Rational Remedies: The Role of Opinion Clarity in the Inter-American Human Rights System.*

Jeffrey K. Staton† Emory University

Alexia Romero‡ Google

March 31, 2019

Abstract

International courts have far from perfect records of compliance. States routinely delay the implementation of policy changes necessary to come into line with international obligations. Some judicial orders are simply ignored in their entirety. Yet judicial orders aimed at potentially recalcitrant states often vaguely express what is required, and thus create conditions for delay and defiance. This paper leverages a detailed, public monitoring system for decisions of the Inter-American Court of Human Rights to evaluate a model of judicial opinion writing that connects the informational challenges associated with effectuating significant policy change to the language that judges adopt in their orders and ultimately the reactions of states. Our results suggest that uncertainty about how precisely to bring about a policy change influences compliance by reducing the clarity of judicial orders. Flexibility in language permits judges to trade-off maximal pressure for compliance for the ability to leverage local knowledge about how to bring a state in line with its international obligations. From this perspective non-compliant outcomes are not necessarily a clear signal of weak judicial institutions but instead a natural piece of the process by which judges manage difficult policy-making tasks.

---

*We would like to thank Karen Alter, Tom Clark, Michael Coppedge, Varun Gauri, Emilie Hafner-Burton, Laurence Helfer, Courtney Hillebrecht, Alexandra Huneeus, Diana Kapiszewski, Olof Larsson, Yonatan Lupu, Daniel Naurin, Emilia Powell, Emily Ritter, Eric Voeten, and Georg Vanberg, for valuable feedback on our paper. We also wish to acknowledge the special encouragement and assistance that we received from Will Moore over many years and many iterations of this paper.

†Jeffrey K. Staton is Professor of Political Science, Emory University.

‡Alexia Romero is Investigations Counsel in Ethics and Compliance, Google.

The study of Latin American politics is the study of judicialized politics. Judges across the region routinely shape the production of public policy where once executives, legislatures and bureaucracies dominated (e.g. Sieder, Schjolden and Angell, 2016; Wilson, 2009; Ríos-Figueroa, 2016). The study of Latin American politics is also the study of transnational and international politics. Most obviously, the protection of human rights in Latin American political systems is regulated by overlapping, often complementary, domestic and international legal obligations, interpreted by judges at both the domestic and international levels, and promoted by international advocacy networks (e.g. Brewer-Carías, 2009; Dulitzky, 2015; Huneeus, 2011; Binder, 2012; Sikkink, 2005). For these reasons, studies of human rights in Latin America often illuminate the complex ways in which policy is produced in a system that is neither fully international nor fully domestic. And yet studies of human rights in Latin America also highlight core features of politics, which transcend the international and domestic levels as well as the complications that naturally follow from the interconnectedness of a transnational legal system.

One of these features is that courts empowered to review a state's policies for consistency with their legal obligations are tasked with two inter-related challenges. They craft rules that detail how to evaluate policies in light of a substantive interpretation of the obligation; and, having found a violation of this obligation, courts often construct remedies in the form of direct orders to states explaining how to bring their policies in line with the obligation. Both of these tasks, especially the remedial task, require the resolution of a "means-ends" problem. A judge may know what outcome she desires, but she may be uncertain about how to ensure that this outcome is realized. Importantly, although common concepts of the rule of law value legal clarity,[1] judges do not always clearly express how they expect states to remedy their violations (Lax, 2012; Schlag, 1985). Deadlines for policy change are set indefinitely; general activities are ordered, as in "investigate this crime" or "build this program," but the details of how precisely to do this are left largely to the state's imagination. The flexibility with which orders can be expressed in natural language offers judges a measure of control over the process linking legal obligations and legal orders to policy outcomes. Vagueness can be used to provide state officials with a measure of discretion, allowing them freedom to use local knowledge about particular processes and actors to reach outcomes

[1]See discussions in McCubbins, Rodriguez and Weingast (2009) and Waldron (2002).

judges desire but would struggle to produce absent better information. Yet, because vagueness invites delays, resistance and sometimes outright defiance (Spriggs, 1997), this ability to manage the means-ends problem through vagueness invites a tradeoff. Informational gains must be traded off against increasing chances of non-compliance.

Our paper considers this tradeoff in the context of decisions of the Inter-American Court of Human Rights (IACtHR or simply the Court) and state reactions to its orders. We leverage a rich source of data made available by the Court's compliance monitoring process to consider implications of a theoretical model designed precisely to evaluate the choice of remedies in political contexts where compliance with judicial orders cannot be assumed. The model that guides us envisions two roles for vagueness. Vague orders permit judges to take advantage of policy expertise in state bureaucracies that can aid in the translation of a legal obligation to a policy outcome. In this way, vagueness operates as does "discretion" in models of legislation (e.g. Huber and Shipan, 2002) as well as does "flexibly" in arguments about the rational design of international institutions (e.g. Koremenos, Lipson and Snidal, 2001). Vague orders also may attenuate the potential costs judges perceive to be associated with non-compliance. All else equal, we might presume that judges prefer to have their decisions respected; and, it is certainly possible that non-compliance, especially so if frequent, might undermine judicial legitimacy or judicial power itself (Carrubba, 2009). Vagueness can attenuate these costs by making it difficult to argue that an order has actually been defied. We find that compliance with the Court's orders demanding reparations for international law violations is more likely when the Court clearly expresses what it requires from the state. We also find that the Court is less likely to issue clear orders when the policy change it desires demands information that the court may not perfectly possess. The policy challenges the Court confronts, mediated by its choice of legal language, condition compliance outcomes in meaningful ways.

Our results provide a new interpretation of international courts' compliance records, including the IACtHR's record, which are often far from perfect (Hillebrecht, 2014; Hawkins and Jacoby, 2010; Basch et al., 2010; Huneeus, 2011). The considerable degree of non-compliance with international judicial orders is not necessarily clear evidence of weak judicial institutions or of a fundamental problem with the rule of law in the international context. It is possible that less than perfect compliance records are better understood as the natural consequence of the substantial policy-making problems and related informational challenges that international judges confront. Far from

perfect compliance records are held by many domestic courts, even those in states characterized by high levels of the rule of law (Vanberg, 2005). Our study suggests that the Court adopts less clear remedies under conditions of uncertainty in part because local practices and knowledge can be essential for turning a policy demand into a desirable policy change. This approach no doubt invites non-compliance, but that may be an acceptable consequence of drafting remedies in uncertain contexts. From this perspective, the Court's imperfect compliance record follows from tasking any judicial body, domestic or international, with challenges that naturally result in the kind of judicial language that encourages resistance.

Our study also has implications for an approach to judicial decision-making that encourages a "dialogue" between courts and political officials (e.g., legislatures, governments, bureaucrats, etc) over the meaning of legal texts and over the means by which legal obligations can be best enforced. Legal scholars from multiple traditions have wrestled with the democratic implications of judicial review. Developments in constitutionalism in Canada, New Zealand and the United Kingdom during the 1990s resulted in a model that seeks to combine judicial protections for fundamental rights and freedoms with parliamentary sovereignty (Hogg and Bushell, 1997). In this new Commonwealth model of constitutionalism, courts and policymakers are meant to jointly interpret the meaning of a constitution through an iterated process of legislation and judicial review, which is interpreted as a form of dialogue. The model has been applied in the Latin American context as well, where scholars have highlighted the judicial construction of "dialogic" remedies, often in the context of individual constitutional complaints, whereby courts declare the violation of a constitutional norm, demand a policy response, but leave open the means through which this can be accomplished (e.g., Rodríguez-Garavito, 2010; Brinks and Forbath, 2010). Dialogic remedies, by their nature, are often vague. Our perspective suggests that there are sound reasons for vagueness even if one adopts one of the many rationales under which judicial review can be conducted quite consistently with a commitment to democracy. Yet at the same time our study reminds us that efforts to engage in dialogue may risk foot-dragging and ultimately a failure to fulfill legal obligations.

We also highlight that the politics of human rights protection in the Americas is neither wholly international nor domestic, and in this way connects to broad themes in this special issue. Gaining compliance with IACtHR orders can involve a complex process linking multiple executive agencies, legislative bodies and courts. Reflecting on the complexity of a political system that is judicialized

at the international and domestic levels, the authors of the framing paper suggest that a general model of the process will likely be "extremely complicated" and potentially "elusive." Instead they suggest a framework for inquiry in which scholars clarify "phases" of judicialized politics, including (1) the process by which political grievances are translated into valid legal claims, (2) the process of adjudication, (3) the process of implementation or compliance, and (4) the process by which actors outside the judiciary react to judicialized politics. Of course, as the authors also suggest, "while venues, actors and politics at each phase differ, actors may attempt to build connections across the phases to achieve their goals." In the context of international human rights norms, these connections are often highlighted through the lens of civil society, where norm entrepreneurs, NGOs, members of the legal profession and academics guide private parties in the identification of legally cognizable claims, fund legal campaigns, and generally help shape the processes by which states are compelled to respond to violations of human rights norms (e.g. Dancy and Michel, 2016; Lopez, Smith and Pagnucco, 1998; Sikkink, 1993).

Our paper focuses on the ways in which international judges manage connections between two phases of the judicialization process. One of the themes of the special issue is that connections between phases are common. Understanding the way that Larsson and Naurin claim that the Court of Justice of the European Union navigates European politics requires connections between the adjudicatory, implementation and reaction phases. Similarly, the logic of Lupu, Verdier and Versteeg's argument of how national courts make use of interpretive canons clearly recognizes connections between these phases. We argue that the successful achievement of judicial goals requires that judges build connections across the adjudicatory and compliance phases (see also Gauri and Brinks, 2008). We focus on how judges make these connections via language. In this way, our study reflects a concern of Pelc's contribution to the special issue, where the author considers how the World Trade Organization panels, as well as the Appellate Body, make use of affect to influence persuasion in legal argument.

We divide the remainder of the paper as follows. In the next section, we situate our study in the broader literature concerned with compliance with international human rights norms and the implementation of international court orders. We then summarize the model that guides our analysis and describe the background for the study, highlighting why the Court represents a useful

case for evaluating implications of the model. We then describe our design, data and results. We conclude by returning to the implications of this study.

Compliance, Implementation and Effectiveness

The primary outcome of interest in our study is the compliance of a state with a remedial order from the Court following a finding that the state has violated the American Convention on Human Rights. These orders identify particular tasks that states must carry out to provide reparations for violations of the Convention. To clarify the normative importance of our study it is useful to briefly define a few terms, especially since "compliance" is used by scholars in somewhat different ways. Generally, compliance is understood as "a state of conformity or identity between an actor's behavior and [a] specified rule" (Raustiala, 2000). In international law compliance is often understood as the relationship between a formal rule as detailed in an international agreement and the behavior of a state that is party to the agreement. A state that fails to offer individuals within its borders the right to "simple and prompt recourse ... to a competent court or tribunal for protection against acts that violate his fundamental rights" has failed to comply with Article 25 of the American Convention on Human Rights. Similarly, in comparative law, compliance often indicates a connection between a constitution and a state's behavior with respect to particular provisions of that constitution (e.g. Law and Versteeg, 2013; Chilton and Versteeg, 2016). Of course, compliance may also refer to the connection between a state's behavior and a judicial order. In this sense, compliance is understood as "full execution of the action (or complete avoidance of the action) called for (or prohibited) in one or more court rulings" (Kapiszewski and Taylor, 2013). Kapiszewski and Taylor's definition reflects our understanding of compliance.

Scholars draw a distinction between compliance and two related concepts: implementation and effectiveness. Implementation refers to the often complicated process by which legal commitments, as say articulated in treaties or constitutions, are put into practice through legislative and administrative reform. Effectiveness, in contrast, is "the degree to which a given rule induces changes in behavior that further the goals of the rule; the degree to which a rule improves the state of the underlying problem; or the degree to which a rule actives its inherent policy objectives" (Raustiala, 2000).

In the context of international human rights law, these conceptual distinctions are not always clear. To comply with orders, states often confront significant challenges of legislative and administrative reform, as when compliance requires the passage of a new law or changes to regulations or bureaucratic procedures. Thus compliance involves implementation. Further, the jurisdiction of the Court can only be triggered after domestic remedies have been exhausted and the Inter-American Commission on Human Rights has completed its own investigation. The cases that the Court hears commonly involve egregious violations of both international and the state party's own domestic human rights standards, following years of advocacy and the direct intervention of the Commission. The behaviors that the Court seeks to change are commonly behaviors that states have proven unwilling to change in the past despite a robust process of deliberation and international intervention. Thus, understanding why states comply with some orders and not others not only involves an investigation of compliance but of effectiveness as well.

Models of State Compliance

Why do states comply with international commitments? A first conjecture is that compliance is simply a measure of a preference for the kinds of commitments that states adopt when they ratify an agreement. Scholars of international politics recognize the possibility that states select into international regimes when they intend to comply with new commitments, or when international agreements require behaviors that states already endorse (Downs, Rocke and Barsoom, 1996; Von Stein, 2005). On this account, compliance is simply a reflection of pre-existing plans to behave in ways described by international agreements. Although the dynamic described by the argument is persuasive and has helped structure the interpretation of compliance outcomes in a number of international regimes, it sits somewhat less easily with the facts in the context of human rights regimes, especially the American regime. There is considerable variation in compliance in the context of human rights, suggesting that states do not only take on international human rights commitments when they intend to comply (e.g. Simmons, 2009).

Setting aside the issue of non-random selection into international treaties, scholars have highlighted four forces that influence a state's compliance with international commitments. As we will see, these forces have also been identified as important sources of state compliance with court orders. States are believed to comply with international norms at least in part because of interna-

tional pressure, whether it comes from reputational concerns (Simmons, 2000), material incentives (Hafner-Burton, 2005; Hashimoto, Forthcoming) or the mobilization of an international network of activists (Neumayer, 2005). An alternative view, associated mostly commonly with Chayes and Chayes (1993), suggests that states often fail to comply with international norms simply because of resource constraints that undermine the effective management of bureaucratic agencies necessary to bring a state into compliance. On this account, states might want to comply but lack the resources or human capital to do so. A third view focuses on the potential for a variety of domestic institutions to prevent leaders from violating constraints on their power or to incentivize positive changes. These institutions include independent judiciaries (Lupu, 2013), legislatures (Lupu, 2015), and empowered prosecutors who are willing to pursue justice claims (Dancy and Michel, 2016). Finally, scholars have raised the possibility that international courts themselves influence compliance with international norms. They do so by interpreting the meaning of treaty provisions in ways such that states are only asked to change behavior when the change is desired by many other, and typically powerful, state parties to the agreement (Carrubba and Gabel, 2014). This account identifies a distinct source of sample selection bias in the study of compliance, one that runs through the judicial interpretation of treaty terms.

These four factors are highlighted by scholars who have focused more specifically on state compliance with court orders. Importantly, scholars often link the forces, as say when domestic institutions influence the way that international pressure can be brought to bear on a state. Cavallaro and Brewer (2008, pp. 792-793) suggest that non-governmental organizations can create pressure for compliance in a case by increasing media coverage. By enhancing publicity surrounding a case, NGOs may be able to raise the states' costs of non-compliance arising from domestic politics, the international community or both. A second line of argument considers the international implications of reactions to international court decisions. Hillebrecht (2010) suggests that states use their responses to the Court to signal commitments to human rights to a variety of external communities. According to the argument, leaders should be increasingly likely to comply as they become constrained domestically by other internal institutions.

A third line of argument focuses on bureaucratic capacity. Both Hawkins and Jacoby (2010) and Voeten (2014) note that the implementation of many policy changes depends on a well-run bureaucratic process as well as high human capital in the state sector. We might then expect that

states with high quality bureaucracies should be less likely to resist orders, simply because they are better equipped to implement them. Finally, (Carrubba, 2005) explicitly models the choice of a state to implement an international court order (also see Larsson and Naurin, 2016; Carrubba, Gabel and Hankla, 2008). Precisely because international courts wish to avoid non-compliance and can anticipate non-compliance, the primary implication of their work is that states are infrequently confronted with compliance requests that are difficult to implement.

The Role of Judicial Writing

Each of the existing models of compliance suggests a convincing mechanism through which an international court order will be successfully implemented by a state. Yet nearly all of them ignore the role that judges themselves might play in this process; and, even scholars like Carruba, Larsson, Naurin, Gabel and Hankla, who consider the possibility that judges might influence compliance through their decision-making processes, largely obscure a critical aspect of the work that judges actually carry out. Judges resolve cases. They explain their resolutions in natural language meant to articulate rationales. And when appropriate, judges prescribe remedies. A single case can bring multiple legal and political challenges that will require distinct solutions, all communicated in written language. To be clear, we do not argue that existing models are incorrect or unhelpful. Far from it. Each of the existing approaches illuminates an important element of the process by which court decisions are obeyed and by which international norms are ultimately implemented. That said, we also believe that the outcomes to legal conflicts are influenced in important ways by choices judges make in their opinions (Basch et al., 2010). These choices are articulated in text and the control over text offers judges a measure of control over the compliance process that is not easily articulated within any of the existing models.

The model of opinion vagueness that guides us is developed elsewhere (Staton and Vanberg, 2008) and we refer the reader to the original paper for details of the argument. It is nevertheless useful to summarize its motivation and the empirical implications we can evaluate here. The model asks how clearly a court ought to express its demands for policy change when it confronts a government whose cooperation is necessary for policy implementation, and where the government prefers the *status quo* to any other policy. The model assumes that the government has an informational advantage over the court – it knows how to translate a policy demand into a particular policy

outcome. The court is uncertain about how that works. Importantly, uncertainty varies so that in some contexts, the court knows approximately as much as the government, but in other contexts the court is considerably uncertain. Thus, the government and the court have conflicting interests over policy outcomes and differ with respect to knowledge of how to turn a particular order into an outcome.

Reflecting many models of judicial politics (Clark, 2010; Cavallaro and Brewer, 2008), the model assumes a public enforcement mechanism for judicial decisions. Specifically, governments pay costs for not implementing the policy the court orders. Yet, the marginal cost of defiance is influenced by the clarity of the court's order. By increasing clarity, the court makes defiance decreasingly attractive, because it is easier for publics to observe non-compliance when the remedy is clear. On the other hand, by limiting discretion, clear orders undermine the ability of a court to leverage informational advantages that state officials might have about how to translate a particular demand into a policy outcome, a problem that becomes increasingly salient when judges lack certainty over how precisely to achieve their goals. And finally, the court itself pays a cost for being openly defied, so that even though clarity increases pressure to comply, any non-compliance will be increasingly costly to the court itself, since it will be more clearly observed. This dynamic provides another incentive to be vague.

An equilibrium in this model can be described by a pair of outcomes (y_1, y_2), where y_1 reflects the clarity of the order that the court makes demanding a policy change and y_2 reflects the degree to which the policy implemented by the government diverges from the policy requested by the court, i.e., y_2 measures compliance. Government reactions (y_2) to the court's demand depend on how clearly the court issues the order (y_1). Clarity pulls the government's response more closely to the court's ideal policy outcome. For a given level of clarity, non-compliance increases in the degree to which the government and the court disagree about the policy outcome and decreases in the publicly imposed costs for defiance that the government perceives.

Judicial choices over order clarity (y_1) anticipate these reactions. As long as the court does not assign an extremely high weight to its own consequences of having non-compliance become public knowledge, the equilibrium level of clarity turns importantly on how uncertain the court is about the translation between law and policy. As the court becomes increasingly uncertain about how exactly to translate an order into an outcome, order clarity decreases. Likewise, the model

suggests that courts will be increasingly clear as government preferences diverge from their own, using clarity to place pressure for compliance potentially intransigent officials.[2]

The equilibrium levels of clarity and compliance are consistent with the following recursive system of equations:

$$y_1 = \quad f_1(x_1, x_2) \tag{1}$$

$$y_2 = \quad f_2(y_1, x_2, x_3) \tag{2}$$

where x_1 denotes judicial uncertainty about how to translate an order into a policy outcome, x_2 denotes the divergence in preferences between the court and the government, and x_3 denotes the costs the government anticipates will be paid for defying an order.[3] The critical point, for empirical reasons at least, is that the system implied by the model is recursive. Observed levels of clarity directly influence the government's reaction to the court's decision. Expectations about likely compliance influence the court's optimal choice of clarity, but these expectations are defined in equilibrium by exogenous parameters.

Our empirical goal will be to generate measures of these concepts and estimate the system implied by the model. The empirical implications we consider are as follows.

> *Uncertainty: Uncertainty about the translation between an order and an outcome is negatively associated with clear orders; and, through this association with clarity, uncertainty is negatively associated with compliance.*

[2]Increasing the public costs of defiance for the government has competing effects on opinion clarity. The reason is that an increase in awareness provides a judge with greater leverage to address the means-ends problem, and this leverage can be translated into either greater order clarity or vagueness. Thus, increases in awareness and order clarity should both be associated with better compliance outcomes; however, it is not the case that greater awareness necessarily increases order clarity.

[3]Interestingly, public support has two competing effects on opinion vagueness, which renders the empirical implication indeterminate in the data to which we will have access. In a section evaluating robustness, we consider the implications of conditioning on a measure of these costs. None of the substantive results we highlight is influenced materially by this choice.

Clarity: Clear orders are positively associated with compliance.

Preference Divergence: A divergence in preferences between a court and a government are negatively associated with compliance.

Costly defiance: High costs associated with defying a court are positively associated with compliance.

Empirical Analysis

The Inter-American system represents a particularly appropriate setting for testing elements of the model. The jurisdiction of the Court creates a wide array of policy challenges for judges to consider. Past behavior of court members suggests that there will be useful variation in clarity and compliance to analyze and the Court's own practices considerably alleviate practical challenges of measurement.

The Inter-American Court of Human Rights

The Court's jurisdiction has been documented carefully elsewhere (Pasqualucci, 2012; Basch et al., 2010; Hillebrecht, 2010; Cavallaro and Brewer, 2008). Still a brief review is useful, because the court's jurisdiction and the cases that comprise its docket clarify the appropriateness of testing the model's implications in this case. The Court enjoys both adjudicatory and advisory jurisdiction pursuant to Articles 61-64 of the American Convention on Human Rights.[4] The court's contentious jurisdiction may be triggered by a member state of the OAS or the Inter-American Commission; however, no member state has submitted a case to the Court (Tan 2008, 248). Consequently, the Commission is the principal institution through which cases arrive.

The Commission receives complaints from individuals and non-governmental organizations alleging state violations of human rights obligations under the American Convention or the American Declaration of the Rights and Duties of Man, the latter of which though not legally binding has been interpreted by the Court as applicable to member states as the "authoritative interpretation of the human rights commitments contained in the OAS Charter" (Cavallaro and Brewer, 2008, pp. 778-779). Following its own investigation, and in the event that the Commission is satisfied

[4]The Court's jurisdiction is developed further by the Statute of the Inter-American Court of Human Rights, adopted by the General Assembly of the OAS in October 1979.

that a member state has violated its obligations and failed to take appropriate measures to remedy the situation, the Commission will either draft a final report to be published in the OAS's Annual Report, or assuming that the state has accepted the Court's contentious jurisdiction, submit the case to the Court (Reisman, 1995; Pasqualucci, 2003). Having ratified the American Convention and made a separate declaration of consent to the Secretary General of the OAS, decisions of the Court are binding on member states (Guzman and Landsidle, 2008). The implication of this jurisdiction is that the docket of the Court is biased toward particularly problematic cases, where there has already been considerable resistance from the defendant state with respect to recommended policy changes.

An important challenge in studying reactions to judicial decisions involves a variety of processes of sample selection that can raise simple questions about inference. Clearly, the cases that the Court hears are not a representative sample of all legal conflicts in a state or even of human rights claims in a global sense. They are not even a representative sample of human rights claims raised in the Americas. As we have discussed, the Court hears only the most extreme cases. Yet this source of sample selection is helpful if we wish to consider compliance with (or the implementation of) court orders in policy contexts where we can be confident that states would prefer not to change their behavior. Another concern about sample selection flows from the kind of model developed by Carrubba (2005). Might it be that the Court only issues orders that it knows will be obeyed? In the model we test, this kind of logic would result in orders that are maximally vague and there were be no variation to explain. We know from past work on the Court that its compliance record is very far from perfect, and we will show that there is considerable variation in the clarity with which the court drafts its remedies. There is also no reason to believe that judges of the Court during the time of our study would have been surprised to learn that some states will resist implementing their orders. These facts suggest that the Court does not single-mindedly pursue compliance above all goals, making the court a useful case for evaluating the implications of the model.

A third useful feature of the Court concerns measurement. A core challenge in evaluating claims regarding compliance with judicial decisions lies in uncovering information about what happens in the aftermath of a litigation. Legal systems typically envision a "fire alarm" mechanism to deal with non-compliance. The parties are incentivized to bring non-compliant behavior to the attention of judicial system in one form or another. In many ways this would seem like an efficient solution

to the problem of monitoring compliance. Yet there are multiple reasons (resources, intimidation, uncertainty about what a judicial order requires exactly, etc.) why this mechanism might fail to identify all or most instances of non-compliance. Consequently, we typically cannot be sure that non-compliant outcomes that make it into the judicial process constitute a representative sample of compliance outcomes generally.

The Court represents a particularly useful empirical context because it has supported a detailed monitoring system for its decisions in contentious cases since 1996. Challenged in 2003 by Panama in the state of Baena Ricardo *et al* case, the court found an inherent power to monitor compliance deriving from its status as an international judicial body, and specific legal authority derived from Articles 33, 62(1), 62(3) and 65 of the American Convention (Hawkins and Jacoby, 2010).[5] During the period studied, the Court's compliance reports were developed by staff specifically charged with this task, primarily via direct communication with the defendant states, the Commission, NGO's and the victims' representatives regarding the status of certain reparations. The Court also collected information from "expert declarations and reports" and on occasion from hearings with defendant states.[6] Monitoring continues over time, allowing the court to evaluate whether states continue to ignore commands as well as giving the court a chance to clarify previous orders.[7]

Data

As Hillebrecht (2009) highlights, because the decisions of international courts typically involve multiple orders, within many cases, across states that appear frequently, measuring compliance with orders of international courts involves important choices about the unit of analysis, and in

[5]See the competence judgment of the court: *Baena Ricardo et al v. Panama*, Inter-Am Ct. (November 28, 2003). Procedures for monitoring compliance with judgments are specified by Article 69 of the court's Rule of Procedures.

[6]See IACtHR Rules of Procedure, Article 69.

[7]Our data reflects the full set of compliance reports which the Court makes available. As we discuss, our first compliance measure reports on the status of an order at the point of the last compliance report. We also measure whether there was ever a problem with compliance discussed in any of the reports.

light of that choice, potentially even more challenging questions about aggregation. For example, suppose that we wish to focus on states as units. Characterizing the compliance behavior of a state will require aggregating up within a case from multiple orders, and then aggregating again within a state from potentially multiple cases. Our study is primarily concerned with the connection between the outcome a judge seeks and the language through which she encourages the outcome. For this reason, the unit of analysis for our study is the Court reparation. We coded final decisions to contentious cases resolved by the court between 2005 and 2009, which also has at least one compliance report. This resulted in 292 remedies across 45 cases. Specifically information about the cases and their reparations was collected using the original, Spanish version of the case. The information was collected by a native Spanish speaking research assistant from the Emory School of Law, trained by the authors.

Measures of Compliance and Clarity

As has been described elsewhere (e.g. Hawkins and Jacoby, 2010; Hillebrecht, 2010), for every case in which the Court issues a decision on the merits and specifies reparations, it indicates in its compliance report whether the state has complied or not with each remedy. Thus, our empirical measure of y_2 is a binary variable, which takes a value of 1 if the court is satisfied that the remedy has been properly implemented and 0 otherwise. Although valuable, this indicator runs the risk of under-estimating the court's compliance problems. Critically, states may comply with orders eventually, but engage in significant delay. If they eventually implement the order, it is certainly reasonable to code the state as having complied, but if compliance followed a protracted period of delay, such a coding will miss what we believe is substantively meaningful defiance. To address this issue, we also consider another binary variable y_2^r, which indicates whether there is evidence of state delay or resistance with respect to a remedy. Specifically, y_2^r takes a value of 1 if there is such evidence and 0 otherwise.

To capture the clarity of each reparation, we asked for a subjective evaluation of each reparation, according to the following rules:

1. Clear Order: What is required of the state is completely unambiguous. For example, consider the first remedy in the Montero Aranguren case, where the Court is requiring monetary

damages to be shared among the victims' families. There can be no question here about what is required.

> *Taking into account the different aspects of the non pecuniary damage caused, the Court awards, on equitable grounds, the following compensations: a) For each of the 37 executed victims, the Court awards a compensation in the amount of US$ 75,000.00 (seventy five thousand United States Dollars).*

2. Somewhat Clear: What is required of the state is fairly clear, but how it should be done is not completely obvious. For example, consider the following remedy from the eighth remedy in Montero. Clearly, a public apology is required and a time frame is given; however, precisely how this is to be done and where is not articulated:

> *The Court orders the State to publicly acknowledge its international liability and ask public forgiveness to the victims next of kin declared in the instant Judgment. Said acknowledgment must be made in the presence of the victims next of kin and the highest-ranked authorities of the State within the term of six months as from the date of service of the instant Judgment.*

3. Vague - The remedy is extremely unclear, leaving considerable discretion to the state. For example, we consider the seventh remedy in Montero to fall into this category. Although it is clear that a program should be developed, many details are left unspecified:

> *[T]he Court deems it appropriate that the State frame and implement a training program on human rights and international standards applied to inmates addressed to police agents and penitentiary officials.*

Importantly, the research assistant who completed the coding of our sample of cases was generally familiar with project goals. This raises the possibility that her coding of order clarity might have been influenced by her reading of the compliance reports and her sense of our empirical expectations. To address this concern we asked a second coder, who was not familiar with project goals and was

not exposed to the compliance data, to replicate our measures of reparation clarity.[8] Below we present results using measures produced by both coders. We have found in repeated evaluations that although there may be reasonable conceptual reasons to justify a trichotomous coding of clarity, as an empirical matter, there is no difference between the first two categories. Thus, in this paper we report the findings associated with a binary variable, which is coded as 1 if the remedy was judged to be clear or somewhat clear and 0 otherwise. This reflects our empirical measure of y_1.

Table 1 displays the bivariate distribution of the response variables, for both the compliance and resistance indicators. The table reflects previous findings suggesting that compliance on the Court is, generally speaking, a problem. We identify resistance with nearly 81% of the remedies. Indeed, the court itself only codes 31% of its remedies as having been implemented to its satisfaction. Of greater theoretical interest, though, is the relationship between remedy clarity and compliance. For both indicators, the proportion of cases in which there is evidence of some sort of resistance is lower for clear remedies. To consider this relationship further, we require a model of the process that is both consistent with theoretical expectations and addresses obvious structural features of the data.

Measures of Uncertainty, Government Preferences, and the Government's Costs of Compliance

We need measures of the theoretical constructs x_i. To capture relative differences in uncertainty the Court might have over the link between policy means and policy ends for particular remedies, we first adopt a relatively coarse approach, which depends on fundamental differences in the policy-making contexts associated with each remedy. We discuss an additional approach below. The intuition behind our measure is that not every kind of reparation involves the same kind of policy-making challenge. For example, pecuniary damages present almost no challenge at all. Any judge in the world, with little to no help, could transfer money between two bank accounts without error. All she would need is access. In contrast, providing clean drinking water or acceptable healthcare solutions present important policy-making challenges, fraught with uncertainties about how precisely to implement the policy. Our measure of uncertainty, x_1 is binary and answers

[8]There is evidence of reasonable inter-rater reliability. Specifically, the kappa statistic for the two coders is 0.69.

	Compliance (Row %)	Non-Compliance (Row %)
Clear Remedy	81 (40)	123 (60)
Vague Remedy	12 (14)	73 (86)
	No Resistance	Resistance
Clear Remedy	52 (26)	151 (74)
Vague Remedy	3 (4)	82 (96)

Table 1: *Relationship between Clarity and Compliance (Resistance). Displays the distribution of Comply and Resist for both values of Clear. Row percentages are shown in parentheses.*

a simple question: Could a reasonable person carry out the task without error in the absence of policy-making expertise? Thus, x_1 is coded 1 if the answer to this question was yes and 0 otherwise.[9]

Measuring the policy differences between the Court and the sitting government at the level of the remedy represents a significant challenge. Our approach assumes that the theoretically relevant policy differences between the court and the defendant government (i.e. how much the court and the government differ over ideal policy outcomes) are likely correlated with a state's underlying human rights record. Better records likely reflect closer connections between the court and the government. States with generally good human rights records will likely require less costly changes to come into line with international obligations. These states may find it easier to make changes, as well. Although we cannot observe this record directly, there are a number of observable indicators

[9]The following remedy types were considered to present little to no problems with implementation: 1) Pecuniary and non-percuniary money damages, 2) Publication of Decision, 3) Acknowledge responsibility for violation, 4) Release information, 5) Strike criminal conviction from record, 6) Sponsor an academic scholarship. The following remedy types received a 1: 1) Undertake criminal proceeding/Investigate, 2) Change or modify a law, 3) Implement programs, 4) Equitable relief (e.g. provide psychological treatment)

of it. We rely on Fariss's (2014) estimates of latent respect for human rights. This variable is measured for each country and year during the period of our study. Specifically, for each state that is challenged in the cases in our dataset, we assign the value of the mean of the posterior distribution of latent respect for human rights for the year that the Court resolved the case.

Finally, we require an indicator of the publicly-induced costs of non-compliance that a government might pay for defying the Court. Ideally, we might appeal to a measure of the Court's legitimacy in each state. Although costs probably vary by case, and perhaps by remedy, this would be a reasonable strategy. Unfortunately, we are aware of no such indicator. In its place, we appeal to the Latinobarometer's judicial confidence measure, which provides a survey-based measure of citizen's confidence in their domestic judiciary.[10] Thus, x_3 is the average, mean centered, judicial confidence for a state in the year of the Court's final decision.

Any indicator of this concept will have to be a proxy, in so far as we are attempting to measure the perceptions of state leaders regarding the likely consequences of failing to implement an Court order. The first assumption we make in appealing to this indicator is that judicial confidence is a reasonable proxy for state official's perceived costs of defying a domestic court order. Helmke (2010) finds that a lack of confidence in the judiciary, as measured by the Latinobarometer items we use, is associated with political attacks on judicial institutions, e.g., threats against judges' positions, politically motivated packing or removals, etc, suggesting that at a minimum, domestic courts face significant political threats when they do not enjoy trust. The second assumption we make is that the perceived costs of defying an international court order are correlated positively with the perceived costs of defying a domestic court order. Trust in judicial institutions is commonly thought to constitute an element of the states whose citizens also value the rule of law (e.g. Prillaman, 2000). In states with strong rule of law values, judicial orders are expected to be obeyed, and it is conceivable that these expectations might be transferred to international court orders. If confidence is associated with attacking domestic judicial institutions, it also seems plausible that political leaders might feel less than compelled to comply with orders from international courts.

We offer two caveats. First, we are not assuming that the costs of defying an international court order are necessarily equal to those associated with defying a domestic court order. It is

[10]Latinobarómetro [2006,2007,2008,2009].

certainly possible that some forms of defiance might be preferred by domestic publics, especially when considering an international institution. Second, given the Court's docket, which again only includes cases that states have been unwilling to resolve for many years, it is possible that all of our cases involve very low costs of defiance. If true, then we will not observe a strong association between our confidence measure and the compliance outcomes we seek to explain. To pursue this possibility further, we consider a few additional measures of the perceived costs of defiance in what follows.

Analysis

The data we analyze present a three-level hierarchical structure, with remedies i nested within cases j which are nested within states k. We begin with a model that includes covariates suggested only by the argument. We thus consider the following system:

$$\begin{aligned} Pr(y_{1i} = 1) &= logit^{-1}(\alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2k[ij]} + \psi^{case}_{j[i]} + \psi^{state}_{k[ij]}) \\ Pr(y_{2i} = 1) &= logit^{-1}(\beta_0 + \beta_1 y_{1i} + \beta_2 x_{2k[ij]} + \beta_3 x_{3k[ij]} + \zeta^{case}_{j[i]} + \zeta^{state}_{k[ij]}) \end{aligned}$$

The ψ and ζ parameters reflect modeled random intercepts for cases and states. The model assumes that each intercept is drawn independently from a normal distribution with zero mean and a unique variance. This system is recursive, and assuming no correlation of errors across equations, we estimate it equation by equation.[11] The models are fit in Stata 14 using the gllamm package.

Results

Figure 1 summarizes the results of fitting the system of equations that we describe above. A complete table of results for these models is available in the appendix. The two columns show the results using two distinct coders' evaluations of opinion clarity. On the left we display the results derived from a measure in which the coder also captured information about the Court's cases. Although this coder was unfamiliar with the theoretical goals of the project, there was an opportunity to guess what might be of interest to us. The results summarized in the second column use a measure of opinion quality generated by another coder who never saw the Court information and was absolutely blinded to all goals of the study except for our intent to measure order clarity.

[11] We consider the sensitivity of our results to this assumption below.

The first row of the figure shows coefficients in the equation estimating the probability of a clear order. The second row shows coefficients in the equation estimating the probability of compliance. The third row shows the coefficients in the model of resistance to the court's orders.

The results in Figure 1 suggest that the Court was considerably less likely to issue clear orders in policy areas that would appear to present empirical challenges with respect to translating an order into a policy outcome; and, compliance would appear to be more likely (and resistance less likely) when the court is clear. These results are strongly significant. Indeed, the Bonferroni corrected p-values (for 8 comparisons) for the uncertainty result in the model of y_1 as well as the opinion clarity results in the models for y_2 and y_2^r are .008. .072 and $< .001$, respectively. The evidence connecting a state's human rights record as measured by the Fariss latent variable estimates and compliance is mixed. Although the coefficients are in the expected direction, results do not reach traditional levels of statistical significance. There is also no evidence that the human rights context of a state influences the clarity of the courts orders. Order clarity would appear to track well the type of policy challenge that the court confronts, but not the human rights record of the state that is the target of the order.

A critical question in this analysis is how does policy uncertainty influence compliance. According to the model, the effect is indirect, flowing through the clarity of orders that the court issues. In the context of our model, ignoring for the moment the random intercepts, the effect of our measure of policy uncertainty (a binary variable) is the difference between two quantities: $\mathbb{E}[y_2|x_1 = 1, x_2, x_3] - \mathbb{E}[y_2|x_1 = 0, x_2, x_3]$ (Greene, 1998), where $\mathbb{E}[y_2|x_1 = 1, x_2]$ is the following sum of probabilities:

$$Pr(y_1 = 1|x_1 = 1, x_j) \cdot Pr(y_2 = 1|y_1 = 1, x_1 = 1, x_j) + Pr(y_1 = 0|x_1 = 1, x_j) \cdot Pr(y_2 = 1|y_1 = 0, x_1 = 1, x_j)$$

for $j = 2, 3$. The first term in the sum reflects the probability of a clear order multiplied by the probability of compliance, given that the order is clear. The second term in the sum reflects the probability of an unclear order multipled by the probability of compliance, given that the order is vague. Both probabilities are for uncertain policy areas. Likewise, $\mathbb{E}[y_2|x_1 = 0, x_2, x_3]$ is given by the following sum.

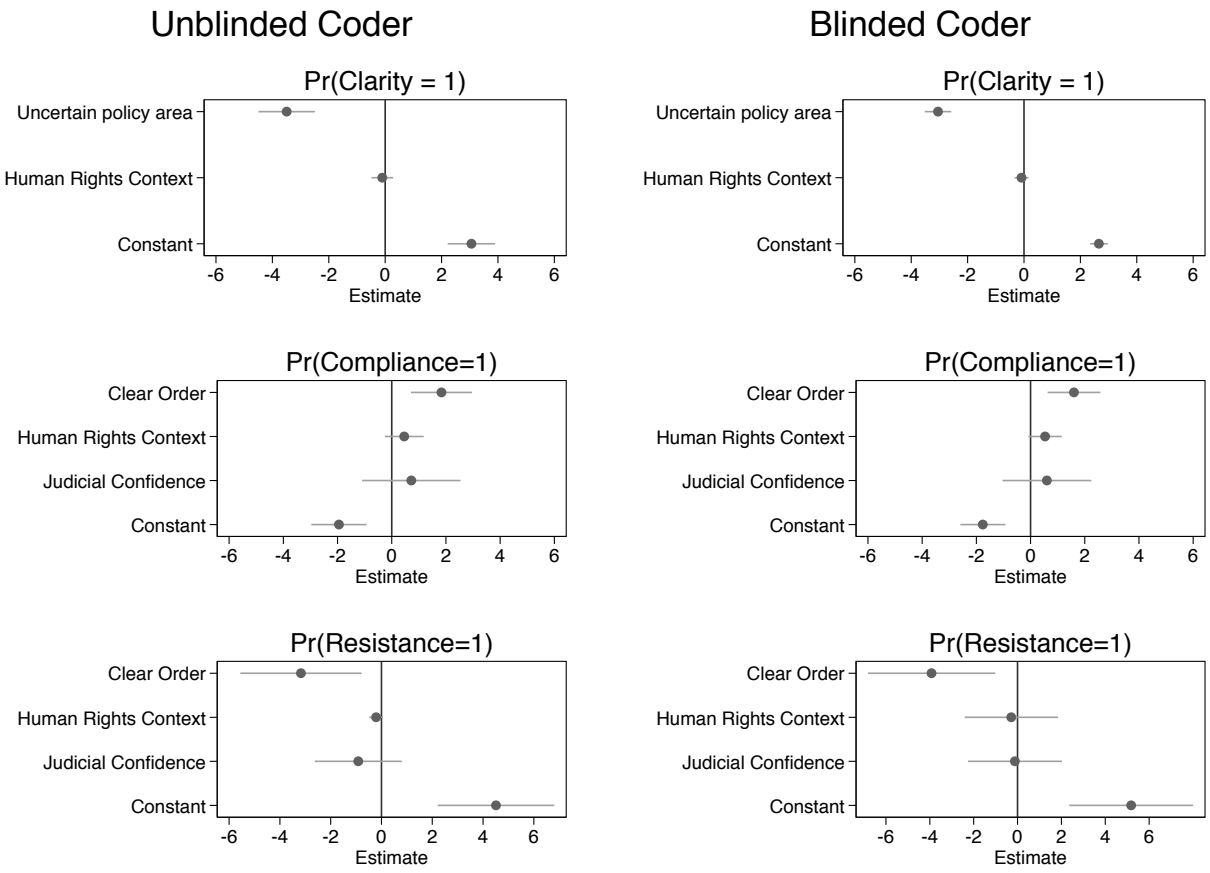

Figure 1: *Results of Clarity and Compliance Models for Unblinded and Blinded Coders. The figure summarizes coefficients from the system of equations with 95% confidence intervals. The left column shows results relying on a measure of opinion clarity constructed by a coder with access to the information about the Court's decisions. The right column shows results relying on a measure of opinion clarity constructed by a coder who was fully blinded to that information. The first row shows coefficients in the equation estimating the probability of a clear order. The second row shows coefficients in the equation estimating the probability of compliance. The third row shows the coefficients in the model of resistance to the court's orders.*

$$Pr(y_1 = 1|x_1 = 0, x_2) \cdot Pr(y_2 = 1|y_1 = 1, x_1 = 0, x_j) + Pr(y_1 = 0|x_1 = 0, x_2) \cdot Pr(y_2 = 1|y_1 = 0, x_1 = 0, x_j)$$

This sum differs from the first only in that the probabilities condition on setting the uncertainty variable to 0.

Given the random intercepts, calculating the predicted probabilities requires further assumptions as we need to establish counterfactuals for a particular case and state. To generate estimates of $\mathbb{E}[y_2]$, we conducted the following procedure. We fit the models described above. We then took 2,000 draws from the multivariate normal distribution of the model's parameters, using the estimated parameters for the vector of means and the estimated variances and covariances for the variance-covariance matrix. We predicted the random intercepts as well as their variances using the *gllapred* package, which provides empirical Bayes predictions of the posterior means and standard deviations of the random effects. We then identified the median intercept for the case in the compliance equation and used that case and its corresponding state to generate predictions. This turned out to be Palamara Iribarne v. Chile. Having identified this case and state, we take 2,000 samples from sampling distribution of the intercepts, using the estimated means and variances just derived. Adding these samples to the samples of the other modeling parameters, we are able to generate 2,000 predictions for each quantity we desire by evaluating the inverse logit function at particular values of the linear predictor, given the parameters and the relevant values of x_i. Predicted probabilities reflect the mean of these samples. The 95% confidence intervals reflect the values of the samples in the 2.5th and 97.5th percentiles. As the table summarizes, setting all covariates to their mean values, the predicted probability of compliance in this case when the policy challenge was deemed to be uncertain was 0.47. For a policy challenge deemed to have little uncertainty, the predicted probability of compliance was 0.65. The estimated total effect of an uncertain policy area is -0.18. The 95% confidence interval for the effect is (-0.44, -0.01).

Alternative Measures of Costs of Compliance

Prior research on compliance with Court orders suggest alternative approaches for measuring state perceptions of the costs of failing to comply with a court order. The first, developed in Cavallaro and Brewer (2008, pp. 792-793) and Simmons (2009) suggests that non-governmental organizations can create pressure for media coverage in a case and in that way increase pressures for compliance.

	Predicted Probability of Compliance	95 % C.I.
Uncertainty = 1	0.47	(0.18 ,0.80)
Uncertainty = 0	0.65	(0.28, 0.91)
$\triangle$	-0.18	(-0.33, -0.05)

Table 2: Effect on compliance of an effort to make policy change in a policy area in which the court is likely to be relatively uncertain.

A second possibility considers constraints on the executive, drawing on the work of Hillebrecht (2010). If only leaders who are constrained by other domestic institutions at home can credibly signal commitments to human rights to a variety of external communities, it is plausible that these leaders will confront lower costs of compliance than leaders who are not constrained. Third, we know that bureaucratic capacity has been found to be an important determinant of compliance outcomes, and it natural to imagine that well-run bureaucracies render compliance with court orders easier (Hawkins and Jacoby, 2010; Voeten, 2014).

To evaluate these alternatives, we first counted the number of NGOs participating in each case on behalf of the complainant. The logic of this measure is that large numbers of NGOs participating in the litigation reflect considerable organizational support for the victims. That support can be leveraged to pressure for compliance by publicizing results. We considered two possible approaches to measuring the ability of executives to make unilateral policy choices, free from domestic pressures. We first appeal to the Polity IV measure of executive constraints (Marshall and Jaggers, 2005). We then appeal to a measure of the independence of the state's judiciary (Linzer and Staton, 2015). Judicial independence has been shown to be positively associated with good human rights practices (e.g. Keith, 2002; Lupu, 2013), and Hillebrecht (2010) suggests that judicial independence is the kind of constraint that should be particularly useful for signaling a human rights commitment. Finally, to assess the role of a well-run bureaucracy, we appeal to the PRS index of bureaucratic quality. The PRS gives a three category measure, indicating that a state's bureaucracy has low, moderate or high quality. We construct two dummy variables from this measure for moderate and high quality bureaucracies.

Figure 2 summarizes the results of fitting GLS models with logit link functions for the compliance equation. We fit three separate models for a number of reasons. The concepts of judicial independence and bureaucratic quality are very closely related to the concept of "executive con-

straints," in the sense that an independent judiciary is often a constraint on an executive. A high quality bureaucracy can be a constraint on executives in some cases, as could be a low quality bureaucracy. On the other hand, a high quality bureaucracy might also allow executives to implement policy choices more easily and thus might reflect a lack of constraint. Research on judicial independence also suggests that judges are more likely to behave independently when government is fragmented, which is a typical source of executive constraint (e.g. Ríos-Figueroa, 2007). Finally, the measure of judicial independence we use is derived from a latent variable model in which the Polity IV score is one of the manifest variables used to estimate it.

As is clear in the figure, we do not estimate significant effects for NGO support or either measure of constrained executives. The coefficients for the NGO variable and the Polity IV executive constraints variable are precisely estimated to be zero. The coefficient on the judicial independence measure is in the expected direction but the estimate is very imprecisely estimated. Bureaucratic quality does seem to influence compliance as expected. Importantly the result on the order clarity variable is robust to these alternative measures.

Does uncertainty influence compliance directly?

The theoretical model assumes that the government is perfectly informed about the link between a policy demand and a policy outcome. Empirically, it is nevertheless possible, frankly likely, that state officials are also uncertain about how to turn a legal obligation into a policy outcomes. To model this possibility we add the uncertainty measure to the second equation, so that there is now the possibility of both an indirect and a direct effect of uncertain policy areas. Figure 3 compares the original results (top panel) to the results of a model in which we condition the compliance probability on our measure of uncertainty in the compliance process. As is clear from the figure, it would seem that compliance is lower in uncertain policy areas. Further, the clarity coefficient is reduced in this specification. Yet the effect of uncertainty is not well captured by this plot in light of the fact that we know that vague orders follow from policy uncertainty.

Table 3 shows a comparison of the effects of uncertainty when we allow for both direct and indirect effects. Although the theoretical model does not anticipate these effects, it seems likely that the features of the policy challenges that encourage vague reparations also directly influence compliance. Simply, it is easier to transfer money than it is to introduce new national programs

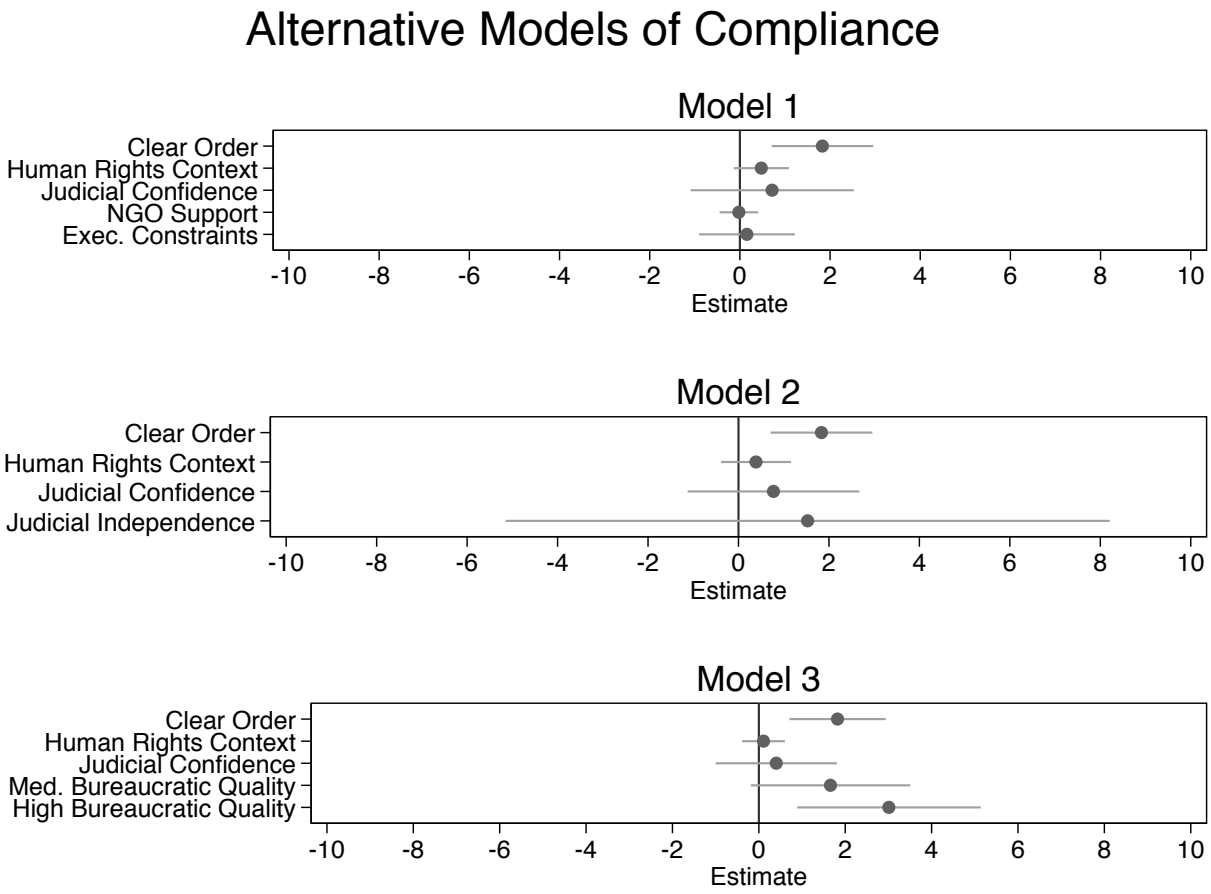

Figure 2: *Alternative Models of Compliance. The figure summarizes coefficients from the compliance equation with 95% confidence intervals. Model 1 includes covariates for the number of NGOs in support of the victims' case as well as a measure of executive constraints. Model 2 includes a measure of domestic judicial independence. Model 3 includes a measure of bureaucratic quality of the defendant state.*

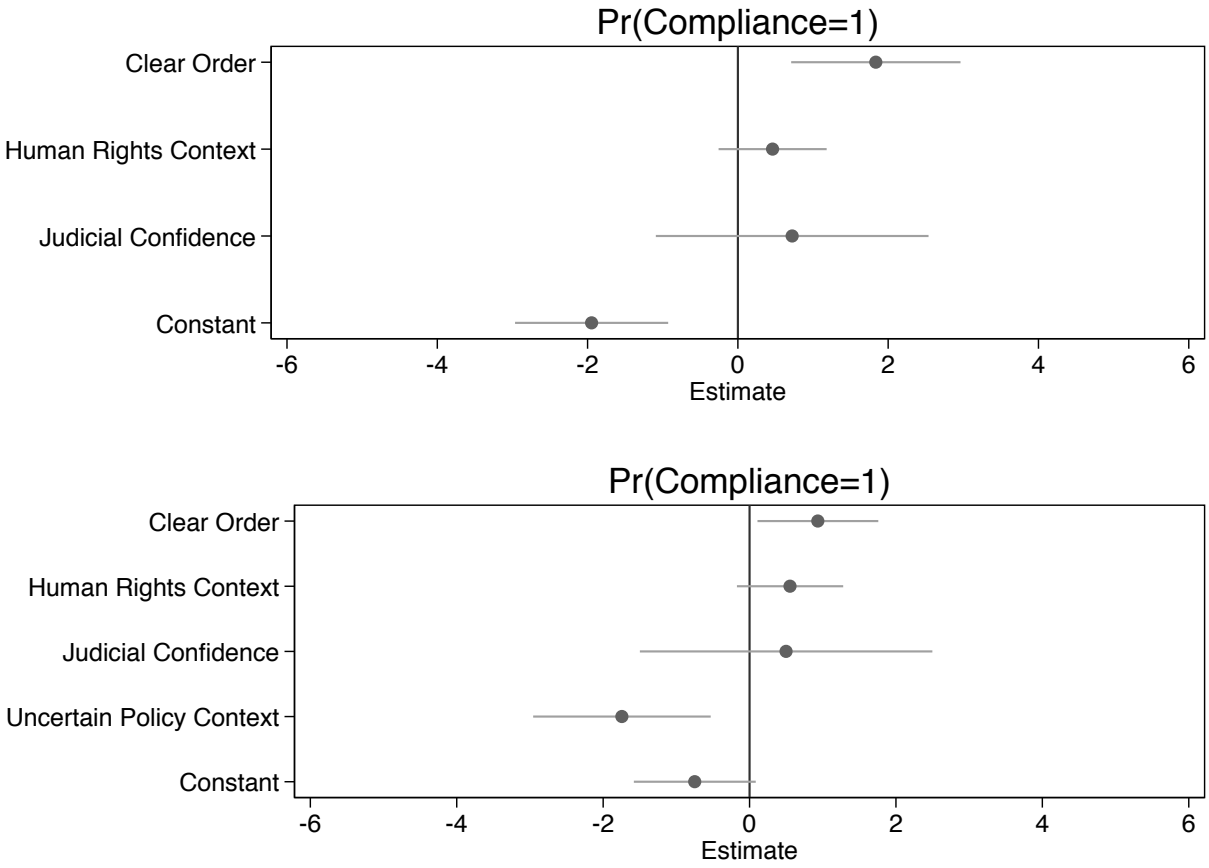

Figure 3: *The top panel contains the original results. The bottom panel includes the uncertainty measure.*

of police training. The results here suggest that once we allow for both types of effects, uncertain policy areas are associated with even larger changes in the probability of compliance. Specifically, the predicted probability of compliance climbs from 0.37 for policy challenges that we believe involve higher levels of uncertainty to 0.78 for policy challenges that should be relatively simple to implement.

		Predicted Probability of Compliance	95 % C.I.
Only indirect effects	Uncertainty = 1	0.47	(0.18 ,0.80)
	Uncertainty = 0	0.65	(0.28, 0.91)
	$\triangle$	-0.18	(-0.33, -0.05)
		Predicted Probability of Compliance	95 % C.I.
Direct and Indirect effects	Uncertainty = 1	0.37	(0.06, 0.81)
	Uncertainty = 0	0.78	(0.38 , 0.97)
	$\triangle$	-0.41	(-0.67, -0.11)

Table 3: *Comparing the effects of uncertain policy areas on compliance when allowing for indirect and direct effects. The top half of the table includes the original results, where uncertainty is constrained to have only indirect effects on compliance. The bottom half of the table shows the results when uncertainty is allowed to have both direct and indirect effects.*

Unpacking the Uncertainty Measure

A final question we ask in light of these results is whether we are obscuring important variation by creating a dichotomous measure of uncertainty. One possibility is that by doing so we have created an alternative measure of order clarity. This would be true if, for example, there was no variation in clarity within a type of policy challenge. The Court does sometimes issue clear remedies when our rough measure would treat the general policy challenge as involving more uncertainty. For example, we treat remedies that are designed to "Investigate criminal activities" as situations in which courts are likely to have greater uncertainty about the means-ends problem relative to, say, asking a state to make an apology. At a bare minimum, the investigative process, done badly, can reinforce harms already done. They may also create new human rights violations, by say inducing officials of the state to intimidate witnesses as they carry out the criminal investigation. Yet in such cases, our coders found that only about one half of the remedies were vague. The standard example of a vague order to investigate, which we expected to find in a larger percentage of classes, is reflected

well by Escher y Otros v. Brasil, a case dealing with the Brazilian state's illegal monitoring of lawyers. Here the court ordered a criminal investigation as follows: "The State must investigate the facts that gave rise to the violations in the instant case." willCompare that order to an order to investigate in the Montero Aranguren case, seeking an end to what it regarded as a continuing *de facto* amnesty, which we qutote only in part here:

> *The State must, within reasonable time, remove all obstacles, de facto and de jure, that perpetuate impunity; provide adequate safety guarantees to the victims, other witnesses, judicial officers, prosecutors, and other relevant law enforcement officials; and use all means at its disposal to expedite the investigation and judicial process, in order to identify, prosecute and punish the perpetrators of the violent events and adopt any necessary emergency measures at the Prison; of the excessive use of force and of the extrajudicial execution of several inmates . . . The State must secure that the victims' next of kin have full access and capacity to act in all stages and proceedings taken in the course of said investigations, according to the domestic laws and the provisions of the American Convention. The outcome of the investigations shall be published by the State, in such a way to allow the Venezuelan society to know the truth about the events of the instant case . . .*

Although our measure of uncertainty is coarse, it is unlikely that the policy-making challenge of the remedy is identical to the way in which the remedy is expressed. If anything, our measure of x_1 contains greater random measurement error than we would like, which may be suppressing a potentially stronger effect than we observe. Nevertheless, it is possible that dichotomizing the policy challenges in the way that the uncertainty measure does may obscure variation useful for evaluating the implication we seek to test. To consider this possibility, we fit an alternative model of Court remedy clarity. Here, we estimate the probability of a clear rule assuming that the rather than being nested within cases and states, the i remedies are nested within j policy challenges. We then model a random intercept for policy challenges.[12] Specifically, we estimate the following model.

[12]The Appendix discusses the results of fitting a linear probability model with fixed effects for order types. The results are substantively similar and statistically significant.

$$Pr(y_{1ij} = 1) \quad = \quad logit^{-1}(\alpha_0 + \alpha_2 x_{2jk} + \psi_j) \tag{3}$$

Here we assume that $\psi_j | x_{ij} \sim N(0, \gamma)$. Table 4 summarizes the empirical Bayes prediction for the random intercepts from this model. The types of orders in the left column were estimated to be less clear on average, given the human rights score for that state, whereas the types of orders in the right column were estimated to be clearer than the average order. We have bolded the types of orders that reflected our concept of policy challenges that would have involved a more challenging mapping between the policy requested and the desired outcome. The Court has been less clear when dealing with policy challenges that appear to involve a less certain mapping between policy instrument and policy outcome. There are two exceptions. In our sample, the court was somewhat less clear about what would ultimately count as an appropriate acknowledgement of responsibility or an appropriate public release of information about the case to the state's public.

Less Clear Orders	Clearer Orders
Modify Law	**Non-Pecuniary Damages**
Acknowledge Responsibility	**Pecuniary Damages**
Release Information	**Publish Decision**
New Programming	**Expunge Criminal Record**
Investigate Crime	**Sponsor Scholarship**
Non-Monetary Restitution	

Table 4: *The table classifies the random intercepts estimated from a model in which orders are assumed to be nested within policy challenges as measured by the type of order being considered. The right column shows types of orders that were estimated to be clearer than the average order whereas the left column shows types of orders that were estimated to be less clear than average. The bolded types of orders were originally coded as reflecting higher uncertainty about the mapping between a policy request and a policy outcome.*

Interpretation

The theoretical model that guides it provides a causal interpretation of the findings, but we wish to be careful not to over-claim about what we can learn from these data. To structure how we interpret the findings, it is worth considering what would appear to be the most salient threat to causal inference in any model supposing that compliance is a function of opinion clarity, i.e., a

model of $y_2|y_1$. Ideally, we would like to randomize the clarity of orders. Although this may be feasible in future applications, the use of the Court's data obviously implies that clarity will not have been randomized; and, absent theoretical structure, we might be suspicious of a simple assumption suggesting that we can treat the clarity of these orders as if they were randomly assigned. A natural concern is that judges are clear when they expect compliance and vague when they suspect non-compliance and thus any positive association between clarity and compliance is simply tapping into the judiciary's concern with hiding non-compliance.

Fortunately, in this application, we are not without theoretical structure. Importantly, the model does identify an incentive to use vagueness to hide non-compliance; however, we also know when that this most likely to happen. This behavior is predicted for courts that are particularly concerned with the institutional implications of being defied. By 2005, in light of its long record of making principled decisions that are routinely resisted and often outright defied (Cavallaro and Brewer, 2008), it seems unlikely that the Court is the kind of court that is particularly concerned with hiding non-compliance through vagueness. Indeed, it is sometimes extremely clear when ordering states to carry out tasks that states never do (see discussion below on the ordering of criminal investigations). Further, if the court had these preferences, the model suggests that there would be no variation in clarity. All orders would be vague. They are not.

Most importantly, the model suggests a system of equations in which opinion clarity is determined by factors that are exogenous to compliance outcomes. The key question is whether it is reasonable to assume that the policy changes that the Court believes are necessary can be treated as if they were exogenous to order clarity. For example, the need for training the police in international norms regarding the interrogation of criminal suspects is something that follows immediately from the recognition of a criminal process that clearly violates these norms. Yet the types of policy changes that the court views as necessary to remedy violations of international law are just another part of the judicial decision-making process. A full treatment of this question would identify all possible reparations that could be required, but it is certainly not true that the Court only picks challenges to address that would result in clear orders that will be obeyed. Indeed, the majority of its clear orders are defied. The bottom line is that whereas we have reason to accept key results as offering *prima facie* evidence of causal relationships, we should retain a healthy degree of skepticism.

Conclusion

In order for courts to fully constrain government behavior, they must be able to issue decisions that command compliance when governments would rather not, at least on occasion. Just like prior studies, we find considerable resistance to Court orders in a variety of cases. Having said that, it is neither true that the Court is always ignored nor that it has no influence over the process. Like other courts in other contexts, the IACtHR can exercise a measure of control over compliance via the clarity with which it expresses itself. Although clarity cannot resolve defiance in all cases, it influences state behavior on the margin, and this may be the best an international court can hope for.

Although international courts may be able to maximize pressure for compliance through unambiguous language, it is not clear that maximal clarity is optimal. We find support in IACtHR data for a theoretical model in which opinion vagueness is driven by pragmatic concerns of policy-making. Pragmatism on the Court, in the form of vague remedies, may seem to undermine the protection of human rights, by rendering a potentially powerful court weak. Yet, this kind of pragmatism may be unavoidable. The primary implication of our study for the promotion of constrained government underscores a central insight from positive political analysis of law and courts (e.g. Epstein and Knight, 1998; Rogers, 2001; Carrubba, Gabel and Hankla, 2008) – in so far as judicial policy-making is a fundamentally interdependent process, judicial decision-making is, indeed it must be, sensitive to external political environments. Although judges may give up some control over policy outcomes, they gain the ability to better tailor some policy responses to particular problems.

Pelc's contribution to this special issue raises a question in light of our findings and research on the use of dialogic remedies. Pelc's own focus on language concerns the potential persuasive effects of affect in written text. We do not concern ourselves with the persuasive effects of clear language, but it is worth asking whether this might be a focus for a future study. From our perspective, the primary problem with vague language is that it is more difficult to identify non-compliance and bring pressure to bear on non-compliant actors when it is not clear what a court has asked for. It is at least plausible that vague remedies fail to persuade states that it is necessary to change their behavior. Rather than inviting experimentation over how to come into line with international obligations, it may be that dialogic remedies undermine cooperation. Answering this question lies

outside the scope of our analysis, but future work might directly address the persuasive aspects of dialogic engagement with states.

Stepping outside of the context positive political analysis of law, our study also illustrates important insights from international law scholarship in the constructivist tradition. In this tradition, legal norms are understood to develop via processes of contestation about meaning (Adler, 2005; Wiener, 2014). In developing an account of legal change, Jutta Brunnée and Stephen Toope (2018) place place emphasis on what they call the "practice of legality." Legal norms are developed via common legal practices that give meaning rules of behavior that are tied to rule of law criteria elaborated by Fuller (1969) including *inter alia* generality, publicity, prospectivity and clarity. Variation in the clarity of language used by judges in pursuit of remedies for legal violations might be understood as a manifestation of contests over legal meaning, which this tradition views as an essential piece of the process by which legal norms are constructed.

Our results also address global concerns for the promotion of judicial independence. If our results meaningfully reflect the context in which international courts like the IACtHR operate, then it may be unwise to continue promoting judicial independence only. Concerning international courts, this argument and a variety of its implications, is developed in greater detail by Posner and Yoo (2005). Posner and Yoo's central concern is that you cannot easily force states to make changes in policies to which they have fundamental commitments. This may be an impossible task in the context of a powerful state. We do not deny the challenge Posner and Yoo raise, but the problem is not just that courts frequently cannot gain compliance. That much is likely true. The critical point is that we may not want to maximize compliance. This is not because of an abstract violation of democratic commitments to majoritarianism or sovereignty. Simply put, courts, certainly international courts, often do not have the kind of information necessary to match policy means to policy ends, and for that reason, they will want to depend on the information domestic actors can bring to bear. Vagueness addresses this problem. The cost of vagueness is to give up some measure of control over policy outcomes. Managing this tradeoff is the central challenge of pragmatism.

If we must have pragmatic judges, we need to ask how to train them to be pragmatic without giving up all concerns with constraining states. Police patrol monitoring schemes like the those of the Inter-American Court and the European Court of Human Rights seem like absolutely necessary

components of such an approach. It is one thing to be vague in order to allow for the efficient use of information. It is quite another to use vagueness simply to avoid being exposed as weak. A vigorous monitoring system (carried out honestly of course) can advance the first goal without promoting pragmatism of the second sort. This is but one institutional approach to addressing the issue. We are not wed to a particular strategy for encouraging more effective legal institutions, but we are confident that the right approach needs to take seriously the political challenges and constraints judges confront as they attempt to enforce legal obligations.

References

Basch, Fernando, Leonard Filippini, Ana Laya, MarianoNino, Felicitas Rossi and Barbara Schreiber. 2010. "The Effectiveness of the Inter-American System of Human Rights Protection: A Quantitative Approach to its Functioning and Compliance with its Decisions." *International Journal on Human Rights* 7(12):9–36.

Binder, Christina. 2012. "The prohibition of amnesties by the Inter-American Court of Human Rights." *International Judicial Lawmaking* pp. 295–328.

Brewer-Carías, Allan. 2009. *Constitutional protection of human rights in Latin America: a comparative study of amparo proceedings*. Cambridge University Press.

Brinks, Daniel M and William Forbath. 2010. "Social and economic rights in Latin America: Constitutional Courts and the prospects for pro-poor interventions." *Tex. L. Rev.* 89:1943.

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

Carrubba, Clifford J. 2009. "A Model of the Endogenous Development of Judicial Institutions in Federal and International Systems." *Journal of Politics* 71(1):1–15.

Carrubba, Clifford J., Matthew Gabel and Charles Hankla. 2008. "Judicial Behavior under Political Constraints: Evidence from the European Court of Justice." *American Political Science Review* 435–452(4):109.

Carrubba, Clifford J and Matthew J Gabel. 2014. *International courts and the performance of international agreements: A general theory with evidence from the European Union*. Cambridge University Press.

Cavallaro, James L. and Stephanie Brewer. 2008. "Reevaluating Regional Human Rights Litigation in the Twenty-First Century: The Case of the Inter-American Court." *The American Journal of International Law* 102:768.

Chayes, Abram and Antonia Handler Chayes. 1993. "On Compliance." *International Organization* 47(2):175–205.

Chilton, Adam S and Mila Versteeg. 2016. "Do Constitutional Rights Make a Difference?" *American Journal of Political Science* 60(3):575–589.

Clark, Tom S. 2010. *The Limits of Judicial Independence*. New York: Cambridge University Press.

Dancy, Geoff and Verónica Michel. 2016. "Human rights enforcement from below: Private actors and prosecutorial momentum in Latin America and Europe." *International Studies Quarterly* 60(1):173–188.

Downs, George W, David M Rocke and Peter N Barsoom. 1996. "Is the good news about compliance good news about cooperation?" *International Organization* 50(3):379–406.

Dulitzky, Ariel E. 2015. "An Inter-American Constitutional Court-The Invention of the Conventionality Control by the Inter-American Court of Human Rights." *Tex. Int'l LJ* 50:45.

Epstein, Lee. and Jack Knight. 1998. *The choices justices make*. CQ Press.

Fariss, Christopher J. 2014. "Respect for human rights has improved over time: Modeling the changing standard of accountability." *American Political Science Review* 108(2):297–318.

Gauri, Varun and Daniel M. Brinks. 2008. "Introduction: The Elements of Legalization and the Triangular Shape of Social and Economic Rights". In *Courting Social Justice: Judicial Enforcement of Social and Economic Rights in the Developing World.*, ed. Varun Gauri and Daniel M. Brinks. Cambridge University Press pp. 1–37.

Greene, William H. 1998. "Gender economics courses in liberal arts colleges: Further results." *The Journal of Economic Education* 29(4):291–300.

Guzman, Andrew and Jennifer Landsidle. 2008. "The Myth of International Delegation." *California Law Review* 96:1693.

Hafner-Burton, Emilie M. 2005. "Trading human rights: How preferential trade agreements influence government repression." *International Organization* 59(3):593–629.

Hashimoto, Barry. Forthcoming. "Why do autocrats accept the International Criminal Court's jurisdiction? Theory and." *International Organization* .

Hawkins, Darren and Wade Jacoby. 2010. "Partial Compliance: A Comparison of the European and Inter-American American Courts for Human Rights." *Journal of International Law and International Relations* 6(1):35–85.

Helmke, Gretchen. 2010. "Public Support and Judicial Crises in Latin America." *U. Pa. J. Const. L.* 13:397.

Hillebrecht, Courtney. 2009. "Rethinking compliance: the challenges and prospects of measuring compliance with International Human Rights Tribunals." *Journal of Human Rights Practice* 1(3):362–379.

Hillebrecht, Courtney. 2010. The Effect of International Human Rights Tribunals on Domestic Practice and Policy PhD thesis University of Wisconsin-Madison.

Hillebrecht, Courtney. 2014. *Domestic politics and international human rights tribunals: the problem of compliance*. Vol. 104 Cambridge University Press.

Hogg, Peter and Allison A Bushell. 1997. "The charter dialogue between Courts and Legislatures (or perhaps the charter of rights isn t sucha bad thing after all)." *The Osgood Hall Law Review* 35:1.

Huber, John D. and Charles R. Shipan. 2002. *Deliberate Discretion: The Institutional Foundations of Bureaucratic Autonomy*. Cambridge University Press.

Huneeus, Alexandra. 2011. "Courts Resisting Courts: Lessons from the Inter-American Court's Struggle to Enforce Human Rights." *Cornell International Law Journal* 3(3).

Kapiszewski, Diana and Matthew M Taylor. 2013. "Compliance: conceptualizing, measuring, and explaining adherence to judicial rulings." *Law & Social Inquiry* 38(4):803–835.

Keith, Linda Camp. 2002. "Judicial Independence and Human Rights Protection Around the World." *Judicature* 85(4):195–200.

Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The rational design of international institutions." *International organization* 55(04):761–799.

Larsson, Olof and Daniel Naurin. 2016. "Judicial independence and political uncertainty: how the risk of override affects the Court of Justice of the EU." *International Organization* 70(2):377–408.

Law, David S and Mila Versteeg. 2013. "Sham constitutions." *California Law Review* pp. 863–952.

Lax, Jeffrey R. 2012. "Political constraints on legal doctrine: how hierarchy shapes the law." *The Journal of Politics* 74(03):765–781.

Linzer, Drew A and Jeffrey K Staton. 2015. "A global measure of judicial independence, 1948–2012." *Journal of Law and Courts* 3(2):223–256.

Lopez, George A, Jackie Smith and Ron Pagnucco. 1998. "Globalizing human rights: the work of transnational human rights NGOs in the 1990s." *Human Rights Quarterly* 20(2):379–412.

Lupu, Yonatan. 2013. "Best evidence: The role of information in domestic judicial enforcement of international human rights agreements." *International Organization* 67(3):469–503.

Lupu, Yonatan. 2015. "Legislative veto players and the effects of international human rights agreements." *American Journal of Political Science* 59(3):578–594.

Marshall, Monte and Keith Jaggers. 2005. *Polity IV Project: Dataset UsersÂ Manual*. Center for International Conflict Data Management. available online at: http://www.cidcm.umd.edu/polity/.

McCubbins, Mathew D, Daniel B Rodriguez and Barry R Weingast. 2009. "The rule of law unplugged." *U of Texas Law, Public Law Research Paper* (158).

Neumayer, Eric. 2005. "Do international human rights treaties improve respect for human rights?" *Journal of conflict resolution* 49(6):925–953.

Pasqualucci, Jo M. 2003. *The Practice and Procedure of the Inter-American Court of Human Rights*. Cambridge Univ Press.

Pasqualucci, Jo M. 2012. *The practice and procedure of the Inter-American Court of Human Rights*. Cambridge University Press.

Posner, Eric A and John C. Yoo. 2005. "Judicial Independence in International Tribunals." *California Law Review* 93(1):1–74.

Prillaman, William C. 2000. *The judiciary and democratic decay in Latin America: Declining confidence in the rule of law*. Greenwood Publishing Group.

Raustiala, Kal. 2000. "Compliance & (and) Effectiveness in International Regulatory Cooperation." *Case W. Res. j. Int'l L.* 32:387.

Reisman, Michael. 1995. "Practical Matters for Consideration in the Establishment of a Regional Human Rights Mechanism: Lessons from the Inter-American Experience." *St. Louis-Warsaw Transatlantic Law Journal* 89.

Ríos-Figueroa, J. 2007. "Fragmentation of Power and the Emergence of an Effective Judiciary in Mexico, 1994–2002." *Latin American Politics & Society* 49:31–57.

Ríos-Figueroa, Julio. 2016. *Constitutional Courts as Mediators: Armed Conflict, Civil-military Relations, and the Rule of Law in Latin America*. Cambridge University Press.

Rodríguez-Garavito, César. 2010. "Beyond the courtroom: The impact of judicial activism on socioeconomic rights in Latin America." *Tex. L. Rev.* 89:1669.

Rogers, James R. 2001. "Information and Judicial Review: A Signaling Game of Legislative-Judicial Interaction." *American Journal of Political Science* 45(1):84–99.

Schlag, Pierre. 1985. "Rules and standards." *UcLA L. REv.* 33:379.

Sieder, Rachel, Line Schjolden and Alan Angell. 2016. *The judicialization of politics in Latin America*. Springer.

Sikkink, Kathryn. 1993. "Human rights, principled issue-networks, and sovereignty in Latin America." *International Organization* 47(3):411–441.

Sikkink, Kathryn. 2005. The transnational dimension of the judicialization of politics in Latin America. In *The judicialization of politics in Latin America*. Springer pp. 263–292.

Simmons, Beth A. 2000. "International law and state behavior: Commitment and compliance in international monetary affairs." *American Political Science Review* 94(4):819–835.

Simmons, Beth A. 2009. *Mobilizing for human rights: international law in domestic politics*. Cambridge University Press.

Spriggs, James F. 1997. "Explaining Federal Bureaucratic Compliance with Supreme Court Opinions." *Political Research Quarterly* 50(3):567.

Staton, Jeffrey K. and Georg Vanberg. 2008. "The Value of Vagueness: Delegation, Defiance, and Judicial Opinions." *American Journal of Political Science* 52(3):504—519.

Vanberg, Georg. 2005. *The Politics of Constitutional Review in Germany*. New York: Cambridge University Press.

Voeten, Erik. 2014. "Domestic Implementation of European Court of Human Rights Judgments: Legal Infrastructure and Government Effectiveness Matter: A Reply to Dia Anagnostou and Alina Mungiu-Pippidi." *European Journal of International Law* 25(1):229–238.

Von Stein, Jana. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance." *American Political Science Review* 99(04):611–622.

Waldron, Jeremy. 2002. "Is the Rule of Law an Essentially Contested Concept (in Florida)?" *Law and Philosophy* 21(2):137–164.

Wilson, Bruce M. 2009. "Institutional reform and rights revolutions in Latin America: The cases of Costa Rica and Colombia." *Journal of Politics in Latin America* 1(2):59–85.

Appendix

	Compliance	Resistance	Clarity
Clear Remedy	1.83^{***}	-3.17^{***}	–
	(0.57)	(1.22)	
Human Rights Context	0.46	−0.21	−0.10
	(0.37)	(0.14)	(0.20)
Judicial Confidence	0.72	−0.91	–
	(0.57)	(0.87)	
Uncertain Policy Area	–	–	-3.49^{***}
			(0.51)
Constant	-1.95^{***}	4.51^{***}	3.06^{***}
	(0.52)	(1.17)	(0.43)
N	266	265	263
Cases	42	42	45
States	15	15	17
$\upsilon^{(case)}$	0.81	4.34	–
$\upsilon^{(state)}$	1.19	5.60	–
$\gamma^{(case)}$	–	–	.38
$\gamma^{(state)}$	–	–	$< .01$

Table 5: *Compliance and Clarity Models. Effect of remedy clarity (Clear), domestic judicial confidence (Costs), and human rights index (Record) on state reactions to Court remedies and effect of policy uncertainty (Uncertain) and human rights record on remedy clarity; GLS estimates shown (standard errors in parentheses); dependent variable for the first model is the probability of the court finding compliance; dependent variable for the second model is the probability of observing state resistance; estimated variances for case andstate random intercepts not displayed; judicial confidence has been mean centered;* $^{***}p \leq 0.01^{**}p \leq 0.05;^{*}\, p \leq 0.10$ *(two-tailed).*

	Model 1	Model 2	Model 3
Clear Remedy	1.83***	1.84***	1.82***
	(0.57)	(0.57)	(0.57)
Human Rights Context	0.47	0.39	0.11
	(0.31)	(0.39)	(0.25)
Judicial Confidence	0.71	0.77	0.40
	(0.92)	(0.97)	(0.71)
NGO Support	−0.21	–	–
	(0.22)		
Executive Constraints	0.16	–	–
	(0.54)		
Latent Judicial Independence	–	1.53	–
		(3.40)	
Med. Bureaucratic Quality	–	–	1.65*
			(0.94)
High Bureaucratic Quality	–	–	3.01***
			(1.08)
Constant	−2.87	−2.70**	−3.33***
	(3.35)	(1.84)	(1.06)
N	266	266	266
Cases	42	42	42
States	15	15	15
$\gamma^{(case)}$	0.76	0.79	0.63
$\gamma^{(state)}$	1.38	1.29	0.91

Table 6: *Alternative Models of the Compliance Equation Original model plus indicators for the number of NGOs in support of the plaintiff, Polity IV's Executive Constraints scale, Tate and Keith's judicial independence measure, and PRS's indicator of bureaucratic quality; GLS estimates shown (standard errors in parentheses); dependent variable is the probability of observing state resistance; estimated variances for case andstate random intercepts not displayed; judicial confidence has been mean centered;* $^{***}p \leq 0.01$ $^{**}p \leq 0.05$;$^{*}\, p \leq 0.10$ *(two-tailed).*

Alternative measure of remedy types Hillebrecht (2014) conceptualizes orders as involving reparations, symbolic measures, accountability, non-repetition and individual measures. Our measure of uncertainty is derived from our measures of the order types. Our coding at the level of the remedy, while not identical to Hillebrecht's conceptualization, allows us to consider results in light of that alternative. A potential concern is that since we know that uncertainty does seem to be associated with compliance, it is possible that some of the particular order types might be strongly associated with compliance as well as clarity. Thus the association between clarity and compliance may be confounded. In fact, we know that this is possible in the context of orders to investigate the security sector, which are strongly and negatively associated with compliance. Of course, since there is variation in clarity in that context it it not clear that we confront omitted variable bias. Nevertheless, we can address this possibility by fitting a linear probability model in which we include fixed effects for the order types. Although this does not involve using the same categories as used by Hillebrecht, it fully addresses the concern since our codings would have to be aggregated up to match Hillebrecht's categories. The positive and strong association between clarity and compliance is robust to adding remedy fixed effects. The model suggests that orders to modify laws, to release information, to provide non-monetary restitution as well as orders to create scholarships are negatively associated with compliance.

Excluding Criminal Investigations The Court has been notoriously unsuccessful in commanding criminal investigations of the security sectors of defendant states (Basch et al., 2010). Our measure of uncertainty treats criminal investigations as the kind of policy challenge that would create significant uncertainties. A judge may know that a criminal investigation should be conducted, but be uncertain about how exactly to carry it out in particular contexts. Yet we know that the Court has simply not be able to get states to investigate their police or military. Figure 4 shows a comparison of the results of our models when we omit orders in which the court has required a criminal investigation. As in clear, there are no meaningful differences in the results.

Coding Resistance This variable is derived from a categorical indicator we have also coded. We use the binary variable in this application in order to balance the desire to keep observations which are coded originally as an "unknown level of resistance" against the need to have to assign a value to such remedies in order to generate an ordinal indicator. The original variable is measured as

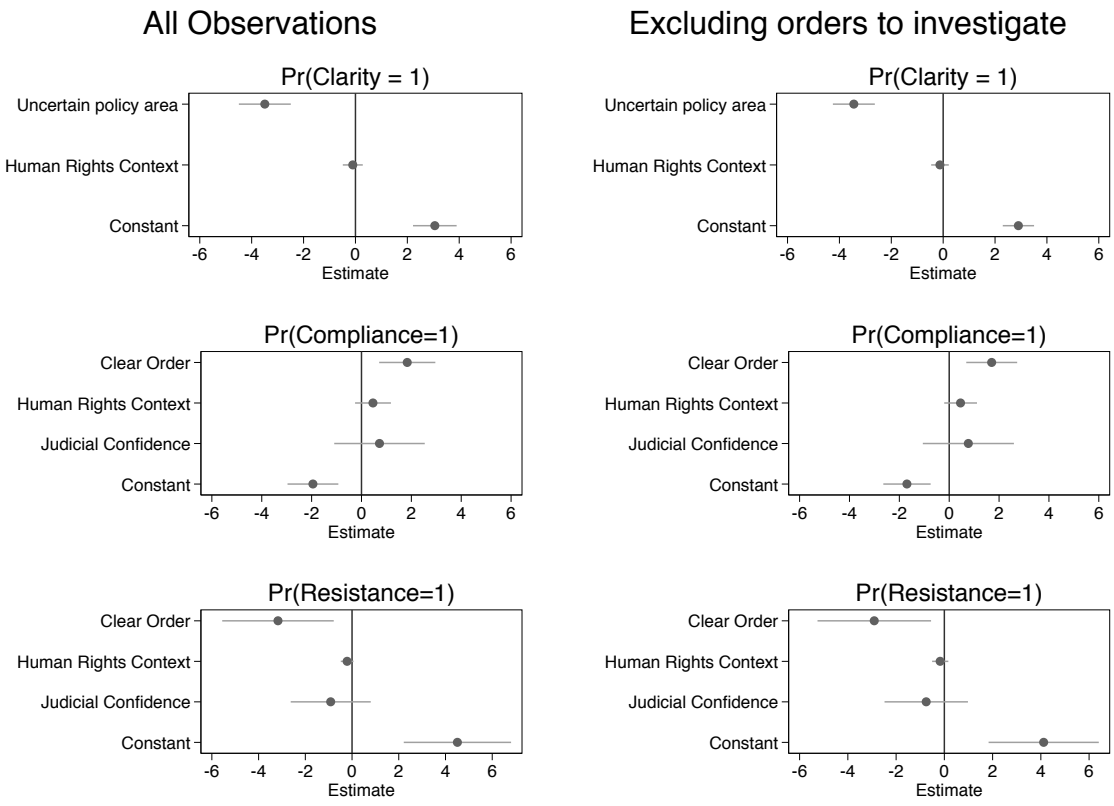

Figure 4: *Comparison of results to models excluding criminal investigations. The figure summarizes models in which we discard orders to engage in criminal investigations. The results in the left column are the original findings. Those in the right reflect the results when excluding criminal investigations.*

follows: (0) *No resistance* – The state complied with the remedy and evidence of that compliance can be seen in the Compliance Report; (1) *Minor resistance* – The state engaged in some delay, but eventually complied with the remedy. No evidence in the compliance reports of the Court expressing significant concern over delays. Compliance is evident by the second compliance report; (2) *Substantial resistance* – The state delayed or continues to delay the implementation of the remedy considerably and theres some evidence in the compliance record of the Court expressing concern over State delay. The implementation, if it occurs it, does not occur until compliance report 3 or later; (3) *Massive resistance* – The State has failed to implement the remedy despite long-term (e.g. 4 Compliance Reports or more) monitoring from the court. Considerable evidence in the compliance record of the Court expressing significant concern over delay. Despite at least 4 compliance reports, there is no evidence that the remedy has been implemented; (4) *Uncertain Resistance* – Some level of resistance is evident since the remedy has not been implemented as of the latest compliance report on file, but due to the small number of reports (2 or fewer) it is difficult to assess the level of resistance. y_2^r is coded such that it is 1 if any resistance is detected.

Clarity Coding We also have coded the number of words the court used to describe each remedy, which has been used by legislative scholars as proxy for discretion (e.g. Huber and Shipan, 2002). Huber and Shipan claim in the legislative context that word count is a reasonable proxy for clarity, claiming that longer legislative language is associated with clearer statements, which in turn provide less discretion to bureaucracies. The mean number of words for the remedies that we code as clear is 165. The mean number of words is 101 for remedies that we code as vague. The p-value for the t-test of the difference of these means is 0.0002. Clear remedies tend to have more words relative to vague remedies, as Huber and Shipan found in the legislative context.