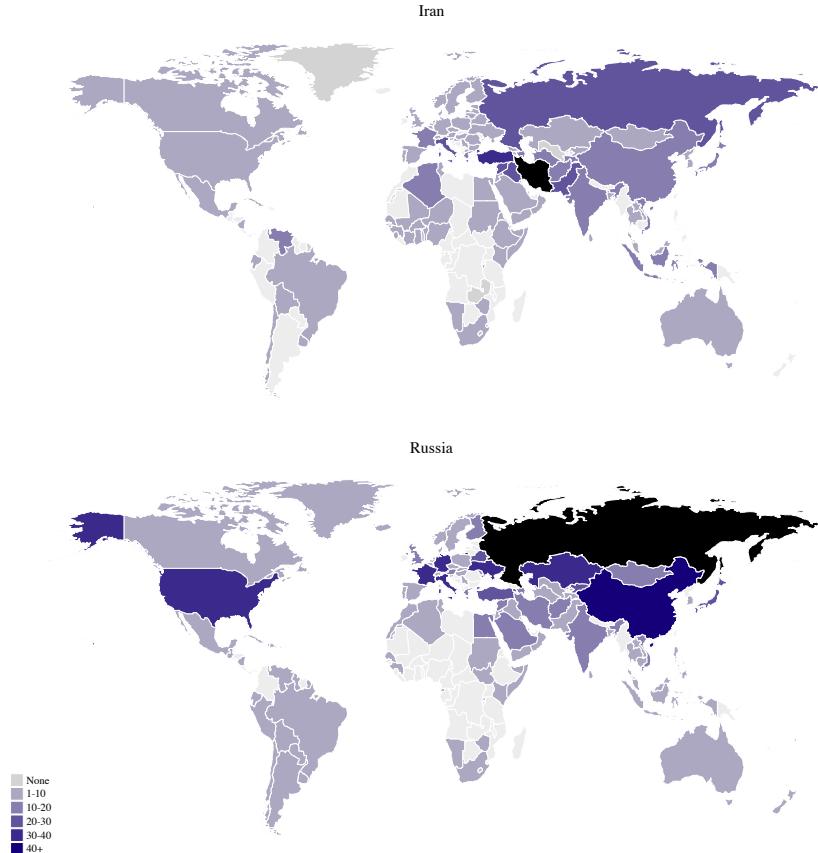
To capture the geographic distribution—that is, who is meeting with whom and how often—we present Figure 10, which describes the pooled count of meetings between countries across the relevant time span for each country. The maps point to different regional focuses and clear differences in the total numbers of meetings. Moreover, the coverage corresponds with expectations regarding Iran and Russia's global coverage. As Figure 10 demonstrates, Russia is far more active on the global stage than its Iranian counterpart, which fits our

---

<sup>27</sup>We exclude Australia given that it is a null case and no formal claims are made; however, the same descriptive statistics are available by request.

prior expectations.

Figure 10: Pooled Geographic Coverage of Meetings



The world maps describe meeting coverage for the relevant time period available for each country. Darker values correspond with a larger number of total meetings. Elite meeting data is pooled for all three elite positions by country. The reference country is filled black.

These descriptives offer confirmatory evidence that our machine-coding procedure consistently and systematically extracts relevant meeting information. As we've noted, the method we outlined here is scalable, making it possible to systematically collect diplomatic events data on an expanded pool of countries.