Two Sides to Every Story: A Theory of Political Contestation and International Institutions *

Stephen Chaudoin

University of Pittsburgh

August 23, 2013

Abstract

A broad class of theories, applied to a wide array of substantive issues, argues that international institutions facilitate compliance by mobilizing pro-compliance domestic groups. I develop a model of political contestation over compliance policy in which international institutions can mobilize *both* pro- and anti-compliance groups. The theory is applicable to a wide variety of issue areas in international cooperation, types of political mobilization, and domestic political institutions. The model predicts institutions are most able to induce compliance where it otherwise would not have occurred when the strength pro- and anti-compliance groups are balanced, *ex ante*. Institutions have a weaker effect on compliance when either group is much stronger than the other. I demonstrate key features of the model using the Kenyan experience with the International Criminal Court. I show how the ICC cemented the political alliance of two indicted candidates (i.e. anti-compliance actors) and helped them mobilize supporters. I also show how, consistent with the model's predictions, the ICC's indictments had the greatest effect on support for the most prominent indicted candidate in regions of Kenya where pro- and anti-indictment forces are balanced.

*I appreciate helpful from advice from Terrence Chapman, Luke Condra, Michael Goodhart, Sarah Hummel, Robert Keohane, James Long, Helen Milner, Susanne Mueller, and Jonathan Woon.

Introduction

The effect of international institutions on the behavior of sovereign nation states is a fundamental question in international relations research. A prominent answer is that international institutions are important because they mobilize, activate, or empower sub-national groups who support policies that are consistent with the institution's goals, i.e. compliance. For example, in 2007, when the International Criminal Court (ICC) first issued arrest warrants for Sudanese politicians suspected of facilitating genocide in Darfur, pro-ICC protests sprung up worldwide as individuals and NGOs increased their efforts to convince elected officials to address the situation in Sudan. As predicted by existing theories, an important effect of the institution (the ICC) was to mobilize pro-compliance groups who wanted the perpetrators of genocide brought to justice. The argument that international institutions mobilize pro-compliance groups is a key feature of broad classes of theories regarding international institutions, such as those based on information provision, audience costs and credible commitments,[1] across a wide array of substantive issues ranging from human rights to international trade.[2]

However, institutions also affect the behavior of groups that support policies that run counter to the institution's rules or goals, i.e. anti-compliance groups. When the ICC issued an arrest warrant for the Sudanese President, Omar al-Bashir, the protests which took place in Sudan's capital, *supporting* al-Bashir, were much larger than the anti-Bashir protests taking place in other parts of the world. Thousands of Sudanese citizens mobilized and rallied in support of their leader and vilifying the ICC as a tool of Western imperialism.[3]

The response to the ICC's actions over Sudan is an example of a common, broader phe-

[1] Carrubba (2005); Mansfield, Milner and Rosendorff (2000, 2002); Rosendorff (2005); Tomz (2007), Chaudoin (Forthcoming a and b).

[2] For example, Simmons (2009); Simmons and Danner (2010) on human rights and Mansfield, Milner and Rosendorff (2000); Buthe and Milner (2008); Elkins, Guzman and Simmons (2006) on trade and investment.

[3] McDoom, Ophera, "Thousands of Sudanese rally in support of Bashir," Arab News/Rueters, July 14, 2008. Rice, Xan, "Sudanese president tells international criminal court to 'eat' arrest warrant," The Guardian, March 3, 2009.

nomenon: institutional actions mobilize *both* sides of an issue. In virtually every issue area affected by international institutions, there are at least two groups with divergent preferences over compliance with the institution's rules. This divergence arises because a government's decision to comply entails distributional consequences, with some groups benefitting more or incurring less costs from compliance than other groups. This phenomenon occurs in many contexts where international cooperation is important. For example:

- Oppressed citizens benefit from the removal of leaders who violate human rights treaties, but members of the incumbent regime's in-group lose.
- Exporting and downstream firms benefit from the removal of trade barriers that violate a trade agreement, but import-competing firms lose.
- "Clean" firms benefit from stricter environmental regulations to meet global abatement targets, but "dirty" producers lose, etc.

Examining the effects of international institutions on the behavior of pro-compliance groups is to tell, at best, only half the story. A country's compliance policy is ultimately a function of contestation between these opposing groups. Groups mobilize and exert costly effort in order to influence the outcome of the contest and attain their preferred policy outcome. Mobilization efforts can range from the purely political (campaign contributions, voting, coalition forming, logrolling, etc.) to the more conflictual (protests, violence, etc.).

I develop a theory in which international institutions can affect the mobilization decisions of both pro- and anti-compliance groups. In response to actions by international institutions, both pro- and anti-compliance groups choose their level of mobilization strategically, i.e. with an eye on influencing compliance policy and with an eye on *each other's* mobilization efforts. If an institution mobilizes pro-compliance groups to work harder to influence policy decisions, then a potential best-response for anti-compliance groups is to increase their mobilization efforts as well.

A theory which accounts for the equilibrium effects of institutions on anti-compliance groups yields testable predictions for the effect of institutions on both groups' efforts and when institutions have the greatest ability to induce compliance. Institutions are most effective at inducing compliance when the two groups are more balanced *ex ante*, i.e. when both groups have comparable valuations of winning the contest and costs to effort. Institutions are less effective when pro-compliance groups are particularly weak or strong. When institutional actions mobilize weak pro-cooperation groups, this may induce greater effort on the part of those groups. However, this also induces anti-compliance groups to ratchet up their effort even more, which minimizes the effect of the institution on the resulting political contest over compliance. While increased mobilization by pro-compliance groups swings the contest towards compliance, the corresponding increase in effort by anti-compliance groups mutes this effect. When pro-compliance groups are stronger, *ex ante*, those groups are already likely to win the political contest over compliance, even without an institutional jolt. In other words, the effect of institutions on the likelihood of compliance is non-monotonic in the relative strength of the two groups.

I demonstrate the key features of this theory by analyzing the recent Kenyan experience with the ICC during the 2013 presidential elections. This situation is particularly well-suited for analyzing the theory described above. In the middle of an important political contest -the 2013 presidential election- the ICC issued indictments against two of the three most prominent candidates. This allows me to analyze the efforts of different groups and the outcome of political contestation before and after a prominent institutional action. I use qualitative data on the decisions of political elites and quantitative data on public opinion to show two things that are consistent with the theory: (1) the ICC increased the effort levels of anti-compliance actors and (2) the effect of the ICC on support for the main indicted politician, Uhuru Kenyatta, was highest in regions where his support was balanced with other candidates, *ex ante*.

The broader implication of this article is to suggest a reframing of how scholars should think about the effects of international institutions on member state compliance decisions. Rather than

think of institutions as way to mobilize pro-compliance groups, we can better understand the effects of institutions if we think of them as mechanisms to affect political contests between pro- and anti-compliance groups. Even beyond the impact on the outcome of a political contest, the efforts of both groups have direct welfare consequences because effort itself is costly. As the stakes of international issues rise and international institutions gain prominence, understanding how institutions affect contestation between groups will be increasingly important to assessing their overall welfare effects.

Accounting for how institutions affect both sides of political contests also has important implications for how institutions should target their own efforts. For example, institutions like the ICC should account for the both pro- and anti-accountability groups when deciding which war criminals to indict. A member of the World Trade Organization should account for both pro- and anti-free trade groups when deciding whether to initiate a WTO dispute against another member. The desire to have the greatest impact might suggest that institutions should take actions in the "toughest cases," where pro-compliance groups are weakest. However, my theory argues that institutions can play the greatest role when they can "tip the balance" in favor of pro-cooperation groups relative to anti-compliance groups. This is consistent with recent calls for international actors to do a better job of "triaging" cases, i.e. prioritizing cases where international actions can have the greatest effect.[4]

In the sections that follow, I first motivate consideration of anti-compliance groups in greater detail, showing myriad examples of how both pro- and anti-compliance groups play an important role in determining compliance policy. In Section 3, I develop an original game-theoretic model in which two groups exert effort in a contest over which group gets to choose their country's compliance policy. An international institution (also strategic), can potentially send a signal that influences one group's value to winning the contest, in turn affecting *both* groups' equilibrium effort levels. Section 4 analyzes these predictions in the context of the 2013 Kenyan presidential

[4]Hafner-Burton (2013).

elections using quantitative and qualitative data. Section 5 concludes.

Two Sides to Compliance

Most international institutions lack direct enforcement capabilities which has led to a growing emphasis on domestic or sub-national enforcement mechanisms. In virtually every issue area concerning international politics, theoretical arguments abound describing how international institutions can help facilitate compliance by activating or mobilizing pro-compliance groups.

In the area of human rights, Beth Simmons (2009) argues that human rights obligations help mobilize citizens to demand better practices and laws from oppressive governments.[5] The theory that institutions can mobilize pro-compliance groups is an important component of arguments linking regime type with human rights behavior. For example, Poe and Tate (1994, pp. 855) argue that democracies are less repressive, since citizens can mobilize against such practices and remove repressive leaders.[6]

In the area of international trade, Mansfield, Milner and Rosendorff (2002) argue that trade agreements act as alarms, triggering citizens to punish elected officials who violate the terms of the agreement.[7] Kono (2006) argues that democracies hide trade barriers from domestic audiences for fear of backlash against the welfare loss caused by protectionism. Pelc (Forthcoming) argues that citizens mobilize to gain information about trade barriers after WTO dispute. In the context of international conflict, Fang (2008) argues that institutions can inform citizens, enabling them to remove bad leaders.[8] Similar examples can be found in the areas of environmental protection, finance, etc.[9]

In addition to the broad array of issue areas, there are a variety of theoretical mechanisms with

[5] Similarly, Simmons and Danner (2010) argue that ICC commitments raise the probability of a "negative reaction of any group in a position to inflict costs... on an actor who defects."

[6] See also: Poe, Tate and Keith (1999); Keith (1999); Keith, Tate and Poe (2009).

[7] See also: Ehrlich (2007); Mansfield and Reinhardt (2008); Pevehouse (2002).

[8] See also: Chapman (2007).

[9] For example, Dai (2005); Lohmann (2003).

which international institutions can mobilize pro-compliance groups. One mechanism is through information provision. Institutions provide information about the occurrence and severity of a government's noncompliance, allowing pro-compliance groups to punish their elected leaders.[10] In another mechanism, institutions provide focal points that help diffuse individuals coordinate their efforts to influence a government's compliance policy.[11] A key argument of Simmons' (2009) theory of human rights treaties is that they increase citizens' value of winning control over human rights practices by inspiring new ideas and helping citizens focus on law. Treaties also raise the probability that pro-human rights groups can favorably influence government policy by precommitting the government to be receptive to citizens' demands and providing new and different resources that can be brought to bear in convincing their government to change its ways. Audience costs are another often-applied theoretical mechanism, where governments who break the rules of an institution suffer increased punishment.[12]

Despite the diversity of issue areas and mechanisms, these theories have a common feature: in a world with a particular international institution (treaty, organization, etc.), pro-compliance groups have a higher likelihood of influencing their country's compliance policy, than in a world without. These arguments all describe how an institution can mobilize latent pro-compliance groups, or give a jolt to those that are already active, in their quest change their government's policies either directly or through leadership change.[13] These theories have been immensely valuable in establishing that international institutions can have significant effects on governments' policies and in describing particular mechanisms through which these effects occur.

Without disputing the value of these arguments, I submit that they tell only half the story. Compliance with the rules of international institutions entails making policy changes that inevitably create winners and losers among domestic groups. Some groups bear higher costs or receive lower

[10] E.g. Mansfield, Milner and Rosendorff (2002); Carrubba (2009).
[11] Simmons (2010).
[12] Tomz (2007, 2008).
[13] For two exceptions, see: Dai (2006, 2005), discussed in greater detail below.

benefits than others, which affects whether they support or oppose compliance. This creates *anti*-compliance groups, who, like pro-compliance groups, can change their behavior in response to the actions of international institutions.

Anti-compliance groups can be found in every issue area and often are equally or more important to determining a country's compliance policy. Consider the context of human rights. In one sense, human rights is an unlikely place to find anti-compliance groups. It is hard to find compelling ethical, normative, or ideological reasons *not* to support human rights. Even so, many attempts to encourage respect for human rights are met with fierce resistance from opposition groups. This resistance is often couched in terms of the defense of "traditional" practices or a defense against Western imperialism. In the late 1800's, British efforts to end the practice of widow burning in India were met by considerable political and violent opposition from groups defending traditional practices. In modern day Japan, attempts to pass implementing legislation for the Convention to End Discrimination Again Women during the 1980's were met with fierce political resistance from those seeking to defend the traditional role of women in society.[14]

Though few would argue in favor of impunity for war criminals, resistance to the ICC in Africa has been particularly serious, driven in large part by the perception that the court is a tool for Western imperialism that only targets African countries. The African Union's (AU) Assembly even went so far as to adopt a declaration that "the AU Member States shall not cooperate pursuant to the provisions of Article 98 of the Rome Statute of the ICC relating to immunities, for the arrest and surrender of President Omar El Bashir."[15] After the ICC issued an arrest warrant for Colonel Gaddafhi during the Libyan crisis, the AU adopted another decision that "[AU] member states shall not cooperate in the execution of the arrest warrant."[16] The pro-Bashir protests mentioned above were littered with chants of "Down, down USA," and even went so far as to burn the ICC chief prosecutor in effigy.

[14]For example, see: Mehta (2004); Marfording (1997).
[15]African Union Assembly Decisions and Declarations, 13th Ordinary Session, July 2009.
[16]African Union Assembly Decisions and Declarations, 17th Ordinary Session, June-July 2011.

Examples of the tension between pro- and anti-compliance groups also abound in international political economy contexts, where policy decisions have direct economic consequences that benefit some groups at the expense of others. Trade policy is a clear example because of the distributional consequences of barriers to trade. In February of 2002, the United States Senate Committee on Finance held hearings regarding protection of U.S. lumber and steel producers. At issue were WTO disputes over U.S. tariff protections for domestic steel manufacturers. Also on the agenda were U.S. tariffs designed to counteract Canadian lumber subsidies- tariffs which would become the subject of a large WTO dispute only a few months later.

While some Senators and hearing participants expressed support for compliance with the WTO and opposition to the tariffs, the hearing was overwhelmingly a platform for tariff supporters and compliance opponents. The Chair of the committee, Senator Baucus (from lumber-producing Montana) opened by lambasting Canadian "give-away prices" before showing contempt for WTO steel disputes saying, "I take particular umbrage when Europe and Japan criticize U.S. trade policy in steel after decades of subsidies and cartels in their own country." Senator Rockefellor (from steel-producing West Virginia) then followed with scathing criticism of the Clinton administration for not having tariffs that were high enough, declaring that "tariffs are the answer." Caught in the middle, Senator Breaux of Louisiana described how he had received two letters- one from a large New Orleans port facility company opposing any tariffs and one from a Louisiana steel producer supporting higher tariffs. He succinctly demonstrated the importance of both pro- and anti-compliance groups, saying, "This is obviously not an easy problem to resolve. We have got very strong feelings on both sides."

Why Analyzing Both Sides Matters

The existence of anti-compliance groups is more than an interesting observational phenomenon. Incorporating anti-compliance groups into a theory of international institutions has implications for when international institutions will be most able to induce compliance, because the policies cho-

sen by sovereign countries are the outcome of political contests between pro- and anti-compliance groups. The idea that government policy choices are the result of contestation between different groups has a rich history, rooted in the study of rent-seeking behavior and lobbying.[17] Groups within society, like special interest groups, assign value to the ability to influence or choose government policy and are willing to exert costly effort to increase that influence. The "prize" for the contest is that the winning group gets to shift government policy closer to that group's most preferred policy. "Effort" could be the lobbying efforts made by the interest groups or contributions (or bribes) made to certain politicians. Groups can vary in their marginal costs or effectiveness of effort.

If one group's valuation of the prize changes, that induces a change in their effort level. If they value the prize more highly, they are willing to exert more effort to attain it. Crucially, this also induces a change in *their opponents'* effort levels, in equilibrium. For example, if side A becomes more willing to exert effort to win a political contest (perhaps in response to an international institution's signal), this also affects the effort of side B. Under certain conditions, side B might decide to increase their effort levels to match A's increase; or side B could decide that A's effort levels are so high that further attempts to influence the contest are futile; or side B could decide to more-than-match A's increases if they value winning enough. This framework is often referred to as a "general equilibrium" approach since effort levels of each group and the outcome of the contest are determined jointly in equilibrium, similar to how production, consumption, and prices are jointly determined in a microeconomic general equilibrium model.

In terms of political contestation, the theories described above argue that institutions can raise pro-compliance groups' value to winning the prize, make their efforts more effective, or make effort less costly. These arguments often resemble a "partial equilibrium approach," because they consider how changes to a particular feature of the model might affect only one side's effort. Since the winner of the contest to influence compliance policy is a function of the effort levels of *both*

[17] Tullock (1967); Krueger (1974); Becker (1983).

pro- and anti-compliance groups, a partial equilibrium approach risks missing important dynamics that affect the policy outcomes we care about.

In the above example of steel and lumber, both pro-tariff (anti-compliance) Senators described possible punishment for politicians who did not support higher tariffs. Senator Rockefellor issued a particularly dire warning to his fellow politicians: "Let me tell you, public opinion is something that you do not want to, gentlemen, take lightly on this matter." Presumably, they feared that lowering tariffs or complying with an adverse WTO ruling would mobilize anti-compliance groups to increase their political efforts- in this case, punishing certain politicians. In 1998, the WTO's dispute settlement body ruled against a popular automotive agreement between the United States and Canada. In response, the Canadian Auto Workers Union and car manufacturers marshalled immense political firepower in support of the agreement. The lobbying effort in support of the auto pact was so strong that when Canada lost the case and was considering appealing the decision, the Canadian government was virtually beholden to pro-agreement groups, with one government official even saying "You have to realize, it's not our decision... Basically we're dealing with the unions and with the industry to see whether or not they want to appeal this."[18]

The approach developed here considers how institutions influence the valuation and costs for pro-compliance groups and how that, in turn, affects the effort levels of anti-compliance groups. The effort levels of both groups then affect the equilibrium probability that each side wins the contest, with pro-compliance groups choosing compliance, and anti-compliance groups choosing noncompliance. A crucial moderator for the effectiveness of international institutions is the relative valuations of each group to winning the prize. When anti-compliance groups value the prize much more highly than pro-compliance groups, an increase in the pro-compliance group's value to winning the prize has two effects. First, and straightforwardly, it increases the effort level of the pro-compliance group. However, increasing the pro-compliance group's value to winning can also increase the effort of the anti-compliance group. Faced with increased effort by their rival, the

[18]Krikorian (2005).

anti-compliance group ratchets up their own effort levels to retain a high probability of winning the prize. As a result, the effect of the institution on the probability that the pro-cooperation group wins the prize is minimal. The anti-compliance group increases their own effort level so that the probability that the pro-compliance group wins is only slightly higher.

When the pro-compliance group values the prize much more highly than the anti-compliance group, an institution also has a minimal effect on the probability that the pro-compliance group wins. This is because the pro-compliance group's high *ex ante* valuation means that they are already likely to choose high effort levels and already have a high probability of winning the contest.

An institution that increases the value of winning to pro-compliance groups has the greatest effect when the two sides' valuations are relatively equal. This is because the institution can best help swing the contest towards favoring the pro-compliance groups. When the two valuations are relatively equal, an institutional signal can induce the greatest increase in the pro-compliance group's effort level, *relative* to that of the anti-compliance group. And this has the greatest effect on the change in the probability that the pro-compliance group wins the contest. The model presented in the next section formalizes these ideas.

I've described these dynamics, so far, in terms of an institution's effect on a group's value to winning the prize. But as will be more apparent in the formal model presented next, my theory also captures the idea that an institution can lower or raise a group's marginal costs to effort or change the effectiveness of their effort levels. For example, Beth Simmons (2009) argues that human rights treaties help oppressed groups influence policy by giving them better access to litigation as a recourse. In other words, the institution (the human rights treaty) helps give pro-compliance groups better "technology" (litigation) with which to try and more effectively influence policy. She uses this theory to generate a non-monotonic prediction that human rights treaties have the greatest effect in countries along the middle of the autocracy-democracy spectrum. I show how all of these mechanisms and accompanying predictions can be captured in a general framework of political

contestation.

The closest related work to mine is from Xinyuan Dai (2006, 2005), who argues that a government's compliance decision is influenced by groups that strongly or weakly support compliance. These groups can support an incumbent or support a challenger, and their influence on electoral outcomes is moderated by their informational capacities, i.e. their ability to discern the effects of their government's compliance decision from stochastic noise. Institutions increase the information available to uninformed pro-compliance groups, increasing their influence over compliance decisions. I build on this work by making different groups' capacity to influence political contests endogenous. Rather than endowing certain groups with stronger or weaker influence on electoral contests, I develop a model where each group strategically chooses how hard it will work (for a cost) to influence the government's compliance decision.

Model

The model presented in this section describes two groups engaging in costly effort to influence a political contest over their country's compliance policy, and an international institution that can potentially influence the groups' behavior. The model is general in three important ways. First, it describes only a pro-compliance and anti-compliance group competing for influence over government policy, which fits many issue areas.The model simply assumes the existence of two groups who have divergent preferences over compliance policy. Second, the model is general to any type of effort or regime type. "Effort" is assumed to have only two features: it is costly and more effort improves, however minimally, a group's chance of influencing compliance policy. Effort thus includes campaign contributions, costly lobbying efforts, voting, etc. (efforts generally associated with democracies) as well as protests or direct opposition to the government (efforts associated with non-democracies). Third, the model also incorporates strategic behavior on the part of the institution by explicitly modeling the institution's preferences and actions. The institution is not

assumed to be a passive or stochastic influence on political contests.

Players, Preferences, and Actions

A society is comprised of two groups: pro-compliance (PC) and anti-compliance (AC) citizens. For simplicity, I consider a representative citizen from each group.[19] Citizens differ in their most preferred government policy, with PC citizens preferring a higher level of compliance than AC citizens. Each assigns value to the ability to set government policy. The AC citizen assigns value $V_{AC} > 0$ to being able to choose their government's policy.

The role of the international institution is to potentially influence the PC citizen's beliefs about the value to setting policy. The value that the PC citizen assigns to this depends on the state of the world- specifically, whether compliance is "beneficial" to the PC citizen or "not beneficial," $\{B, \sim B\}$. When compliance is beneficial, the utility the PC citizen gains from setting policy is v_{PC}. When compliance is not beneficial, I normalize their utility to zero. The probability that pro-compliance policies are beneficial, $p \in (0, 1)$, is commonly known. In other words, the PC citizen has an *ex ante* expected value of getting to set compliance policy, $V_{PC} = pv_{PC}$, but is uncertain about the exact value.[20]

The institution receives a private signal about the state of the world, denoted $\{b, \sim b\}$. The probability that the institution's private information is correctly reflects the state of the world is q. In other words, $Pr(b|B) = Pr(\sim b| \sim B) = q$. I assume that the institution's signal is "accurate enough," such that $q \in (\frac{1}{2}, 1)$. To make terminology consistent, I say that b is a "positive signal," indicating that compliance is beneficial. After observing their private signal, the institution chooses whether to send a positive public signal, S, indicating that compliance is highly beneficial, or to not send a signal, denoted $\sim S$.

The informational environment thus matches features of real world situations. Citizens might

[19] For analysis of collective action problems in contestation games with multiple groups, see Esteban and Ray (2001).

[20] Later, I discuss the possibility that an institution can affect both side's valuations.

not know the value to compliance, e.g. they may not know whether their leaders are likely guilty of war crimes, whether tariff barriers harm the economy, etc. The institution has an informational advantage about the state of the world: the ICC gathers private information over whether a politician has committed war crimes; the WTO gathers information on trade barriers, etc. The possibility that the institution's private information is wrong reflects the imperfection or possibility of bias in these information gathering processes.

Below, I consider equilibria where the institution's public signal is informative, meaning a positive signal from the institution increases the PC citizen's expected utility for setting compliance policy from V_{PC} to V'_{PC}. Conversely, when this institution does not send this signal, it lowers the PC citizen's expected utility for setting policy to V''_{PC}. The features of this equilibrium thus reflect the dynamics described by existing theories. The institution's signal potentially causes the pro-compliance citizen to update her beliefs about the state of the world according to Bayes rule, which induces a change in her expected utility for setting compliance policy.

After the institution's signaling decision, each citizen can engage in costly political activities to try and "win the prize," the ability to set compliance policy. The AC citizen can exert effort, denoted e_{AC}, to win the ability to choose a noncompliant policy. The PC citizen's effort is denoted e_{PC}.

Exerting effort is costly for each citizen, and I allow marginal costs to differ for each group. The costs associated with effort are a linear function of that citizen's effort level. For the PC citizen, $c_{PC}(e_{PC}) = c_{PC}e_{PC}$, and for the AC citizen, $c_{AC}(e_{AC}) = c_{AC}e_{AC}$.[21]

The probability each citizen wins is a function of their effort levels. The probability that the pro-compliance citizen wins, is $\phi_{PC}(e_{PC}, e_{AC}) = \frac{e_{PC}}{e_{PC}+e_{AC}}$, and $\phi_{PC}(0,0) = \frac{1}{2}$. The probability that the anti-compliance citizen's efforts prevail is $\phi_{AC}(e_{PC}, e_{AC}) = \frac{e_{AC}}{e_{PC}+e_{PC}}$, with $\phi_{AC}(0,0) = \frac{1}{2}$.[22]

[21] Results similar to those below can be derived from more general cost functions (Corchón, 2007).

[22] This is the familiar ratio form of contest success functions. Skaperdas (1996) derives this form from appealing axioms and Jia (2008) derives this form from a stochastic setting. In political science, examples of this contest success function in the study of armed conflict include: Slantchev (2010); Garfinkel and Skaperdas (2000); Hegre (2004).

The expected payoffs for a particular effort level chosen by the PC citizen, given the effort of the AC citizen, are thus: $\Pi_{PC}(e_{PC}, e_{AC}) = \phi_{PC}(e_{PC}, e_{AC})V_{PC} - c_{PC}(e_{PC})$. Similarly, the AC citizen's expected payoffs are: $\Pi_{AC}(e_{PC}, e_{AC}) = \phi_{AC}(e_{PC}, e_{AC})V_{AC} - c_{AC}(e_{AC})$.

The international institution's payoffs are affected by whether they choose to send the public signal, S, and whether the pro-compliance citizen prevails in situations when compliance is beneficial. When compliance is beneficial, the institution prefers when the pro-compliance citizen wins the contest, and receives utility of $V_I > 0$. When compliance is not beneficial or when the AC citizen wins the contest, the institution receives a payoff of 0. The institution must also pay a cost, $0 < k < V_I$, if they choose to send a public signal.[23] The institution also values its legitimacy or reputation. All else equal, the institution prefers not to have its signals ignored. If they send the signal and the anti-compliance group prevails, the institution pays a legitimacy cost, $l > 0$.[24]

The sequence of the game is as follows: (1) the institution receives its private information about the state of the world and chooses whether to send a signal, (2) the citizens observe the institution's signal and simultaneously choose their effort levels, (3) a "winner" is realized, and that winner gets to choose their most preferred compliance policy. A perfect Bayesian Nash equilibrium consists of (a) the institution's decision over whether to signal $S, \sim S$ and (b) for each $S, \sim S$, a pair of effort levels, e^*_{PC}, which maximizes Π_{PC} given e_{AC} and e^*_{AC}, which maximizes Π_{AC} given e_{PC}.

Equilibrium Analysis

This section first characterizes an equilibrium in which the institution can send a signal that changes the efforts of both citizens. I then derive optimal effort levels and the probability that each side wins the contest *without* the institution. This generates intuition on how effort levels change in

[23]The costs can be thought of in terms of direct, institutional resources. For the ICC to indict someone consumes the finite time of ICC lawyers, requires expensive travelling for lobbying and consultations with governments, and large expenses for gathering evidence, protecting witnesses, and educating local citizens about the ICC process. There are also opportunity costs- prosecuting one trial consumes resources that could be used on other trials, etc. Similar arguments could be made for virtually all institutions.

[24]For a similar assumption, see Carrubba (2005) where an international regulatory regime wants to maximize compliance.

response to changing valuations and costs for each player. The derivations for optimal effort levels are general to any institutional action that affects either side's value to winning the prize or costs of effort. I then add the institution, describe the effects of an institutional signal on effort levels and use this to characterize the institution's signalling decision.

Note, for arguments where it is not necessary to distinguish between the two citizens, I refer to the two citizens generically as "citizen i" and "citizen j" and replace the PC and AC subscripts with i and j accordingly.

I consider an equilibrium where an institution's signal increases the pro-compliance citizen's expected value of winning the political contest. I call this an "informative equilibrium."[25] Formally,

Proposition 1. *Informative Equilibrium: There exists a Subgame Perfect Bayesian Equilibrium where*

(i) The institution chooses $S|b$ *and* $\sim S| \sim b$

(ii) Citizen i *chooses* $e_i^{*\prime}|S$ *and* $e_i^{*\prime\prime}| \sim S$

(iii) The PC citizen beliefs are $Pr(B|S) > Pr(B| \sim S)$.

Optimal Effort Levels and Win Probabilities

To understand the moving parts of this equilibrium, I start by deriving a general form for each citizen's optimal effort levels given their valuations of winning the political contest. To do this, it is helpful to first transform each citizen's payoffs. Consider a linear transformation of citizen i's ($i \in \{PC, AC\}$) payoffs by dividing Π_i by V_i.[26] Further, define d_i as $d_i \equiv \frac{c_i}{V_i}$. This transformation recasts each citizen's maximization problem in terms of the citizen's relative costs and benefits of effort. d_i represents the ratio of costs to benefits for citizen i: as their value to win-

[25]To be sure, there are multiple equilibria in this game, as in other signalling games. I only analyze this one since it has the intuitive features associated with existing theories of international institutions. Proofs of all propositions and the conditions for the existence of this equilibrium are contained in the supplementary appendix.

[26]Nash equilibria are preserved by linear transformations in payoffs. This approach is from Corchón (2007).

ning increases or marginal cost to effort decreases, d_i *decreases*. Citizen i's maximization problem is thus: $max_{e_i}\Pi_i(e_i, e_j) = \frac{e_i}{e_i+e_j} - d_i e_i$. The accompanying first order condition for citizen i is $\frac{e_j}{(e_i+e_j)^2} = d_i$.[27]

Proposition 2 characterizes e_i^* as a function of d_i and d_j, and Corollary 1 which describes how optimal effort changes with each parameter.[28]

Proposition 2. *In equilibrium, the optimal effort level for citizen i is:*

$e_i^* = \frac{d_j}{(d_i+d_j)^2}$.

Corollary 1. *In equilibrium:*

(i) $\frac{\partial e_i^*}{\partial d_i} = \frac{-2d_j}{(d_i+d_j)^3}$ *and (ii)* $\frac{\partial e_i^*}{\partial d_j} = \frac{d_i - dj}{(d_i+d_j)^3}$.

Intuitively, according to (i) of Corollary 1, the optimal effort level for citizen i is decreasing in d_i. As the value of winning the contest, V_i, increases, citizen i's optimal effort level increases. As effort becomes more costly for citizen i, higher c_i, they exert less effort.

Figure 1 shows the equilibrium effort levels for each player, e_{AC}^* and e_{PC}^*, as V_{PC} increases along the horizontal axis. Looking first at the PC citizen's efforts (blue line), as V_{PC} increases, so too does their equilibrium effort level, as in (i) of Corollary 1. The red line shows how the AC citizen's effort is at first increasing as V_{PC} increases. When the V_{PC} is lower, relative to a fixed V_{AC}, the AC citizen's optimal effort level is *increasing* in the PC citizen's value to winning the prize. When V_{PC} is higher, relative to a fixed V_{AC}, the AC citizen's optimal effort level is *decreasing* in the PC citizen's value to winning the prize. The two curves cross when $V_{PC} = V_{AC}$.

Turning to the AC citizen, the effect of the PC citizen's costs and valuation on the AC citizen's optimal effort level is nonmonotonic. Looking at left side of the blue line, as V_{PC} increases, the

[27] The derivations describe optimal effort levels in any subgame perfect Nash equilibrium, so I temporarily drop the $'$ and $''$ superscripts.

[28] The proof uses the above FOC and the sum of the two FOCs.

AC citizen's effort also increases. However, when V_{PC} gets high enough, further increases in V_{PC} result in decreases in the AC citizen's effort. This non-monotonicity arises because of the key feature of a theory of contestation: that effort levels are determined jointly in equilibrium. When V_{AC} is higher relative to V_{PC}, increases in V_{PC} may cause the PC citizen to increase her own effort level, but they also cause a more-than-proportionate increase in the AC citizen's effort. This non-monotonicity can also be seen mathematically in the expression for (ii) of Corollary 1.

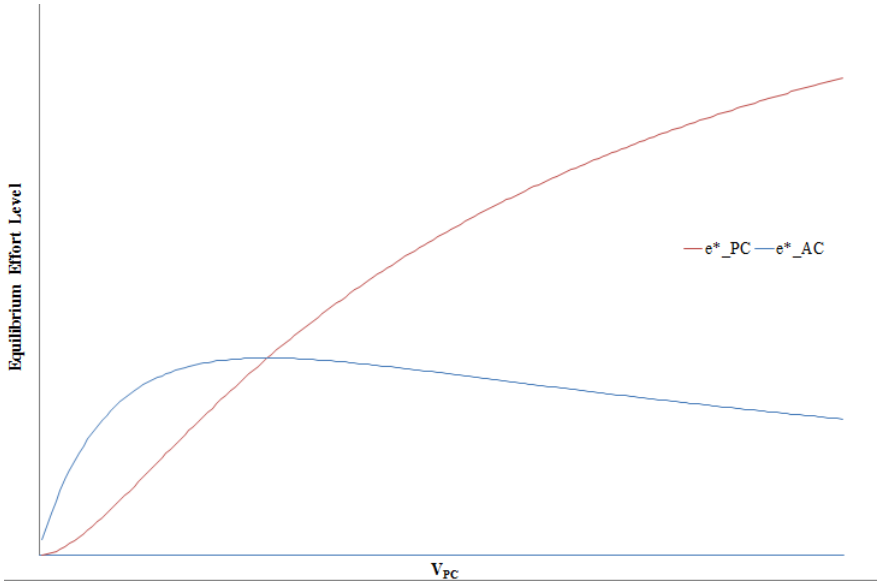
This effect is akin to deterrence. When the AC citizen values winning the contest very highly, she is willing to respond to small increases in her opponent's effort levels with more-than-proportionate increases in her own effort levels, in order to retain her high probability of winning the contest. This is what happens on the left hand side of Figure 1. An increase in the PC's citizen's valuation increases her own effort level, but it also causes an even steeper increase in the AC citizen's effort.

On the other hand, when the PC citizen values winning the contest very highly, increasing her value to winning the contest can decrease the optimal effort of the AC citizen. This is what happens on the right hand side of Figure 1. When the PC citizen very strongly wants to win the contest, and chooses a correspondingly high level of effort, the AC citizen has very little chance to win the contest, and the marginal cost of effort can outweigh the marginal gain in probability of winning. As the PC citizen increases her effort level to reflect her increased desire to win the contest, this drives down the marginal value of the AC citizen's effort. This figure is in terms of the PC citizen's value to winning, but an analogous figure could display a similar relationship considering the PC citizen's costs to effort.

Having characterized how changes in players' valuations and costs affect their optimal effort levels, how do changes to valuations and costs affect who wins the contest? I denote the probability that player i wins the contest as $\phi_i(e_i, e_j)$. Proposition 3 and Corollary 2 describe the effect of changes in d_i and d_j on the equilibrium probability of each player winning.

Proposition 3. *In equilibrium, the probability that citizen i wins the contest is:*
$\phi_i(e_i^*, e_j^*) = \frac{d_j}{d_j + d_i}$.

Figure 1: Equilibrium Effort Levels as V_{PC} Varies



Corrollary 2. *In equilibrium:*

(i) $\frac{\partial \phi_i(e_i^*, e_j^*)}{\partial d_i} = \frac{-d_j}{(d_i+d_j)^2}$

(ii) $\frac{\partial \phi_i(e_i^*, e_j^*)}{\partial d_j} = \frac{d_i}{(d_i+d_j)^2}$.

Intuitively, ϕ_i is decreasing in d_i and increasing in d_j. For example, as the PC citizen values the prize more (decreasing d_{PC}), their effort level increases, and the corresponding probability of winning also increases, even taking into account any corresponding increases in the effort level of the AC citizen (part i of Corollary 2). Similarly, as the AC citizen values the prize more, they increase their effort level, lowering the probability that the PC citizen wins the prize (part ii of Corollary 2).

More broadly, proposition 3 and Corollary 2 say that the institution has at least *some* positive effect on the probability that the PC citizen wins. Even when increasing the PC citizen's value to winning the prize leads to a more than proportionate increase in the AC citizen's effort, the

probability that the PC citizen wins the contest still increases.

Before continuing, it is worth noting that the results derived above concerning optimal effort levels and the probability of each side winning are independent of any particular signalling role of the institution. This model describes a particular way in which an institution can affect one side's valuation to winning a prize: Bayesian updating over an unobserved state of the world. To be sure, there are many other ways that an institution could affect these parameters outside of a signalling mechanism. All of the results for effort and winning probabilities apply to any mechanism through which an institution could affect the value of a prize or the cost of effort.[29]

The Institution's Signal

What affects the institution's decision? The institution weighs their expected gains and costs from sending a signal. The gains arise from the possibility that a signal will cause the PC citizen to increase her effort level and increase her probability of winning the contest. An institutional signal causes the PC citizen to update her beliefs about the state of the world and the expected value of compliance. After observing a signal, the PC citizen's expected value of compliance is higher. And if she observes no signal, her expected value of compliance is lowered. Taking into account the effect of an institutional signal on V_{PC}, I denote her updated d as, $d_{PC}\gamma'$ after an institutional signal, and $d_{PC}\gamma''$ when the institution does not send a signal. Note that because the institution is "honest" (part i of Proposition 1) and if the institution's signal is accurate enough,[30] it implies that $0 < \gamma' < 1 < \gamma''$, and by implication, $d_{PC}\gamma' < d_{PC} < d_{PC}\gamma''$.[31]

This change in the PC citizen's valuation represents the institution's expected gain to sending the signal. Since the institution values compliance, they may want to send the signal in order to induce the PC citizen to exert higher effort and have a higher chance of winning the contest. On

[29] For example, these results would be consistent with a constructivist account in which an institution directly shaped the preferences of an actor or socialized them into more strongly supporting compliance.

[30] Specifically, $q > \frac{1}{2}$.

[31] Full representation of γ is in the appendix.

the other hand, the institution has to pay a fixed cost for sending this signal, and sending the signal also risks their loss of legitimacy if the PC citizen loses the ensuing contest.

To gain intuition on how the institution weighs these costs and benefits, Proposition 4 and Corollary 3 characterize the difference between the institution's expected utility for sending the signal and not sending the signal.

Proposition 4. *In an informative equilibrium, when the institution receives a private signal b, the difference between the institution's expected utility for sending a signal and not sending a signal is:*

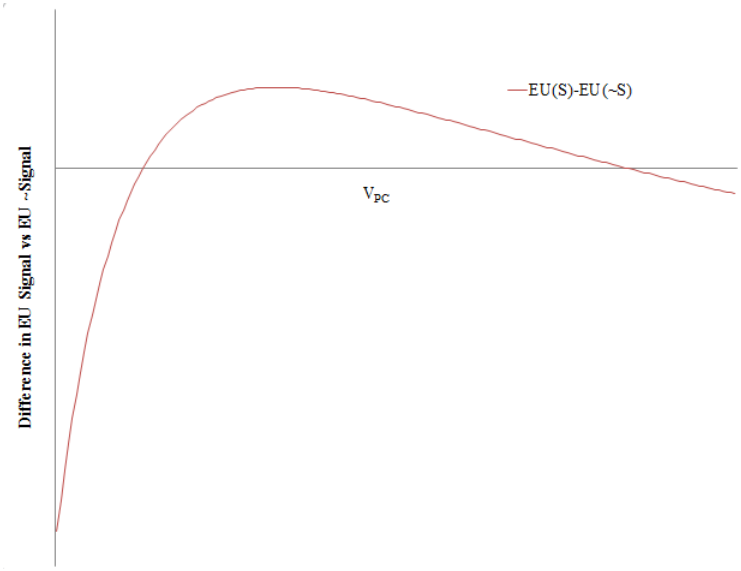
$$EU_I(S|b) - EU_I(\sim S|b) = Pr(B|b)[\phi_{PC}(e^{*\prime}_{PC}, e^{*\prime}_{AC}) - \phi_{PC}(e^{*\prime\prime}_{PC}, e^{*\prime\prime}_{AC})]V_I - \phi_{AC}(e^{*\prime}_{PC}, e^{*\prime}_{AC})l - k$$

Corrollary 3. $\frac{\partial EU_I(S) - EU_I(\sim S)}{\partial d_{PC}} = \frac{pV_I}{\gamma'}[\frac{\gamma'' d_{AC}}{(\gamma'' d_{PC} + d_{AC})^2} - \frac{\gamma' d_{AC}}{(\gamma' d_{PC} + d_{AC})^2}] - l[\frac{\gamma'}{\gamma' d_{PC} + d_{AC}} - \frac{(\gamma')^2 d_{PC}}{(\gamma' d_{PC} + d_{AC})^2}]$

Corollary 3 shows that the institution's optimal decision, S versus $\sim S$ is non-monotonic in d_{PC}. This is not immediately apparent from the expression, so Figure 2 displays this relationship graphically. In words, the institution's signal has the largest effect when the two sides' valuations are roughly equal. This is for two reasons. First, when the anti-compliance citizen values the prize much more than the pro-compliance citizen, she is already exerting a large amount of effort. And even though the institutional signal might induce some increase in the pro-compliance citizen's efforts, the resulting change in the probability that she wins and choose a compliant policy is less likely to outweigh the institution's costs of sending the signal. At the other extreme, when the pro-compliance citizen values the prize much more than the anti-compliance, she is already exerting a larger amount of effort, and is already likely to win the contest, regardless of the signal's effect on her effort level.

The second reason uses the intuition of Proposition 2, part (ii). When the anti-compliance values the prize much more than the pro-compliance citizen, the signal raises the pro-compliance citizen's effort level, but also raises the anti-compliance citizen's effort level by even more. This

Figure 2: Expected “Gain” from Signal



further decreases the marginal benefit, from the institution’s standpoint, of sending the signal to pro-compliance citizens when the anti-compliance citizen values the prize disproportionately much.

Conversely, the institution’s signal has the largest effect when the two valuations are roughly equal. When the two valuations are roughly equal, the increase in the pro-compliance citizen’s effort level is more likely to be pivotal and to swing the contest in her favor. Additionally, when the two valuations are roughly equal, the signal can push the pro-compliance citizen’s effort level above that of the anti-compliance citizen. And when this occurs, according to Proposition 2, increases in the pro-compliance citizen’s effort levels also cause decreases in the anti-compliance citizen’s effort levels.

The effects of other parameters describing the institution’s preferences are straightforward. As the institution’s cost of sending a signal, k, decreases, the curve depicted in Figure 2 shifts upwards, meaning that the institution is willing to send a positive signal over a broader range of values for V_{PC} (or, implicitly $V_{AC}, c_{PC},$ or c_{AC}). The costs also affect the possibility that an informative

equilibrium exists in the first place. At the extreme, if the institution was "publicity seeking" and had very low or even negative k, and thus wanted to send positive public signals regardless of its private information, this would hinder the possibility of an informative equilibrium.[32]

As the institution's fear of losing legitimacy increases, the range of values over which it wants to send the signal also shrinks. Unlike changes in the costs of a signal, this effect is not simply to shift the institution's expected gains from a signal up or down. As legitimacy costs increase, the institution must be particularly careful when d_{PC} is low since this is when the risk of loss of legitimacy is highest. Increasing l makes the left side of the institution's expected gain from a signal steeper.

The conditions for the informative equilibrium to exist are also straightforward and are described formally in the appendix. Informally, they say that the institution's public signal must induce a large enough change in the PC citizen's effort levels to justify the institution's fixed costs of a signal and the risk of legitimacy loss. The degree of change in the PC citizen's effort levels induced by the institution's signal is a function of how accurate the institution's signal is, q, the relative valuations and costs of winning for the two citizens, and the citizens' prior beliefs about the expected value of compliance. For the informative equilibrium to exist, the institution's costs to sending a public signal (both the immediate cost k and the possibility of a loss of legitimacy, l) must be high enough to deter the institution from wanting to send a positive public signal even when it does not receive a positive private signal. And they must be low enough so that the institution wants to send the signal when it receives a positive private signal. Similarly, the PC citizen's costs to effort must be high enough to keep them from exerting the high level of effort ($e^{*\prime}_{PC}$) even when the institution does not send a signal. And the costs must be low enough to make the PC citizen want to choose the higher level of effort when they do observe the institution's signal.

As noted above, this model described how an an institution might shock one side's valuation

[32]The first prosecutor of the ICC was thought by many (fairly or unfairly) to be a publicity hound who plunged the institution into situations with more regard for the preeminence of the institution as opposed to its effects on the ground.

to winning the prize, but it did not incorporate "two sided" shocks, e.g. where the institution affected both the PC and AC citizens' value to winning. It is worth noting that, even if the model incorporated such a feature, many of the propositions above would not change. All of the derivations for optimal effort levels and equilibrium probabilities of one side winning would obtain. In a model with two-sided shocks, the institution's calculus would likely be affected by whether an institutional signal affected the valuation of one group more than the other. I leave such a model to future work.

There are also reasons to think that the model described above, where the institution only shocks the PC group's valuation, reflects real-world dynamics. AC groups (like protectionist special interest groups or incumbent politicians) are generally thought to hold informational advantages over and be better mobilized than PC groups, absent institutional signals. That is, even without an institution, these groups likely have a better idea of the true state of the world. As a result, institutions may not change their valuations as much. This is even more apparent for arguments about how institutions help lower the costs of effort for PC groups, for example, by opening up litigation channels against human rights violations.[33] Such shocks are likely to be one-sided, as modeled above; a human rights treaty helps the PC group litigate against human rights violations but doesn't help AC groups better defend violations.

Application: The Kenyan Presidential Election and the ICC

What can be learned from a model which considers political contestation and effort by two, opposing sides? The goal of this section is to corroborate empirically two key features of the theoretical model: (1) institutional signals have important effects on the effort levels of anti-compliance groups and (2) the effects of institutional signals differ according to the *ex ante* strength of opposing groups. I will demonstrate these using data from the 2011-2013 Kenyan presidential election

[33] E.g. Simmons (2009).

cycle. This is a good laboratory because the election itself is a political contest: opposing candidates exert effort to try and persuade voters, in order to win the prize of controlling an elected office. This particular election is significant because, in the early stages of the electoral cycle, a signal sent by an institution, the ICC, became an important campaign issue.

Background

Before continuing, it is worth giving a brief background to the 2013 Kenyan presidential election and the ICC's involvement. The ICC's involvement in Kenyan arose because of violence that occurred in the aftermath of the previous presidential elections, in 2007. After general voting in the 2007 elections, the Kenyan Electoral Commission declared the incumbent, President Mwai Kibaki of the PNU party, the winner of the election. But supporters of the challenging candidate, Raila Odinga of the ODM party, charged that electoral results had been manipulated. The ensuing tension erupted into violence committed by supporters of Odinga against supporters of Kibaki, and vice versa. While estimates vary, it is generally thought that this violence, which occurred over several months, directly resulted in over 1,000 deaths and internally displaced as many as 600,000 people. The violence largely subsided after a UN-moderated power-sharing agreement was reached, with Kibaki remaining president and Odinga becoming prime minister.

In early 2010, the ICC's Pre-Trial Chamber granted Chief Prosecutor Louis Moreno-Ocampo permission to open an investigation into possible crimes against humanity committed during the 2007 post-electoral violence. In March of 2011, after the investigation, the ICC's Pre-Trial Chamber issued a "summons to appear" for six individuals. This list included Uhuru Kenyatta, the privately-wealthy son of one of Kenya's founding president, Jomo Kenyatta. Uhuru Kenyatta had previously been allied with the PNU and was accused of organizing and financing gang members to enact violence against supporters of the ODM. The ICC also issued a summonses for William Ruto, who was then the Education Minister, and who was accused of supporting violence by ODM

members against PNU supporters.[34]

The fact that the ICC issued summonses for Kenyatta and Ruto is notable because, in March of 2011, jockeying for the 2013 Kenyan presidential election was well under way. Both Kenyatta and Ruto had declared themselves as candidates, and Kenyatta in particular was widely considered to be leading candidates to oppose Raila Odinga, who was also a front-runner candidate. The ICC Chamber's decision forced the issue onto the national radar and into the election. There was widespread media coverage of the ICC Chamber's decision and public awareness of the issue was generally high.[35]

There is deep suspicion about the validity of the official results, but Kenyatta ultimately won in the first round of the election by the slimmest of margins. He received 50.05% of the vote, just barely reaching the 50% threshold by a mere 6,110 votes in an election with over 12 million votes officially cast.

It is also worth emphasizing that the Kenyan case is of inherent interest, even apart from the theoretical model. These indictments were the first to become a prominent issue in a major election. Previous indictments had largely targeted rebel leaders who were not in control of their country (e.g. Joseph Kony),[36] recently ousted dictators (e.g. Gaddafi in Libya), or partisans of countries not party to the Rome Statute (e.g. al-Bashir in Sudan). Kenyatta would later win the election, meaning that the ensuing ICC trials are now the first pertaining to the head of state of a country that is party to the Rome Statute. His election, combined with the indictments, has had a significant impact on Kenyan politics, both internal and external.

[34]The others who initially received summonses were, Joseph Arap Sang, a radio presenter, Francis Muthaura, Head of Civil Service, Hussein Ali, Post Master General and former head of the police, and Henry Kosgey, an ODM minister and MP. The ICC would later decide to drop the charges against Ali, Kosgey, and Muthaura.

[35]For example, a poll conducted by South Consulting in February of 2012 found that approximately 80% of people were aware of the trials, and among those citizens, 97% and 94% could identify Kenyatta and Ruto as suspects, respectively.

[36]See also: Ginsburg (2008).

Effort and Political Coalitions

The first feature of the model that I want to demonstrate empirically is that institutional signals have important effects on the effort levels of anti-compliance groups. In this case, I categorize Kenyatta and his supporters as anti-compliance actors.[37] The election represents an important way in which Kenya's compliance policy was contested. By winning the 2013 election, Kenyatta put himself in a much stronger position to resist the ICC.

In the theoretical model, effort referred to any action that had two features: (1) it was costly and (2) it increased the probability of winning the contest. This section describes a particularly important instance of how the ICC's actions changed the efforts of the politicians that were indicted. Specifically, the ICC played an important role in shaping the pre-election political coalition choices of the indicted candidates, Kenyatta and Ruto. I argue that the specter of the ICC trial helped cement their unlikely political alliance and spurred them to run on the same ticket in order to win the election in the first round. The decision of Kenyatta to join with Ruto had the two key features of effort: it was costly and it increased the probability of their side winning.[38]

Before continuing, two further background features of Kenyan politics are relevant for this argument. First, Kenyan politics are characterized by a large number of official parties, of varying sizes, who often combine to form coalitions before elections and break apart during or after the election. As the actual election date approaches, the frontrunner candidates scramble to assemble large coalitions of smaller political parties, to enhance their chances of winning. Coalition members are promised certain cabinet positions or an allocation of MP seats, and in return, they are expected to marshal their loyal voters behind the main candidates.

The 2013 electoral cycle was no exception. In December 2010, there were over 25 candidates

[37] While Kenyatta publicly states his full support for the ICC, his campaign messages and subsequent actions were decidedly anti-ICC. For example, [xx].

[38] Note, the data here are culled from secondary sources, namely Kenyan newspapers. In some contexts, effort might be easily observable or measurable. For example, in a U.S. election, an individual might have to declare substantial political contributions to a candidate. This is the focus of some of my ongoing work on the political contributions of firms affected by U.S. trade policy. In other contexts, like that of the Kenyan elections, effort is difficult to measure, especially at the individual level.

from various political parties who had expressed interest in running for president, with ten of them receiving support from at least 2% of respondents' indicated support.[39] By January 2013, however, most of them had joined together in one of a few larger coalitions. The main two coalitions were the Jubilee coalition, with Kenyatta as its presidential candidate and Ruto as vice president, and the CORD coalition with Odinga as president and Kalonzo Musyoka as vice president.

The second relevant feature concerns the rules and timing of the 2013 elections. The 2010 Kenyan constitution establishes that, to win the presidential election, a candidate must receive the most votes overall and receive at least 25% of the vote in at least 24 of the nation's 47 counties. To win in the first round of the election, the candidate must meet these two criteria and also receive over 50% of the national vote. If no candidate has over 50% of the vote or failed to meet the county criteria, there would be a runoff between the top two candidates.

When the ICC confirmed its charges against Kenyatta, Ruto, and two others, in January of 2012, they set a date for the trials to begin: April 10, 2013.[40] This timing was particularly significant, because it fell squarely in between the scheduled times for the Kenyan first round voting and second round run-off (should a run-off have proved necessary). First round voting was scheduled for March 4, 2013 with second round voting to take place a little over a month afterward. In other words, the original trial date created the possibility that, if Kenyatta and Ruto failed to win in the first round, they would be legally bound to be present at the Hague, Netherlands at the exact time they were supposed to be campaigning for a second round victory. The timing of the elections and trials therefore created a situation where Kenyatta and Ruto very much needed to ensure victory in the first round.[41]

In what ways was the Kenyatta-Ruto alliance costly to them? First, this alliance was very

[39] Infotrak 2010.

[40] The trials would later be delayed at the request of the Prosecutor. However, for the timeframe analyzed here, the actors were working under the assumption of an April trial date.

[41] Whether they would have actually left the campaign to deal with their trials, is, of course, something we'll never know. But suffice it to say, that dealing with an ICC trial concurrent with a national second round campaign would have been extremely thorny at best.

unlikely to have occurred in the absence of the ICC indictments because of the history between the two principals. In the 2007 elections, Kenyatta and Ruto had been on opposite sides of a bitter political battle that ultimately turned violent. After all, the ICC indictments were concerned with Kenyatta's alleged role in supporting violence against supporters of Ruto's party in his homeland, and Ruto was accused of supporting violence against Kenyatta's co-ethnics, the Kikuyu. To put it mildly, it is difficult to imagine two candidates representing such opposed groups with such a recent history of intense violence coming together on the same ticket. One political commentator labeled Kenyatta-Ruto "an unholy alliance,"[42] while another called the alliance "a platypus... a strange beast, consisting of two such different parts that had been thought to exist only in fantasy."[43] Macharia Munene, professor of politics at Nairobi's United States International University, said " The political alliance is a gimmick...the two individuals are in a marriage of convenience as both have questions to answer at the ICC."[44]

There had been intimations of a budding political alliance between Kenyatta and Ruto a little bit before they first were mentioned as ICC targets in late 2010. Ruto had largely fallen out with his previous political ally (Odinga), and Ruto and Kenyatta, along with (ironically, Odinga's future running mate), Musyoka, were loosely linked by the unfortunately named KKK alliance.[45] However, this was always described as a loose alliance, without formal or concrete associations.[46] When the ICC issues summonses for Kenyatta and Ruto, this alliance becomes more cemented,[47] with the two making joint appearances and overtly supporting one another. Some commentators explicitly linked the deepening of the KKK alliance with the ICC process.[48]

The alliance was also costly in more concrete terms. For Kenyatta, including Ruto in an alliance

[42]Titz, Christoph "Ethnic Violence Overshadows Kenyan Campaign," Spiegel Online International, December 13, 2012.

[43]Waweru, Daniel "The Rise of the 'Uhuruto'," *African Arguments*, December 5, 2012.

[44]*Agence France Presse*, "Unity or impunity?" December 1, 2012.

[45]The K's stand for the names of the three tribes that each person came from.

[46]The Nation (Nairobi) Imanyara Pushes for Another Attempt At Tribunal February 5, 2011

[47]Daily Nation (online) "Kenyan leader to address ethnic 'reconciliation' rally in northwestern town," January 21, 2011. Also, The Nation (Nairobi) "Leaders Back Alliance for 2012 Poll" December 5, 2010.

[48]The Star "Kenya and the ICC: Fact Versus Fiction" January 24, 2011.

required that he and his party (TNA) give up a disproportionate amount of the "spoils" of winning the election to Ruto's party (URP). He first had to accept Ruto as a running mate, though others were thought to have been his preferred choice because of similar ideologies and ethnic ties.[49] The block of voters that Ruto was expected to deliver was much smaller than the block expected to follow Kenyatta. Yet in a pre-election agreement between TNA and URP, the two sides agreed to split all public appointments and cabinet positions evenly. In reality, Ruto's party received more than half of the cabinet positions, because, as the Jubilee alliance tried to lure in additional coalition members, it was agreed that TNA would use some of its allocated cabinet positions to pursuade (purchase) additional coalition members. The two sides also divided the country into zones and agreed to not contest elections in each other's allocated zones, meaning the URP got a disproportionately large share of MPs in parliament as well.[50]

The second feature of effort in the model was that it increased that side's chances of winning the contest. In what ways did the alliance increase Kenyatta's probability of winning the contest? The most direct effect of the alliance was that Kenyatta and Ruto delivered their expected votes. There was particularly strong turnout in their home regions. Kenyatta and Ruto were able to successfully marshal these pivotal votes, in part, by using the ICC as an issue to rally their supporters. They characterized the ICC trials as outside actors trying to interfere the elections. Initially, they used thinly-veiled anti-ICC rhetoric. At a political rally in late 2012, Kenyatta told supporters that "We are telling our foreign friends that Kenya has come of age. Kenyans should be left alone to elect leaders of their choice and once that is done, they should be ready to respect Kenyans' decision."[51] Over time, Kenyatta's use of ICC-opposition as a way to rally support become more overt and direct. At a January 2013 rally, he urged supporters to use the election as "a vote of no confidence

[49]Selassie, Gus. "Election 2013: Presidential aspirants seek winning tickets in Kenya." Global Insight December 4, 2012.

[50]The Star (Nairobi) Ruto Is Big Winner in Uhuru Deal November 29, 2012.

[51]The Standard (Online) "Kenyan deputy premier, ex-minister said to sign pre-election deal 4 December," Dec. 3, 2012.

in the ICC."[52]

Polling data of public opinion on the ICC shows that this facet of Kenyatta's campaign was successful. Figure 3 plots the percent of respondents who indicated that they were happy with the ICC process over time.[53] The left pane shows the trends for Kenyatta and Ruto's home regions. The right pane shows Odinga's home region (Nyanza) and another other regions in which Odinga received overwhelming support in the 2007 elections (Western).

Public support for the ICC starts at a very high level in all regions. In three out of the four regions, over 75% of respondents said they were happy with the ICC, and even in Ruto's home region, the Rift Valley, 68% of respondents said they were happy. Over time, however, support for the ICC plummets in the Central and Rift Valley regions. On the other hand, in the two regions associated with Odinga, support for the ICC stays strong.[54] Kenyatta and Ruto's campaign appears to have been successful at blunting any negative impact of their ICC indictments in their home regions, which likely helped their electoral prospects.

Institutional Signals and *Ex Ante* Support

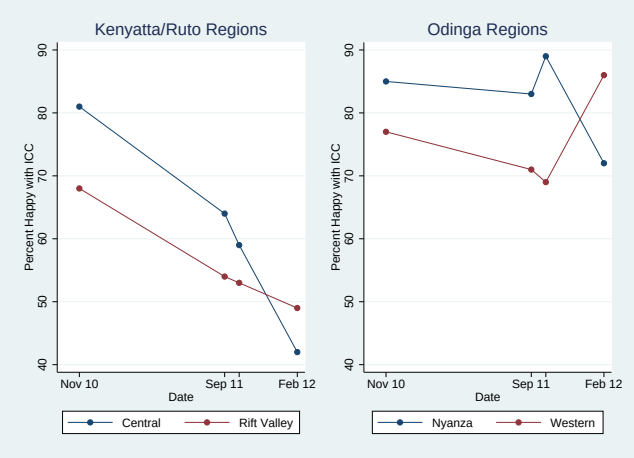
The previous section analyzed changes in effort caused by institutional signals. This section analyzes the effect of institutional signals on the probability of a particular side winning the contest. The theoretical model argued that institutional signals can increase the probability that pro-compliance groups win political contests, and that this effect is greatest "in the middle," i.e. when the *ex ante* strength of pro- and anti-compliance groups are balanced. Institutions have the weakest effect when the anti-compliance groups are stronger, *ex ante*, because any additional effort by the pro-compliance group is swamped by corresponding increases by the anti-compliance group.

[52] Africa News January 31, 2013 The Independent (Kampala).

[53] These data are from South Consulting, which asked "How happy or unhappy are you that the ICC is investigating perpetrators of post election violence?" The surveys averaged between 200 and 900 respondents per region, per survey.

[54] Interestingly, one region that was pro-Odinga in 2007, the North Eastern region, shows the same pattern as the Central and Rift Valley regions, with decreasing support for the ICC over time. However, this is likely influenced by the fact that Kenyatta and Ruto successfully recruited a prominent North Eastern MP, Charity Ngilu, to join their political alliance.

Figure 3: Support for ICC Across Regions



This figure shows the percentage of respondents who indicated that they were happy with the ICC process over time, from data collected by South Consulting.

When pro-compliance groups are very strong, *ex ante*, the institution's has little marginal effect because that group was already likely to prevail. This section assesses evidence of the non-monotonic relationship between *ex ante* support for compliance and the effect of institutional signals on support for pro- and anti-compliance groups.

Kenyan politics are well suited to assess this argument because there is significant geographic variation in *ex ante* support for political candidates. Some regions are strongly predisposed to candidates representing tribes from that region, while others are strongly opposed, and still others are in the middle. This is because voters' political preferences are very concentrated along ethnic lines.[55] Political candidates draw most of their political strength from their ability to marshal their co-ethnic voters to their cause or to support a coalition of which the candidate is a member. Ethnic groups are heavily concentrated in particular regions and political candidates are strongly associated with their home regions. Across the various regions, support for particular candidates or coalitions containing those candidates varies significantly. In exit polls of the 2007 presidential elections, Odinga won 83% of the vote in one of Kenya's 8 regions (Nyanza) compared with less

[55] Gibson and Long (2009).

than 3% in another (Central).[56]

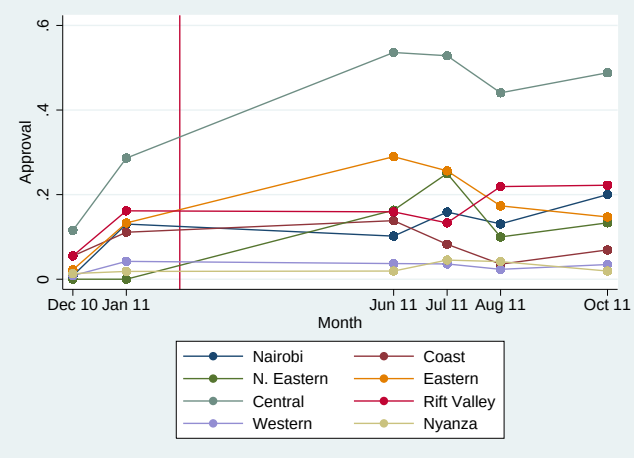
I leverage this variation to examine whether the effect of the ICC's signal on support for Kenyatta varies according to how strongly a region supported Kenyatta *before* the ICC's signal. The signal I analyze is the ICC's decision to issue summonses for Kenyatta and others, in March 2011. This event is significant because it is the first time that the ICC officially incriminates Kenyatta and because it meant that Kenyatta was very likely to go to trial. For the ICC to have the desired effect on the political contest, it should decrease support for Kenyatta. The theoretical model predicts that the ICC should decrease support for Kenyatta most in regions where he and Odinga's support is balanced, *ex ante*. And the ICC should have decrease support less in regions where either Kenyatta or Odinga enjoy very strong support, *ex ante*.

Figure 4 shows public support for Kenyatta, across regions, before and after the ICC's summonses against him. The figure summarizes data from a set of nationally representative polls conducted by Infotrak, a Harris-affiliated Kenyan polling firm. At irregular times, Infotrak conducted public opinion polling in each region of the country. One item on their survey asked, "Apart from President Kibaki, whom would you vote for as your President if presidential elections were held today?" Respondents then indicated their preferred presidential candidate. Each of the six surveys polled an average of 1,300 respondents. Each dot in Figure 4 show the percentage of respondents from that region who indicated that Kenyatta was their most preferred candidate during that particular survey.

Figure 4 shows the two features that make the data particularly useful for testing the theory. First, there is regional variation in the level of support for Kenyatta *ex ante*, before the summonses. In the Central region, which is home to many of Kenyatta's Kikuyu tribe members, he enjoyed his strongest support. In regions like Nyanza, home to a substantial number of Odinga's Luo tribe, and the Western and North Eastern regions, which strongly supported Odinga in the 2007 elections, Kenyatta receives virtually no support. Other regions fall somewhere in between.

[56] Gibson and Long (2009).

Figure 4: Kenyatta Support by Region Over Time



Percent of respondents in each survey answering that Kenyatta was their most preferred candidate for six surveys. Surveys were conducted by Infotrak in Kenya.

Second, there is variation across regions in how, if at all, support for Kenyatta changes after the summonses. If the ICC's signal had a uniform effect on support for Kenyatta, then we would expect a uniform dip in his support. However, consistent with the theory, the effect of the signal appears smallest in regions with very high or low *ex ante* support and highest in regions with middling support. In the Central region, the summons appear to slow the rise in Kenyatta's popularity. In the Western and Nyanza regions, his low level of support stays low or dips slightly. In the middling regions, Kenyatta's rising popularity plateaus or even drops, as in the Coastal region.

This pattern is also present in statistical analysis of the effect of the ICC's signal on the probability that an individual respondent indicates support for Kenyatta over time. Analyzing the effect of the ICC on respondents' choice of preferred candidate is difficult because we can only ever observe one world- the world in which the ICC *did* issue summonses. The challenge thus lies in constructing an estimate of the relevant counterfactual- how that respondent might have answered if the ICC had not issued summonses.

To gain leverage over this counterfactual, I first use the data from the surveys occurring before the ICC's signal to construct an estimate of an individual's latent support for Kenyatta, absent

the signal. In other words, I use the data from the pre-ICC signal surveys to "train" a model that predicts whether a respondent will support Kenyatta based on that respondent's observed characteristics, like where they live and other demographic variables. I then compare this predicted level of support to whether the individual indicated that Kenyatta was their most preferred candidate after the signal. Any effect of the ICC's signal should be found in the difference between the predicted and observed levels of support.

To be sure, this approach requires assumptions that are not testable. It is possible that events other than the ICC summons affected a respondent's likelihood of choosing Kenyatta as their preferred candidate. However, these confounding events would have to have two features. They would have to be important events that occurred in the timeframe in question- between January and June of 2011. The effect of the event on respondents would also have to vary with *ex ante* support in the same way predicted by the theory. For example, a non-ICC event that increased or decreased Kenyatta's popularity equally across regions would not explain these results.

Let k_i be an indicator variable that equals 1 if respondent i chose Kenyatta as their most preferred candidate. The surveys also asked a variety of demographic questions, like the respondent's sex, age, whether they lived in an urban or rural area, their religion, and their region of residence.[57] Let X_i denote the matrix containing these variables (excluding region), where each row corresponds to a particular respondent. Let r_i^j be binary variable that equals 1 if respondent i lives in region j, and zero otherwise.[58]

I first use the data from the pre-summons surveys to estimate a probit regression.[59] The regres-

[57]Different surveys asked different combinations of questions. These were the questions that were common to all the surveys analyzed.

[58]There are 8 regions in Kenya as mentioned above. However, in the December 2010 and January 2011 surveys, no one in the North Eastern region indicated that Kenyatta was their most preferred candidate. This lack of variation means that I can't construct estimates for the parameters related to that region. For the seven remaining regions, the index for the regions goes from $j = 1$ to $j = 6$ to indicate that the seventh region is withheld as the base category.

[59]I did not use matching here or a simple post-ICC-summonses dummy variable because there are obvious time trends in the data. The time period in question is early on in the electoral cycle, and candidate support is in flux. I want to be able to leverage information about the rate at with Kenyatta's support is increasing in particular regions. It would be difficult to assume balance on unobservables if there is an unobserved trend. [xx- Try polywog].

sion models k_i as a function of the respondent's observed demographic characteristics, the set of region indicators, and a region-specific time trend, as in Equation 1. t is a counter variable that starts at zero and measures the month that the respondent was surveyed in, beginning in December 2010.[60]

$$k_i^* = X_i\beta + \sum_{j=1}^{6} \gamma_j r_i^j + \sum_{j=1}^{6} (\delta_j r_i^j * t) + \epsilon_i \tag{1}$$

$$k_i = \begin{cases} 1 & \text{if } k_i^* > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Denote the resulting vector of coefficients as $\hat{\beta}$. As a slight abuse of notation, I will refer to all the estimated coefficients and explanators (including the region indicators and trends) as $\hat{\beta}$ and X_i. For each individual in the four surveys taking place after the ICC's summonses, I calculate the respondent's predicted "level" of support for Kenyatta, $\hat{k}_i = X_i\hat{\beta}$. I call this their predicted level of support to denote that it is the linear prediction from the coefficients derived from the two pre-summonses surveys, rather than a predicted probability of choosing Kenyatta as their most preferred candidate. $\hat{k}_i$ describes the individual's latent support for Kenyatta, as predicted by the covariates observed for that individual.

For each of the post-ICC respondents, I then calculate a measure of the degree to which the pre-ICC model over- or under-predicts that individual's support for Kenyatta. To construct this measure, I use a particular form of the difference between the individual's predicted level of support and observed electoral preference. Specifically, I construct this "residual:" $r_i = \Phi(\hat{k}_i) - k_i$, where Φ indicates the cumulative standard normal distribution function.

Higher values of r_i indicate the ICC had a greater effect in lowering that individual's support for Kenyatta. To see why this quantity captures possible ICC effects, consider an individual in

[60] ϵ_i are assumed to be distributed i.i.d., standard normal.

a post-event survey who did *not* choose Kenyatta, i.e. $k_i = 0$. In this case, r_i is positive by construction. The magnitude of r_i gives a measure of how surprised we are that they did not select Kenyatta, based on the coefficient estimates from the model of the pre-ICC data. The pre-ICC model predicts the probability that individual would have chosen Kenyatta, $\Phi(\hat{k}_i)$, and compares that to the individual's observed choice. Conversely, if the post-ICC individual in question did select Kenyatta, i.e. $k_i = 1$, then $r_i < 0$, by construction, and the residual measures the degree to which the individual's observed support for Kenyatta is higher than expected.

I thus have the two pieces necessary to test the prediction of the theoretical model. I have a prediction of the individual's *ex ante* level of support for Kenyatta, $\hat{k}_i$, and a measure of the effect of the ICC summonses, the difference between the individual's observed and predicted support, r_i. The theory predicts that the residuals should be higher for individuals in the middle of the distribution for supporting Kenyatta and lower for those that are either very likely or very unlikely to support Kenyatta.

Figure 5 - Figure 7 show how r_i varies with $\hat{k}_i$.[61] Each figure focuses the window of analysis closer and closer to the date of the summonses. Figure 5 uses all four of the post-ICC surveys. Figure 6 uses only the two surveys that occurred immediately after the ICC summonses, in June and July of 2011. Figure 7 only uses the June 2011 survey. I present this sequence of Figures because the potential for confounding events decreases as there is less time in between the ICC event and observed support for Kenyatta after the event. Zooming in in this way also has the benefit of muting the effect of the region specific time trends on predicted support for Kenyatta. Since there are only two surveys before the ICC event, I can only estimate linear time trends, and in the later months this causes the predicted levels of support for Kenyatta to be artificially high. The residuals are generally positive, meaning that observed support for Kenyatta is generally weaker than predicted.

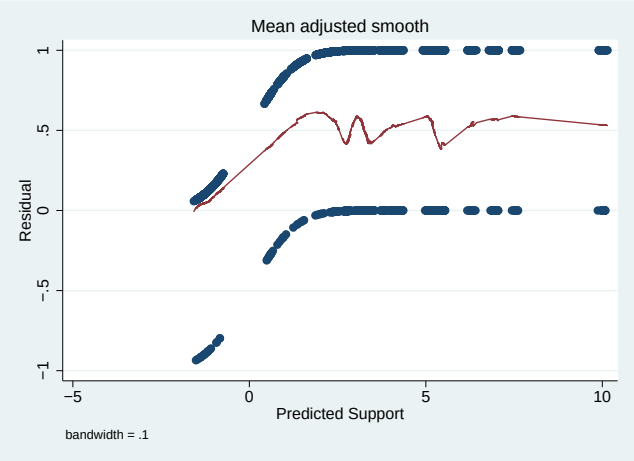
[61] Since each dot in the figure corresponds to an individual respondent, and since the dots tend to be tightly clustered, I use a Loess smoother to help show overall trends. The smoothing algorithm is constrained so that the mean of the smoothed values equals the mean of the values on the vertical axis.

The model's theoretical predictions receive support, though with one caveat. As predicted, the estimated effect of the ICC is highest for individuals who fall in the middle of the distribution of predicted support for Kenyatta. The highest estimated effect of the ICC is found in the Western region, which was considered to be a "swing region" in the election. The estimated effect of the ICC is also particularly low for individuals on the lower end of this distribution. If the pre-ICC empirical model predicted that you were very unlikely to choose Kenyatta, then there is not evidence that the ICC lowered your support for Kenyatta. For example, the estimated effect of the ICC is lowest in Nyanza, which is Odinga's homeland and the base of his support.

However, the second part of the non-monotonic relationship predicted by the theory receives only partial support. In some regions that we would expect to have extremely high support for Kenyatta, the effect of the ICC is low, as predicted. For example, the estimated effect of the ICC is low in the Central region, which is Kenyatta's homeland. However, I say that this part of the prediction receives only partial support because, looking at the right hand side of the Figures, individuals on the high end of the predicted support distribution also appear to have higher ICC effects. For those who appeared most likely to support Kenyatta, the ICC seems to also have lowered their support, almost as much as those in the middle of the predicted support distribution.

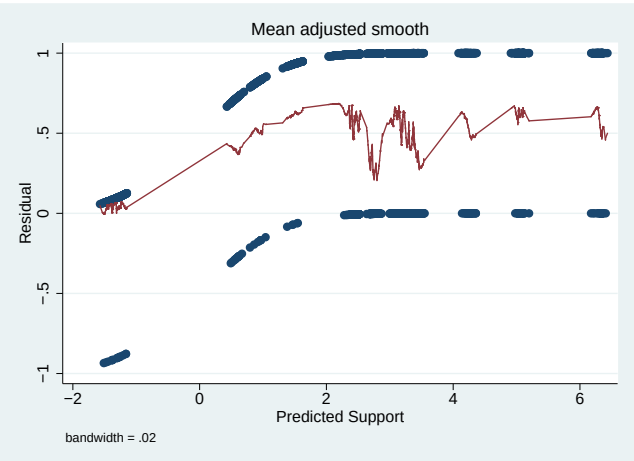
This is largely driven by the responses of individuals in Nairobi, where the pre-ICC empirical model predicts very strong support for Kenyatta. The model likely over-predicts this support to a greater degree than in other regions because of the region-specific time trends. Before the summonses, support for Kenyatta in Nairobi goes from less than 1% to over 13%. This is likely an over-exaggerated swing. Figure 8 replicates the Figure 7, the figure using only the June 2011 survey, and excludes respondents living in Nairobi. The smoothed line is more downward bending on the right hand, at higher predicted values of support for Kenyatta, as predicted by the theory.

Figure 5: Predicted versus Actual Support, All Post-Event Surveys



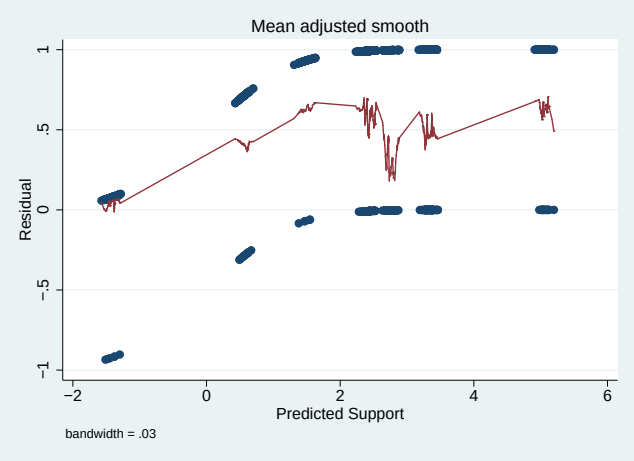
The horizontal axis is the linear prediction of latent support for Kenyatta using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 6: Predicted versus Actual Support, First Two Post-Event Surveys Only



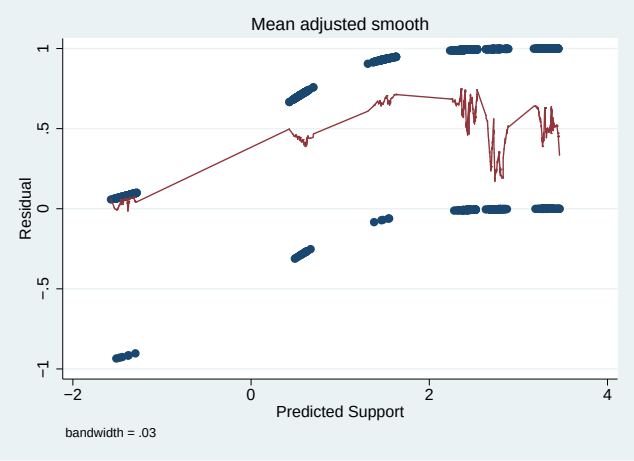
This figure only uses data from the June and July 2011 surveys. The horizontal axis is the linear prediction of latent support for Kenyatta using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 7: Predicted versus Actual Support, First Post-Event Survey Only



This figure only uses data from the June 2011 survey. The horizontal axis is the linear prediction of latent support for Kenyatta using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 8: Predicted versus Actual Support, First Post-Event Survey Only, Excluding Nairobi



This figure only uses data from the June 2011 survey and excludes respondents from Nairobi. The horizontal axis is the linear prediction of latent support for Kenyatta using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Conclusion

A large and valuable body of existing work argues that international institutions are a force for compliance because they help mobilize domestic groups who support compliance. This paper developed a theory in which international institutions can affect the behavior of both pro- and anti-compliance domestic groups. Institutions can influence one group's value to affecting compliance policy, which can cause both groups to exert more or less costly effort. This theory was general to any issue area governed by international institutions in which there are groups with divergent preferences over whether or how to comply with institutional rules. The theory was also general to types of effort and regime types. Effort described any actions that domestic groups could take that were costly and increased their probability of winning influence over a contest. This could describe everything from campaign contributions in a democracy to protests in an autocracy.

The theory predicted that institutional signals should increase the efforts of anti-compliance groups and that signals should have the greatest marginal effect when strength of pro- and anti-compliance groups was balanced, *ex ante*. I found empirical support for the first prediction by tracing how the ICC's indictment of two Kenyan politicians during the 2013 presidential campaign cemented an alliance between the two and helped them rally supporters against the ICC. I found support the second prediction by analyzing individual level data, showing that the effect of the ICC on support for the main indicted candidate was greatest in regions where he would have otherwise expected middling support.

The policy implications of this research are significant. International institutions often focus on the "worst of the worst" violators of institutional rules, perhaps (optimistically) because of their altruistic desire to do good where it is needed most or perhaps (cynically) because of their desire for additional prestige or resources. This research suggests that while targeting the worst violations is admirable, it might not be the best course of action from the most important perspective: achieving compliance where it wouldn't otherwise occur. Proponents of international law right-

fully argue that achieving widespread compliance is a long-term journey. However, to the extent that international law desires to affect compliance *now*, its proponents should focus on easier cases where pro- and anti-compliance groups are balanced, rather than hard cases. Recently, Emilie Hafner-Burton (2013) suggested that steward countries should "triage" the cases of international human rights violations, and focus on those where they are most likely to have an effect. This research is very much in line with that call, and suggests possible ways to think about the likelihood of success. Rather than tilt at windmills, international institutions and their proponents might better benefit from focusing on cases where they can most effectively sway political contests towards compliance.

References

Becker, Gary S. 1983. “A Theory of Competition Among Pressure Groups for Political Influence.” *The Quarterly Journal of Economics* 98(3):371–400.

Buthe, Tim and Helen V. Milner. 2008. “The Politics of Foreign Direct Investment into Developing Countries: Increasing FDI through International Trade Agreements?” *American Journal of Political Science* 52(4):741–762.

Carrubba, Clifford J. 2005. “Courts and Compliance in International Regulatory Regimes.” *The Journal of Politics* 67(3):669–689.

Carrubba, Clifford James. 2009. “A Model of the Endogenous Development of Judicial Institutions in Federal and International Systems.” *The Journal of Politics* 71(01):55–69.

Chapman, Terrence L. 2007. “International Security Institutions, Domestic Politics, and Institutional Legitimacy.” *Journal of Conflict Resolution* 51(1):134–166.

Corchón, LuisC. 2007. “The theory of contests: a survey.” *Review of Economic Design* 11(2):69–100.

Dai, Xinyuan. 2005. “Why Comply? The Domestic Constituency Mechanism.” *International Organization* 59(02):363–398.

Dai, Xinyuan. 2006. “The Conditional Nature of Democratic Compliance.” *Journal of Conflict Resolution* 50(5):690–713.

Ehrlich, Sean D. 2007. “Access to Protection: Domestic Institutions and Trade Policy in Democracies.” *International Organization* 61(3):571–605.

Elkins, Zachary, Andrew T. Guzman and Beth A. Simmons. 2006. “Competing for Capital: The

Diffusion of Bilateral Investment Treaties, 1960-2000." *International Organization* 60(4):pp. 811–846.

Esteban, Joan and Debraj Ray. 2001. "Collective Action and the Group Size Paradox." *The American Political Science Review* 95(3):pp. 663–672.

Fang, Songying. 2008. "The Informational Role of International Institutions and Domestic Politics." *American Journal of Political Science* 52(2):304–321.

Garfinkel, Michelle R. and Stergios Skaperdas. 2000. "Conflict without Misperceptions or Incomplete Information: How the Future Matters." *Journal of Conflict Resolution* 44(6):793–807.

Gibson, Clark C. and James D. Long. 2009. "The presidential and parliamentary elections in Kenya, December 2007." *Electoral Studies* 28:497–502.

Ginsburg, Tom. 2008. "The Clash of Commitments at the International Criminal Court." *Chicago Journal of International Law* 9:499–514.

Hafner-Burton, E. M. 2013. *Making Human Rights a Reality*. Princeton University Press.

Hegre, Håvard. 2004. "Size Asymmetry, Trade, and Militarized Conflict." *Journal of Conflict Resolution* 48(3):403–429.

Jia, Hao. 2008. "A stochastic derivation of the ratio form of contest success functions." *Public Choice* 135:125–130.

Keith, Linda Camp. 1999. "The United Nations International Covenant on Civil and Political Rights: Does It Make a Difference in Human Rights Behavior?" *Journal of Peace Research* 36(1):pp. 95–118.

Keith, Linda Camp, C. Neal Tate and Steven C. Poe. 2009. "Is The Law a Mere Parchment Barrier to Human Rights Abuse?" *The Journal of Politics* 71(02):644–660.

Kono, Daniel Y. 2006. "Optimal Obfuscation: Democracy and Trade Policy Transparency." *The American Political Science Review* 100(3):369–384.

Krikorian, Jacqueline D. 2005. *Managing the challenges of WTO participation: 45 case studies*. Cambridge University Press chapter Canada and the WTO: Multilevel Governance, Public Policy-Making and the WTO Auto Pact Case, pp. 134–149.

Krueger, Anne O. 1974. "The Political Economy of the Rent-Seeking Society." *The American Economic Review* 64(3):pp. 291–303.

Lohmann, Susanne. 2003. "Why Do Institutions Matter An AudienceCost Theory of Institutional Commitment." *Governance* 16.

Mansfield, Edward D. and Eric Reinhardt. 2008. "International Institutions and the Volatility of International Trade." *International Organization* 62(4):pp. 621–652.

Mansfield, Edward D., Helen V. Milner and B. Peter Rosendorff. 2000. "Free to Trade: Democracies, Autocracies, and International Trade." *The American Political Science Review* 94(2):305–321.

Mansfield, Edward D., Helen V. Milner and B. Peter Rosendorff. 2002. "Why Democracies Cooperate More: Electoral Control and International Trade Agreements." *International Organization* 56(3):477–513.

Marfording, Annette. 1997. "Cultural Relativism and the Construction of Culture: An Examination of Japan." *Human Rights Quarterly* 19(2):431–448.

Mehta, Pratap Bhanu. 2004. "Hinduism and Self-Rule." *Journal of Democracy* 15(3):108–121.

Pelc, Krzysztof. Forthcoming. "Googline the WTO: What Search Engine Data Tell Us About the Political Economy of Institutions." *International Organization* .

Pevehouse, Jon C. 2002. "Democracy from the Outside-In? International Organizations and Democratization." *International Organization* 56(03):515–549.

Poe, Steven C. and C. Neal Tate. 1994. "Repression of Human Rights to Personal Integrity in the 1980s: A Global Analysis." *The American Political Science Review* 88(4):pp. 853–872.

Poe, Steven C., C. Neal Tate and Linda Camp Keith. 1999. "Repression of the Human Right to Personal Integrity Revisited: A Global Cross-National Study Covering the Years 1976-1993." *International Studies Quarterly* 43(2):pp. 291–313.

Rosendorff, B. P. 2005. "Stability and Rigidity: Politics and Design of the WTO's Dispute Settlement Procedure." *American Political Science Review* 99(03):389–400.

Simmons, Beth A. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics*. Cambridge University Press.

Simmons, Beth A. 2010. "Treaty Compliance and Violation." *Annual Review of Political Science* 13:273–296.

Simmons, Beth A. and Allison Danner. 2010. "Credible Commitments and the International Criminal Court." *International Organization* 64(02):225–256.

Skaperdas, Stergios. 1996. "Contest success functions." *Economic Theory* 7:283–290.

Slantchev, Branislav L. 2010. "Feigning Weakness." *International Organization* 64(03):357–388.

Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(04):821–840.

Tomz, Michael. 2008. "Reputation and the Effect of International Law on Preferences and Beliefs." Stanford University.

Tullock, Gordon. 1967. "The Welfare Cost of Tariffs, Monopolies and Theft." *Western Economic Journal* 5(3):224–232.

Technical Appendix

This appendix contains the proofs for the formal model. For simplicity, I first characterize optimal effort levels in any subgame perfect Nash equilibrium. As noted in the text, this proofing strategy follows Corchón (2007). I then include the institution and show existence of the equilibrium I discuss.

Optimal Effort Levels

First, redefine citizen i's optimization problem as follows, by dividing her payoffs by V_i:

$$max_{e_i} \Pi_i(e_i, e_j)$$
$$max_{e_i} \frac{e_i}{e_i+e_j} V_i - c_i * e_i$$
$$max_{e_i} \frac{e_i}{e_i+e_j} - \frac{c_i}{V_i} e_i$$
$$max_{e_i} \frac{e_i}{e_i+e_j} - d_i e_i$$

Differentiating with respect to e_i yields:

$$\frac{e_j}{(e_i+e_j)^2} = d_i$$

Note, summing the two citizens' first order conditions and simplifying yields:

$$\frac{e_i+e_j}{(e_i+e_j)^2} = d_i + d_j$$

$$e_i + e_j = \frac{1}{d_i+d_j}$$

Using equation this summation and the first order condition yields e_i^* as a function of d_i and d_j and 2.

$$e_i^* = d_j(e_i + e_j)^2$$

(From the FOC)

$$e_i^* = d_j[(\frac{1}{d_i+d_j})^2]$$

(From the summation, substituting)

$$e_i^* = \frac{d_j}{(d_i+d_j)^2}$$

For Corollary 1, this expression generates comparative statics relating d_i (and by extension c_i and V_i) to the optimal effort level, e_i^*.

$$\frac{\partial e_i^*}{\partial d_i} = \frac{-2d_j}{(d_i+d_j)^3}$$

We can also generate comparative statics relating d_j to i's optimal effort level.

$$\frac{\partial e_i^*}{\partial d_j} = \frac{d_i - dj}{(d_i+d_j)^3}$$

Substituting the optimal effort levels into the contest success function and simplifying yields 3. Taking derivatives yields Corollary 2.

Optimal Effort Levels With/Without Institutional Signal

We can express the effects of an institutional signal (or absence of signal) by using the results above and incorporating the effect of the signal on the PC citizen's expected value of winning the contest. Recall, the PC citizen's prior expected value to winning is $V_{PC} = pv_{PC}$, and her "prior" d_{PC} is $d_{PC} = \frac{c_{PC}}{pv_{PC}}$

For ease of notation, let $\gamma' \equiv \frac{1+2pq-q-p}{q}$ and $\gamma'' \equiv \frac{p+q-2pq}{1-q}$. Using Bayes rule, the PC citizen's updated beliefs that compliance is beneficial, after a signal are:

$$Pr(B|S) = \frac{pq}{pq+(1-p)(1-q)}$$
$$d'_{PC} = d_{PC}\frac{1+2pq-q-p}{q}$$

Using this expression, we can write the PC citizen's "updated" d_{PC} as:

$$d'_{PC} = d_{PC}\frac{1+2pq-q-p}{q}$$
$$d'_{PC} = d_{PC}\gamma'$$

Similarly, when no signal is sent, the pro-compliance honest citizen updates her beliefs and expected value, denoted d''_{PC}.

$$Pr(B|\sim S) = \frac{p(1-q)}{p(1-q)+(1-p)q}$$
$$d''_{PC} = d_{PC}\frac{p+q-2pq}{1-q}$$
$$d''_{PC} = d_{PC}\gamma''$$

This allows us to simplify the optimal effort levels of the PC and AC citizen, with and without the signal.

$$e'^{*}_{PC} = \frac{d_{AC}}{(\gamma' d_{PC}+d_{AC})^2}$$
$$e'^{*}_{AC} = \frac{d_{PC}}{(\gamma' d_{PC}+d_{AC})^2}$$

$$e''^{*}_{PC} = \frac{d_{AC}}{(\gamma'' d_{PC}+d_{AC})^2}$$
$$e''^{*}_{AC} = \frac{d_{PC}}{(\gamma'' d_{PC}+d_{AC})^2}$$

Equilibrium Winning Probabilities and Institutional Utility

The equilibrium winning probabilities described in 3 and in Corollary 2 follow directly from the optimal effort levels described above and the contest success function. And this expression is general to any d, so it can be modified to account for institutional signals by adding the appropriate γ to the appropriate place.

$$\phi_i(e^*_i, e^*_j) = \frac{d_j}{d_i+d_j}$$

The institution's expected utility for sending a signal given that it gets a positive private signal is:

$$EU_I(S|b) = Pr(B|b)\phi_{PC}(e^{*\prime}_{PC}, e^{*\prime}_{AC})V_I - \phi_{AC}(e^{*\prime}_{PC}, e^{*\prime}_{AC})l - k$$

The institution's expected utility for not sending a signal given a positive private signal is:

$$EU_I(\sim S|b) = Pr(B|b)\phi_{PC}(e^{*\prime\prime}_{PC}, e^{*\prime\prime}_{AC})V_I$$

Combing these two expressions yields 4:

$$EU_I(S|b) - EU_I(\sim S|b) == Pr(B|b)[\phi_{PC}(e^{*\prime}_{PC}, e^{*\prime}_{AC}) - \phi_{PC}(e^{*\prime\prime}_{PC}, e^{*\prime\prime}_{AC})]V_I - \phi_{AC}(e^{*\prime}_{PC}, e^{*\prime}_{AC})l - k$$

Taking derivatives and simplifying (and noting that $Pr(B|b) = \frac{p}{\gamma'}$) yields Corollary 3:
$\frac{\partial EU_I(S) - EU_I(\sim S)}{\partial d_{PC}} = \frac{pV_I}{\gamma'}[\frac{\gamma'' d_{AC}}{(\gamma'' d_{PC} + d_{AC})^2} - \frac{\gamma' d_{AC}}{(\gamma' d_{PC} + d_{AC})^2}] - l[\frac{\gamma'}{\gamma' d_{PC} + d_{AC}} - \frac{(\gamma')^2 d_{PC}}{(\gamma' d_{PC} + d_{AC})^2}]$

Existence of Equilibrium

The conditions for the existence of the equilibrium in 1 are:

- $q > \frac{1}{2}$
- $Pr(B|b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I \geq \phi_{AC}(e'_{PC}, e'_{AC})l + k \geq Pr(B|\sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I$
- $Pr(B|b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_{PC} \geq c_{PC}(e'_{PC} - e''_{PC}) \geq Pr(B|\sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_{PC}$

The second condition comes from the institution's decision (a) to send the signal when they receive a positive private signal:

$$EU_I(S|b) \geq EU_I(\sim S|b)$$
$$Pr(B|b)\phi_{PC}(e'_{PC}, e'_{AC})V_I - \phi_{AC}(e'_{PC}, e'_{AC})l - k \geq Pr(B|b)\phi_{PC}(e''_{PC}, e''_{AC})V_I$$
$$\phi_{AC}(e'_{PC}, e'_{AC})l + k \geq Pr(B|\sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I$$

and (b) to not send the signal when they do not receive a positive private signal:

$$EU_I(\sim S| \sim b) \geq EU_I(S| \sim b)$$

$$Pr(B| \sim b)\phi_{PC}(e''_{PC}, e''_{AC})V_I \geq Pr(B| \sim b)\phi_{PC}(e'_{PC}, e'_{AC})V_I - \phi_{AC}(e'_{PC}, e'_{AC})l - k$$

$$\phi_{AC}(e'_{PC}, e'_{AC})l + k \geq Pr(B| \sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I$$

Combining conditions (a) and (b) yields:

$$Pr(B|b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I \geq \phi_{AC}(e'_{PC}, e'_{AC})l + k \geq Pr(B| \sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I$$

Note that $q > \frac{1}{2}$ ensures that $Pr(B|b) > Pr(B| \sim B)$. Also, Proposition [xx] guarantees that $\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC}) > 0$. Since $V_I > 0$, there exist a pair $\{l, k\}$ small enough for both conditions (2) to be met.

The third condition is similar, but for the PC citizen. It says that the PC citizen wants to exert "high effort" iff they observe a positive institutional signal and low effort iff they do not observe this signal. The two analogous expressions are (a):

$$EU_{PC}(e'^*_{PC}|b) \geq EU_{PC}(e''^*_{PC}|b)$$

$$Pr(B|b)\phi_{PC}(e'_{PC}, e'_{AC})V_{PC} - c_{PC}e'^*_{PC} \geq Pr(B|b)\phi_{PC}(e''_{PC}, e''_{AC})V_{PC} - c_{PC}e''^*_{PC}$$

and (b):

$$EU_{PC}(e''^*_{PC}| \sim b) \geq EU_{PC}(e'^*_{PC}| \sim b)$$

$$Pr(B| \sim b)\phi_{PC}(e''_{PC}, e''_{AC})V_{PC} - c_{PC}e''^*_{PC} \geq Pr(B|b)\phi_{PC}(e'_{PC}, e'_{AC})V_{PC} - c_{PC}e'^*_{PC}$$

Conditions (a) and (b) combine for condition (3) above. Meeting this condition requires that the costs of effort, relative to the value of winning the prize, be "just right." They have to be small enough to allow the PC citizen to increase her effort after a signal and large enough to keep her from simply exerting that high effort level regardless of the signal.

$$Pr(B|b)[\phi_{PC}(e'^*_{PC}, e'^*_{AC}) - \phi_{PC}(e''^*_{PC}, e''^*_{AC})]V_{PC} \geq c_{PC}(e'^*_{PC} - e''^*_{PC} \geq Pr(B| \sim b)[\phi_{PC}(e'^*_{PC}, e'^*_{AC}) - \phi_{PC}(e''^*_{PC}, e''^*_{AC})]V_{PC}$$