A Public Strategy for Compliance Monitoring

Varun Gauri
Jeffrey K. Staton
Jorge Vargas Cullell

The World Bank
Development Research Group
Human Development and Public Services Team
June 2013

Abstract

Theoretical accounts of compliance with court orders emphasize the importance of transparency. Most empirical studies of compliance center on high profile political cases, largely ignoring the high-volume, quotidian claims against the state for basic services that constitute the largest share of court dockets in many jurisdictions. This paper uses a unique dataset on compliance with orders from the Constitutional

Chamber of the Supreme Court of Costa Rica to examine the determinants of compliance in low salience cases. It finds that orders issued just after the Court announced, in a press conference, that it was monitoring compliance were implemented roughly two months sooner than orders issued just prior to the press conference. These findings suggest that publicity can motivate compliance even in low salience cases.

This paper is a product of the Human Development and Public Services Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The author may be contacted at vgauri@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

A Public Strategy for Compliance Monitoring¹

Varun Gauri
Development Research Group
The World Bank

Jeffrey K. Staton
Department of Political Science
Emory University

Jorge Vargas Cullell
Director Adjunto
Programa Estado de la Nación

JEL: K00, K40

Keywords: Costa Rica, litigation, constitutions, compliance, transparency

Sector Boards: Public Sector Management, Human Development

¹ For comments, the authors thank participants at workshops on Judicial Institutions at Princeton University, on Global and International Studies at George Washington University, at the Latin American Studies Association, and at the Law and Society Association. The views and findings in this paper are those of the authors alone and do not necessarily represent the views of the World Bank or its Executive Directors. The authors can be reached at vgauri@worldbank.org (Varun Gauri), jkstato@emory.edu (Jeffrey Staton), and jorgevargas@estadonacion.or.cr (Jorge Vargas Cullell).

Introduction

In October 2009, the Constitutional Chamber of the Supreme Court of Costa Rica (colloquially, the “Sala Cuarta” or “Sala IV”) launched an innovative program to monitor compliance with direct orders made under its amparo (constitutional rights) and habeas corpus jurisdiction.² The majority of cases the Court now monitors involve *individual* claims for medications, pensions, or labor rights, rather than the typically more salient concerns of high politics. These are not cases typically covered by the media, and the officials subject to the orders did not originally anticipate that their behavior would become public knowledge. But surprisingly, six months after beginning the program, on March 2, 2010, the Sala IV held a press conference announcing its preliminary results. The event, advertised one week in advance, was well-attended and received careful coverage in the media.³

Figure 1 shows a PowerPoint slide from that press conference.⁴ The final column of the table revealed a shockingly low percentage of orders for which the Sala IV was able to document certain compliance. The key problem, summarized in the fifth column, was that the court was unable to get public officials to answer requests for information. Even the highly respected Department of Social Security (the “Caja”) had a level of documented compliance far below what had been expected. The Sala IV promised to continue monitoring cases and revealed a plan to post compliance rates on its website. The Sala IV’s compliance system and its decision to go public with its early results offer a unique opportunity to evaluate compliance in the high-volume, quotidian matters of constitutional law that make up a large share of dockets for many courts, especially those in Latin American constitutional systems.

²This program was developed following a proposal from Vargas Cullell and Staton for a field experiment to evaluate the effect of increasing beliefs about issue transparency on compliance. Although all three of the authors of this paper have played a consultancy role in some aspects of the system, the Sala IV’s monitoring protocol was developed and is administered by the Sala itself.

³ “Sala IV presentará plan para confirmar cumplimiento de sentencia, Radio Relo, 24 Feb. 2010.; “Seguimiento de Sentencias,” La Nación, 3 Mar. 2010; “Gobierno con baja nota en ejecución de fallos de Sala IV,” Diaro Extra, 3 Mar. 2010.

⁴ Sala Constitucional, “Sistema de seguimiento de sentencias de la Sala Constitucional,” 2 March 2010.

**Sala IV: Sentencias por certeza de cumplimiento según institución:
Octubre – Noviembre 2009**

INSTITUCION	Cumplimiento	Cumplimiento parcial	Incumplimiento	Ignorado	Total	% certeza cumplimiento
<u>Ministerio de Educación Pública</u>	20	1		93	114	18%
Caja Costarricense del Seguro Social	70	2	3	27	104	67%
Ministerio de Educación Pública	4	1		22	27	15%
<u>Ministerio de Gobernación, Policía y Seguridad Pública</u>	7			6	13	54%
Ministerio de Justicia	6			7	13	46%
Poder Judicial	9			1	10	90%
<u>Ministerio de Obras Públicas y Transportes</u>	1			7	8	13%
Autoridad Reguladora de los Servicios Públicos	2			4	6	33%
Ministerio de Ambiente, Energía y Telecomunicaciones	1			5	6	17%
Ministerio de Trabajo y Seguridad Social	1		1	4	6	17%
Resto	20	0	3	93	117	17%
Total	141	4	7	269	424	33%
	33%	1%	2%	63%	100%	

Figure 1: *Compliance Rates in October-November, 2009. Compliance information by institution as presented by the Sala IV compliance team at their March 2, 2010 press conference announcing the monitoring system.*

Analyses of compliance with court orders have made considerable progress in recent years. It is now widely understood that judicial orders are frequently not self-enforcing, and that their proper implementation depends on choices and efforts made by a variety of actors outside the judiciary, including the executive branch, private firms, and citizens (Rodríguez Garavito 2011; Vanberg 2005). The area of constitutional review, where judges often depend on one arm of the state to ensure that another arm complies, is especially challenging. Explanations of compliance with constitutional orders have focused on circumstances in which political fragmentation and the proliferation of veto players creates political roadblocks, facilitating the emergence of judicial power (Ames, 2001; Mainwaring, 1999; Whittington, 2005); the value courts provide to

policy makers when they solve monitoring problems in decentralized settings (Kagan, 2001; Miller and Barnes, 2004); and the ability of courts to ally with key interest groups, or with the public at large, for addressing popular causes, often in conjunction with the media (Gauri and Brinks, 2008; Staton, 2010; Carrubba, 2009).

This last line of scholarship is of particular interest for present purposes. But popular causes typically require high visibility, whereas high volume constitutional review conducted in many jurisdictions, including constitutional claims in Costa Rica, Colombia, and Brazil and administrative courts in many other countries, have extremely low public visibility. Ordinarily, the principal actors in these cases believe it unlikely that their actions will become public. It is in this regard that the Sala IV's press conference is particularly interesting. We leverage the unexpected nature of the event, its heightened media coverage and explicit plan for publicizing non-compliance, to evaluate whether increasing the transparency of constitutional review influences compliance outcomes.⁵ We find that orders issued after the press conference were implemented roughly two months sooner than orders issued prior to the press conference. We attempt to rule out other, unobserved causes by estimating this effect within smaller and smaller temporal windows surrounding the press conference, and by examining other "fake" press conference dates for discontinuities. We also find similar results when controlling for observable features of the cases that might have changed following the press conference event, including the

⁵For certain international courts, there are a few empirical studies of the determinants of compliance in low salience cases. Staton and Romero (2011) find that clarity in the orders of the Inter-American Court of Human Rights lowers state resistance to implementation. Gre-wal and Voeten (N.d.) find that while bureaucratic capacity seems to matter for some cases, it becomes much less important for high salience cases involving controversial issue areas, such as torture and inhumane treatment. Of course, the related literatures on compliance with high salience international court rulings (von Staden, 2009), and with international human rights treaties (Simmons, 2009; Dai, 2002), have found that domestic checks and balances may promote compliance (while also, perhaps, reducing the incentive to ratify). But there are, to our knowledge, no systematic empirical studies of compliance with low salience court rulings against the executive branch in domestic jurisdictions.

clarity with which the court articulated its orders and the type of demands litigants brought to the Court.

The next section of our paper develops the theoretical structure of the study, drawing on existing models of judicial politics. We then discuss constitutional review in Costa Rica and introduce the Sala IV's compliance monitoring system. The results of our study follow. We conclude our paper by discussing a few questions our study raises.

Compliance Monitoring

The great majority of courts worldwide rely on indirect monitoring of compliance in areas of administrative law and constitutional rights violations. That is, once a ruling supporting a litigant's claim has been issued, judges assume that the litigant will return if the executive branch fails to provide relief. If the litigant does not return, courts assume that she is satisfied. Indeed, this process is a special case of a broader "fire alarms" approach to monitoring the implementation of public policy (McCubbins and Schwartz, 1984). The assumption that aggrieved parties will return to court if they are not satisfied may be justified for cases in which litigants enjoy privileged access to the legal system (as in constitutional clashes among government actors), cases in which litigants are well-resourced or well-organized (such as corporations, certain NGOs, unions, and class action cases), or cases in which the cost of non-compliance is, for the litigant, much higher than the cost of more litigation (such as claims for protection against torture). But the assumption may not hold for a class of cases, typically claims against the state for resources, information, or action, whose expected net value may not justify further litigation, and for certain classes of litigants, including those who are resource-

constrained. These take the form of administrative law claims in many jurisdictions and constitutional actions in others.

A more direct or “police patrol” approach to monitoring compliance with court rulings might take a variety of forms, arising on the initiative of various actors. Executive branch agencies could assess the extent to which their own frontline bureaucrats comply with administrative law judgments. External organizations, such as the ombudsmen of northern Europe or the Ministério Público in Brazil, could audit executive agencies for compliance with court orders. The courts themselves could systematically review compliance with their orders and initiate *suo moto* proceedings (or referrals to prosecuting authorities) in cases of non-compliance, maintain cases in their dockets even after verdicts are rendered so that evidence can be collected during the implementation phase (as in the South African “structural interdict” or Indian public interest litigation), or conduct informal meetings with executive agencies who are the frequent targets of administrative law or constitutional law claims. Finally, courts could review compliance rates of various agencies and publicize their findings.

The Costa Rican Sala IVs approach to compliance monitoring has included several of these elements, but a key aspect of its process, and the one on which we focus here, has been to generate publicity about compliance patterns – to raise awareness among agencies and to make it easier for the media to follow up on problem areas. The most obvious manifestation of that strategy was its March 2, 2010 press conference. Has this strategy been successful, and if so, why?

Transparency in Low Stakes Constitutional Review

Models of constitutional review in comparative politics suggest that the transparency of constitutional cases is a critical condition for judicial power.⁶ Courts are less likely to be

⁶ By transparency, we refer to a extent to which actors immediately involved in the case believe that impartial third parties are both likely to learn about the case and understand what is required by the decision.

deferential to governments when the public is likely to be aware of non-compliance because officials are more likely to comply when they are being watched (Carrubba, 2009; Vanberg, 2005). Empirically, scholars have found that non-compliance is more likely in cases that are not covered by the media (Staton, 2010) and when orders are designed in ways that make it difficult to conclude that an official has defied a court (e.g. Spriggs, 1997; Staton and Romero, 2011). In other words, non-compliance is more likely in less transparent contexts.

Yet existing theoretical arguments about the role of transparency have been developed in a context of relatively high stakes constitutional review, where a significant proportion of the docket deals with the actions of public officials who hold high office and where the complaints deal with relatively salient public policies. We believe that the logic of these accounts applies in the context of lower level bureaucratic responses, as well, though the channels of influence are indirect and varied. Among the mechanisms we envision is the potential link between the credibility of the court's promise to publicize non-compliance to elected officials' reputational concerns, and ultimately to their ability to direct street level bureaucrats to respond effectively.

What are the conditions under which publicity of this sort might affect the behavior of relevant parties? First, relevant parties must believe that the court's commitment to monitor and publicize compliance rates in the future is credible. That credibility, in turn, depends on the autonomy of the court and its perceived effectiveness. In other words, publicity of this sort is likely to affect relevant parties much more strongly in environments in which judicial power is relatively high. Costa Rica is such an environment (Wilson and Rodr'iguez Cordero, 2006). Figure 2 displays Linzer and Staton's (2012) *de facto* judicial independence estimates for the Americas, where judicial independence is scaled from zero (low) to one (high). The Costa Rica series is highlighted by a red box. Reflecting regional expert knowledge, these estimates suggest that the Costa Rica's judiciary has much more in common with the judiciaries of the North American states (upper left) than with many other Latin American systems, especially those of Central America.

Second, we assume that senior public officials care about their reputations. For political appointees, reputations likely affect the chances of being reelected and/or impact upon their ability to enact important policies and programs while in office. For senior civil servants, reputations often affect career prospects, and sometimes affect their bargaining power in negotiations over agency budgets and policy outcomes.

Third, we assume that the public at large attaches significance to executive branch compliance with constitutional orders from the courts. Repeated failure to comply with court orders, in the eyes of the public, sometimes constitutes a salient breach of the promises and oaths public leaders have committed to support. Because many of the amparo cases in Costa Rica, and elsewhere in Latin America, involve health, social security, and labor rights, the public can also interpret repeated instances of non-compliance as general indifference to the cause of social justice. While non-compliance in any particular case might be excusable due to circumstances beyond the control of the executive branch, the demonstration of repeated and routine non-compliance is assumed to be troubling to citizens in polities where there is widespread commitment to the constitutional order and/or to the particular rights enshrined in the constitution.

Finally, we assume that street level bureaucrats, who ultimately issue the directives to comply with court orders (e.g., to provide a particular person with medical treatment, a pension, or a bail hearing), are responsive to pressures from their superiors.

The press conference in which the court characterized compliance rates and committed itself to further monitoring would, based on the assumptions above, activate pressure on bureaucrats to increase compliance rates. But there are other possible mechanisms through which transparency could increase compliance. For instance, it is quite possible that non-compliance occurs, in many cases, for reasons related to bureaucratic capacity rather than bureaucratic or political resistance. The agencies might simply be overwhelmed. Under those circumstances, when courts assess compliance by calling agencies to inquire about the status of particular amparo cases, that might have the direct effect of making those cases and claims

more salient to street level bureaucrats, somewhat independent of pressures placed on them by senior public officials. Another possibility is that, under circumstances of limited capacity or limited funding to respond to demands for social rights demands, agencies develop internal procedures for prioritizing or rationing service delivery. When courts assess compliance with individual cases, agencies receive information that the litigant is still dissatisfied and may be a source of further trouble (e.g., further litigation or further direct pressure on the agency). Seeking to avoid such trouble, street level bureaucrats in charge move those litigants to the front of the service delivery queue.

In summary, although models of constitutional review that emphasize transparency as a critical condition for compliance and ultimately judicial power were developed in the context of high stakes constitutional review, we believe that these arguments apply well in cases of relatively low stakes review. The core empirical implication of the argument is that officials should be more likely to implement orders faithfully and quickly when they believe that non-compliance is likely to be made public. For this reason, we would expect better compliance outcomes in Costa Rica in the wake of the Sala IV's compliance monitoring system announcement.

Figure 2: *Linzer and Staton Measure of Latent Judicial Independence for the Americas (and a few Europeans because we were too lazy to remove them), 1960-2010.*

Constitutional Review in Costa Rica

Although the practice of judicial review in Costa Rica goes back to the late 19th century, a series of changes in 1989 dramatically increased its utilization and scope. A constitutional amendment that year created a new chamber of the court with the power of nationally centralized judicial review. It also lowered the requirement for a vote of unconstitutionality from two-thirds to a simple majority of the Court. At the same time, the enabling law for the newly created Sala IV eased the barriers to access so that anyone in the country, even a foreign national, could file a writ petition twenty-four hours a day, without the need for legal representation, and without charge (Wilson, 2011). Finally, reformers with ties to the Inter-American Court of Human Rights influenced the early decision making of the newly created Sala IV, encouraging the enforcement of social and economic rights that, while incorporated into the 1949 Constitution, had previously been understood to be non-justiciable and merely aspirational.⁷

As a consequence of these changes, the caseload of the Sala IV increased sharply, and nearly linearly with time, from about 1,600 amparo filings in 1989 to about 20,000 in 2010. In a typical year, the Sala IV rejects a bit more than half of all amparo filings, and then supports about two-thirds of the cases that go on to receive a vote on the merits. For the public, the popularity of the procedure stems from the low costs of filing, the ability of petitioners to lodge constitutional claims directly and in the first instance (without first having exhausted the available administrative law procedures), the practice of a quick judicial response, and the historical willingness of the Sala IV to consider and then enforce new kinds of constitutional rights. In turn, the popularity of the procedure burnishes the image of the Sala IV, which, along with the requirement that significant changes in its structure and functioning require a super-majority vote of the Legislative Assembly, has protected it from political backlash.

⁷ Julio Jurado, personal communication, 2011.

At the time that the Court began to monitor compliance with amparo orders, estimates of compliance rates ranged from nearly universal compliance (a view expressed by some members of the legal and judicial community) to widespread defiance (a view expressed by many critics and by the cynical press). It is noteworthy that even in a country whose constitutional court enjoyed substantial power and popular support, compliance rates were unknown, and that there many who believed that the Court was unable to enforce compliance. The difficulty, for the Court, of legally enforcing compliance has stemmed from the fact that the principal procedure for enforcing compliance is the *desacato* (loosely, contempt of court), which in Costa Rican law is a criminal charge under the control of the public prosecutors and not the Sala IV, and which requires a demonstration that there exists a non-compliant individual who knowingly and intentionally violated a court order. In most cases, it is nearly impossible to establish that a particular bureaucrat purposefully denied a benefit to someone when defense claims of bureaucratic capacity and accidental oversight are widely thought to be credible.

The Sala IV's Compliance Monitoring System

The Sala IV's monitoring system departs from a simple concept: asking the parties involved whether a Court's order has been fulfilled. The system begins with a compliance team working in the Centro de Jurisprudencia Constitucional (CJC), an administrative office of the Sala IV. The CJC identifies rulings issued by the Justices (magistrados) in the previous month and records the time frame or deadline (*plazo*) attached to each sentence.

When the *plazo* comes due, lawyers on the compliance team call the claimant on behalf of the Sala IV to inquire into the status of the claim. If the claimant is satisfied that the order has been implemented, lawyers certify the answer and register it, including date and time of calling.

On the other hand, if the claimant reports an instance of non-compliance, the team calls the responsible authority for an explanation. Calls are repeated until the CJC has a specific answer about the case's status; however, after five calls, the case is registered as an instance of non-compliance. All calls and answers are recorded in the system in order to have a detailed track record for each case.

The final step in the process involves the CJC lawyers grading the status of compliance based on the specifics of the court's order and the information collected. The CJC sends a detailed monthly report to the Sala's Presidency and to each of the Justices for follow-up in the event that responsible authorities refuse to comply. Legally, the Sala IV can accuse the defendant of *desacato*, but is up to the Prosecutor General to take the case into the criminal justice system; and this has proven challenging.

Critically for our purposes, the CJC records detailed information on the parties to the case, the nature of the complaint and order, the response of the responsible authority, as well as a number of additional features of the process, including the timing of votes and responses.

Analysis

Our analysis makes use of the CJC monitoring data, beginning in October, 2009, and ending in December, 2011. We seek to estimate two kinds of information. Given enough time, the vast majority of *amparos* are implemented per the Sala IV's requirements;⁸ however, there is considerable variation in the time it takes to implement an order. In so far as delayed implementation is a form of non-compliance (Staton, 2010), we focus our analysis on estimating

⁸As of May 2, 2013, the CJC had tracked 11,052 of the 11,363 orders issued by the Sala IV between October, 2009, and March 25, 2013. Among the orders that were tracked, the CJC was able to verify compliance in 9,391 instances (85%).

the duration of non-compliance via an event history approach. Second, we also estimate the probability of an order being defied at all. The CJC compliance team indicates, for each order, whether there has been compliance, probable compliance, partial compliance, or non-compliance. Although we find the results estimating the existence of noncompliance illuminating, our decision to focus on durational data is particularly appropriate in light of an unfortunate feature of the CJC's early data collection strategy, which makes it easier to interpret time until compliance than compliance itself. The critical problem is that the CJC "wrote over" some data, so that although we know very well when a case has been fully implemented, there is some uncertainty about whether there was ever a form of non-compliance prior to implementation.

Critically for the duration analysis, the CJC compliance team recorded the month and year of the Sala IV's vote in each case and the date of compliance. Figure 3 displays a histogram of the number of months orders spent in the CJC system prior to a registration of certain compliance. The minimum number of months is 1 (when compliance is documented within a month of the Sala IV's decision) and the maximum is 25. The median duration is 7 months. Our data are right-censored in all cases of unresolved partial or complete noncompliance. There are no cases that are left-censored, as we know the date of the vote for every case in the sample. Because we are only able to measure the month and year of compliance, the data set gives rise to a number of tied events, i.e., there are a number of months in which multiple orders were resolved, which present their own challenges in an event history context.

Figure 3: *Months in Compliance Monitoring System: Histogram of the number of months orders spent in the compliance monitoring system prior to compliance or right censoring. The mean 8.1 and the median is 7.0.*

The Press Conference

We estimate Cox proportional hazards models of the months to compliance (denoted t), accounting for ties with the Efron method. The hazard function for the i^{th} amparo or habeas order in the Cox model is given by:

$$hi(t) = h0(t) \exp(\beta'xi),$$

where $h0$ is the baseline hazard function, permitted to take any form, and then scaled by a linear combination of covariates. Coefficients are estimated via the partial likelihood method and standard errors are clustered at the level of the case.⁹

The key independent variable for us is a dummy variable coded 1 if the Sala IV vote was held on or after April 1, 2010, and coded 0 otherwise. The logic of this design is that claimants who bring an amparo order to a responsible authority will get more favorable treatment in the wake of the press conference than they would have prior to the event. The press conference was held at the beginning of March. Because we do not know the exact date of the each decision (only that they were decided in March), it was not clear how to code cases in March. Clearly, all votes in April happened after the press conference; thus, we can be sure that all orders in cases voted upon in April will have been issued after the Sala IV's attempt to increase the transparency of the amparo and habeas process. Of course, all orders, except those resolved prior to March 2, 2010, will have been exposed to the press conference.¹⁰ If all orders were influenced by press conference in the same way, independent of when the Sala IV's vote occurred, the effect will be an attenuation of the April vote date estimate. Thus the design is biased toward a null finding. A similar attenuation effect follows from treating all cases in March as if they were voted on prior to the press conference.

⁹ Tests of the proportional hazards assumption were inconsistent. Results for log-logistic models of duration were consistent with what we report here.

¹⁰ The immediate implication of this point is that we cannot treat the date as an assignment variable in a regression discontinuity design. Thus, causal inference will not be as clean as it would be in a stronger design.

Figure 4 shows hazard ratios for the April month threshold for four samples of orders. Every hazard ratio is above one, with the largest at nearly two, indicating a doubling in the odds of an order coming into compliance after the press conference than before. The upper left panel shows the results for all votes in the database. Moving from west to east and then south, the figure shows the results for smaller and smaller windows of data, closing in on the period directly surrounding the press conference. A key concern is that there might have been some unrelated change in the bureaucracy that took place during 2010, which makes it look like something changed in response to the press conference, even if the conference itself was irrelevant. We make a number of efforts to address this issue, the first of which involves limiting the sample to those cases that were resolved essentially during the same period. For this reason, we believe that the most conservative estimate of the press conference effect lies in the lower right corner of the figure.

Figure 4: *Cox Proportional Hazard Models: Panels display estimated hazard ratios with 95% confidence intervals. From the northwest to the southeast, panels show results for increasingly restricted samples, closing the window around the votes taken in the months immediately before and after the press conference.*

Likely Sources of Endogeneity

The results in Figure 4 are suggestive, but their persuasiveness depends on believing that orders were assigned randomly to the periods prior to and following the press conference. There is, quite obviously, no way to instrument for the press conference. In a sense, we (the authors) were the instrument, as our request generated the monitoring system itself. Absent the ability to travel backward in time and convince the Sala IV to randomize the agencies whose records were reported, and without a way to instrument for the event itself, we will have to rely on weaker designs, focusing on observable indicators.

There are two very plausible threats to the assumption that the press conference can be treated as if it were randomly assigned to the stream of case files the Sala IV resolves. The first is that litigants brought distinct kinds of cases before and after the press conference. A simple argument is that prior to the press conference, individuals were less likely to bring cases for which they perceived the likelihood of compliance was low than after the press conference, anticipating the transparency effect discussed above. Such a process would mean that the Sala IV confronted more difficult cases to resolve after the press conference, which again would attenuate any effect we find. Nevertheless, it is worth considering that possibility. The second threat is that the Sala IV itself started issuing different kinds of orders following the press conference. If the Sala IV anticipated more transparency after the press conference, it might have been more willing to express clearly what it wanted after the press conference; and, if clarity produces more compliance (e.g. Spriggs, 1997), this would explain the results in Figure 4.¹¹

Before describing how we address these effects, it is worth noting that a full assessment of a public strategy for compliance would include them, rather than identifying only the effects of the press conference on the bureaucracy through public pressure. Over the medium- to long-term, it is to be expected that transparency with regard to compliance will change the

¹¹ Theoretical results in Staton and Vanberg (2008) suggest that although clarity ought to increase compliance, the effect of an increase in public support (as say via transparency) on the clarity of an order is indeterminate.)

composition of cases that litigants bring to courts, and will change how judges themselves write their opinions (and perhaps how they vote, as well). These should not be understood as collateral, surprising consequences of transparency but rather as *anticipated and intended* effects of a public strategy for compliance. If judges write clearer orders in response to transparency, and if that makes it easier for agencies to comply, that is as much a benefit of transparency as pressure being brought to bear on street level bureaucrats. These effects would likely take time to crystallize, however, because both litigants and judges will not change their expectations and ingrained habits quickly. For that reason, in the short run (in the months immediately following the press conference) the effects of transparency are more plausibly attributed to the effect of transparency on the bureaucracy than on litigant behavior or judicial practice. Still, to identify the effect on the bureaucracy, we attempt to control for observables that measure case composition (the outcome of litigant behavior) and the clarity of the orders (the result of judicial practice).

The ease or difficulty of implementing cases can be measured by two kinds of features of the orders. First, we include a series of dummy variables for salient “plazo” categories. Orders that give public entities considerable time to work out how to rectify a constitutional violation may reflect policy challenges, e.g. the fixing of a drainage system, that are simply difficult to resolve. We develop indicators for the following plazos: One week, One Week to Two Weeks, Two Weeks to One Month, One Month to Six Months, More than Six Months, and Sin Plazo. The “immediate” plazo reflects our base category. The primary rationale for cutting the plazo variable into categories is that roughly 38% of the sample is without a plazo, terms that afford great flexibility in implementation but raise questions about how to precisely code them relative to other plazos, especially relatively long ones. Second, we also include dummy variables indicating particular agencies that typically confront relatively easily implemented orders or relatively difficult ones, or which have reputations for responding quickly to judicial orders. Specifically, we include dummy variables for the Caja, the Ministry of Health, and the Ministry

of Public Works, entities that the CJC itself had suggested were particularly likely to either resolve cases faster (the Caja) or slower (Ministry of Health and Public Works).¹²

To measure the clarity of the order, we make use of a binary scale where 1 indicates that the action required of the target of the order was “clear and definite” and 0 otherwise. We assigned the orders randomly to two research assistants, with the exception of 200 orders, which they coded together. The coders agreed on 85% of the cases. For these orders, we randomly coded of one of the two coders.

To illustrate, the following order, Order No. 16059 of 2009, was not considered clear and definite: “It is ordered that Alberto [last name withheld] and Maria [last name withheld], in their respective capacity as Director of Human Resources and Chief of the Department of Accounting, both of the Ministry of Public Education, or whosoever is acting their offices, do the necessary so to pay the claimant, Teresita [last name withheld], identity number [withheld], the necessary salary adjustments arising from disability, in installments, taking reasonable account of her own needs and those of her family.” The last clauses of the sentence introduce elements of vagueness with regard to whether compliance has occurred.

By contrast the following order, Order No. 13941 of 2011, was coded as clear: “It is ordered that Hilda [last name withheld], in her capacity as Director General, and Jaime [last name withheld], in his capacity as Chief of Urology, both of the hospital San Juan de Dios, or whosoever is acting in their offices, take the necessary steps and execute the relevant actions, within their respective powers and competencies, so that the claimant Juan [last name withheld] receives the transurethral resection of the prostate that he needs, within two months of the date of this communication.”

¹² In addition, we fit models with controls for whether the defendant was a municipality, as well as a number of other agencies. Only the three we include influenced the estimated baseline hazard function. Finally, we have fit models with only the April 1 2010 threshold. Results of those models are only stronger than those reported, suggesting that correcting for the mix of cases by plazo and agency was important.

Overall, the coders were able to identify the clarity of 72% of the orders in our dataset. The remaining 28% could not be coded for clarity because they required actions whose complete characterization relied on other legal proceedings not in the database (e.g., orders to respond to previous right to information orders that were not themselves described in the *amparo* orders in our database). Of the 3121 orders that were coded for clarity, 2556 (82%) entailed actions that were “clear and definite”; and 565 (18%) were not clear.

Models with Control Variables

Figure 5 displays the results for the Cox models with the agency, plazo and clarity control variables. The baseline plazo is an order where the term is immediate. Thus, observing hazard ratios below 1 in these models suggests that every plazo longer than an immediate plazo is likely to decrease the hazard of compliance. The strongest effect is for plazos that are longer than six months, which have a hazard rate estimated to be 40% lower than the baseline. The hazard ratios for the Caja are consistently greater than 1, suggesting increasing hazards of compliance for that agency, whereas the Ministries of Public Works and Health both seem to have decreasing hazards of compliance. The only radically different finding is for the Public Works Ministry in the sample centered on the press conference date. The clarity measure also suggests an increasing hazard of compliance. Finally, across all samples, the April 2010 threshold variable has a hazard ratio between 1 and 2, as would be true if orders voted on after April 1, 2010 have hazard rates that are higher than those voted on prior to April 1, 2010.

Figure 6 plots the estimated survivor functions for the 2010 models with all other variables set at their modal values. The median survival time for such orders is 11 months. As the northwest panel suggests, it is 9 for the orders in cases voted on after the press conference (consider the number of months when the survivor function equals .5 for the median estimated survival time). A similar effect is estimated for the Ministry of Health, though the survival time increases. The Caja seems to have had a very strong effect. Indeed, the median survival time for an order being responded to by the Caja is only 6 months. And unsurprisingly, the plazos

themselves have strong estimated effects. Indeed, there is a 5 month difference in estimated survival times for a case where the Sala IV has ordered immediate compliance and a case in which the plazo is greater than 6 months.

Figure 5: Cox Proportional Hazard Models: Panels display estimated hazard ratios with 95% confidence intervals. From the northwest to the southeast, panels show results for increasingly restricted samples, closing the window around the votes taken in the months immediately before and after the press conference.

Figure 6: *Estimated Survivor Function: Panels plot the estimated survivor function with respect to the vote threshold variable, indicators for the Caja and Ministry of Health, and for three types of plazas, with all other variables held at their modal values.*

Fake Dates

Our study suggests that something likely changed in the Costa Rican constitutional compliance process during March 2010. This is consistent with an effect of increased transparency, caused by the Sala IV's press conference and associated publicity. Of course, it is possible that the change, in so far as there was one, occurred much earlier or much later, and had nothing to do with the press conference. We consider four alternatives. Figure 7 shows estimated hazard ratios for four models. The first column shows models for a January and May 2010 thresholds. The second column shows models for August and November thresholds. There is no evidence that there was a change at the beginning of the 2010, well before the press conference, or at the end of 2010, well after it.

Figure 7: *Cox Models for Fake Press Conference Dates: Panels display estimated hazard ratios with 95% confidence intervals.*

Instances of Non-Compliance

Perhaps the simplest approach for evaluating the effect of the press conference would involve comparing compliance rates prior to and after the event. In so far as the CJC compliance team indicates, for each order, whether there has been compliance, partial compliance, or non-compliance, this would appear straightforward. An unfortunate data management strategy makes the interpretation results derived from such a strategy difficult. Consider Table 1, which summarizes the compliance record for the Caja in October and November of 2009. The second column presents the frequencies of compliance outcomes as reported in March 2010 (compare to the second row of the table in Figure 1). The third column shows the same information as reported by the CJC compliance team in their cleaned data set published in November 2011. Critically, whereas the Sala IV reported a 67% certain compliance record in March 2010, by November 2011, their cleaned dataset suggested that it was 100%.

Level of Compliance	Press Conference (March, 2010)	Cleaned Dataset (November, 2011)
Compliance	70	106
Partial Compliance	2	0
Non-Compliance	3	0
Unknown	27	0
Total	104	106
Percentage certain compliance	67	100

Table 1: *Two Pictures of the Caja's Compliance Record for October-November, 2009.*

Why explains difference? A part of the story has to do with the “Unknown” category. Between March 2010 and November 2011, the CJC was able to track down information about the *Caja's* reactions. But this is not the entire story, as it does not explain what happened to the five orders that were coded as *partial compliance* or *non-compliance*. The answer lies in data management software. The CJC was using MS Excel to manage their

data and they simply wrote over instances of non-compliance after the agency successfully implemented an order. Thus, although we know that the *Caja* implemented all orders voted on in October and November of 2009, we do not know when they did so or how many cases involved significant delay.

Table 2 reinforces the problem. As is clear, the percentage of cases in which there was an incident of overt non-compliance increases over time – indeed, it jumped 117% between 2009 and 2011. The reason is that, as we saw with the *Caja* example, there were many cases in 2009 and 2010 that were once coded as non-compliant but have since been changed. An insufficient amount of time had passed in 2011 to reduce these instances of non-compliance via overwriting data.

Year	Non-Compliance %
2009	6
2010	8
2011	13

Table 2: Percentage of orders that resulted in non-compliance, by year

Having noted these concerns, it is nevertheless worth considering what the data have to say. Figure 8 summarizes the results of two models. The first model is a logit regression of a binary measure that takes the value of 1 when the CJC team codes the order as involving an instance of “non-compliance;” it is 0 otherwise. We do not include the plazo indicators as we believe they are best for picking up orders that may be difficult to carry out quickly, but not necessarily difficult to carry out per se. The second model is an ordered logit regression of the four-category CJC indicator (0 for non-compliance, 1 for partial compliance, 2 for probable compliance and 3 for compliance). We restrict the sample to cases resolved in the period immediately surrounding the press conference, i.e., from February to May, 2010. As the models suggest, the press conference threshold is positive for the ordered logit model of compliance, but essentially zero for the logit model of certain non-compliance. Of particular note is the extremely

strong and negative effect of order clarity on the probability of complete non-compliance. Indeed, the predicted probability of non-compliance for a clear order is only .03 (.01, .07), whereas this probability jumps to .11 (.03, .19) for a vague order.

Figure 8: *Logistic regression models: Panels display estimated coefficients with 95% confidence intervals. The top panel shows results for the binary indicator of non-compliance. The April month dummy should be negative in this model, reflecting a decrease in the probability of complete non-. The bottom panel shows results for the four category indicator of compliance. The April month dummy should be positive in this figure, reflecting an increase in the probability of a case taking on one of the codes associated with compliance*

Conclusion

We find evidence that the Sala IV's decision to publicize the preliminary results of its monitoring program had demonstrable effects on the timing of compliance under its amparo and habeas corpus jurisdiction. Specifically, orders issued after the press conference were estimated to be implemented roughly two months earlier than orders issued before the press conference, controlling for the order's plazo, the order's clarity, and whether the order was being acted upon by agencies suspected of being below or above average responders. These findings are consistent with models of judicial politics that afford considerable causal weight to issue transparency, even in a legal context of high volume, low stakes constitutional review.

Nevertheless, it is worth being careful when interpreting the findings. Absent a strong design (e.g. a regression discontinuity or an experiment) it is difficult to conclude definitively that the March 2, 2010 press conference is solely responsible for a change in compliance times. During the same period, the Sala IV did more than hold a simple press conference announcing the results – the President of the Constitutional Chamber discussed their findings with agency heads and gave interviews on the subject more broadly. We prefer to view the press conference date as the most visible manifestation of the period in which, and the general strategy through which, the Sala IV went public with its findings. The April 1 threshold dummy variable captures that understanding and it is in this sense that we view the evidence.

Although the evidence is consistent with models of judicial politics in which transparency is critical, it is certainly true that the study cannot adjudicate perfectly among more precise statements about the effect of transparency. On one account, the Sala IV simply made it more likely that bureaucrats and their superiors would have to explain to a skeptical public why they were failing to implement court orders. The mechanism connecting this kind of transparency to

compliance runs though bureaucrats expected disutility from being publicly exposed. An alternative understanding is that the Sala IV's efforts generated an honest reevaluation within agencies about their own processes of compliance, a reevaluation entirely disconnected from the potential (external) concerns of a disapproving public. Adjudicating among these two alternatives will likely require data on bureaucratic responses to the press conference, data that might reveal the extent to which agents engaged the public with their efforts to rectify their problems. It is also unlikely that distinct public entities responded in the same ways. If the public regarding account is correct, we should be most likely to see public responses from responsible authorities that are most directly accountable to voters, e.g. ministries.

A number of open questions remain. For one, we might ask why the court would issue unclear orders when the clarity of those orders seems to be powerfully related to non-compliance. We might wish to know if police patrol monitoring itself has an impact on compliance. Originally, the Sala IV had only monitored cases in which responsible authorities know they were being monitored. Recently, the CJC completed an effort to go back and monitor the outcome of a sample of cases about which it had not previously informed the responsible authorities. The results of this study will help quantify the effect of direct monitoring.

We might also ask what impact the system has had on the Sala IV itself. What has it learned? Have these lessons influenced their jurisprudence? Have they affected their political relations with other actors in the Costa Rican state? And surely it is worth asking about the ways in which direct monitoring has influenced litigants. In what ways, if any, did the initial findings mobilize civil society? Has there been an impact on the mix of cases coming to the Sala IV?

Finally, we ought to ask whether there is something special about Costa Rica that explains these results. Given that the institutional structure for prosecuting non-compliance functions less well here than it does in, say, Mexico, it is possible that direct monitoring in the way the Sala IV has moved forward would be less impactful where institutions of accountability work better. The opposite could be true, as well. That is, direct monitoring may be a complement to the fire alarm mechanisms of a typical constitutional system. Answering these questions obviously requires a more comprehensive design, one that is at once cross-nationally comparative and yet highly detailed in its research design, from theoretical argument to strategies for data integrity. Given the pervasiveness of high volume constitutional review in Latin American and the importance of understanding compliance in legal systems, we hope that answering these questions will draw interest from a variety of scholars.

References

- Ames, Barry. 2001. *The Deadlock of Democracy in Brazil*. Univ of Michigan Press.
- 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.
- Dai, Xinyuan. 2002. "Political Regimes and International Trade: The Democratic Difference Revisited." *American Political Science Review* 96(1):159–165.
- 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.
- Grewal, Sharanbir and Erik Voeten. N.d. "The Politics of Implementing European Court of Hights Judgements." Working Paper.
- Ho, Daniel E, Kosuke Imai, Gary King and Elizabeth A Stuart. 2007. "Matching as non-parametric preprocessing for reducing model dependence in parametric causal inference." *Political analysis* 15(3):199–236.
- Kagan, Robert A. 2001. *Adversarial Legalism: The American Way of Law*. Harvard University Press.
- Linzer, Drew and Jeffrey K. Staton. 2012. "A Measurement Model for Synthesizing Multiple Comparative Indicators: The Case of Judicial Independence." Paper presented at the Annual Meeting of the American Political Science Association.
- Mainwaring, Scott. 1999. *Rethinking party systems in the third wave of Democratization: The Case of Brazil*. Stanford Univ Press.

- McCubbins, M.D. and T. Schwartz. 1984. "Congressional oversight overlooked: Police patrols versus fire alarms." *American Journal of Political Science* pp. 165–179.
- Miller, Marc C. and Jeb Barnes. 2004. Putting the Pieces Together: American Lawmaking from an Interbranch Perspective. In *Making Policy, Making Law: An Interbranch Perspective*, ed. Marc C. Miller and Jeb Barnes. Georgetown University Press.
- North, Douglass, John Wallis and Barry Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge University Press.
- Simmons, Beth A. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics*. New York: Cambridge University Press.
- Spriggs, James F. 1997. "Explaining Federal Bureaucratic Compliance with Supreme Court Opinions." *Political Research Quarterly* 50(3):567.
- Staton, Jeffrey K. 2010. *Judicial Power and Strategic Communication in Mexico*. New York: Cambridge University Press.
- Staton, Jeffrey K. and Alexia Romero. 2011. "Clarity and Compliance in the Inter-American Human Rights System." Paper presented at the American Political Science Association Meetings, September 1-4, 2011.
- 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.
- von Staden, Andreas. 2009. "The Politics of the Rule of Law." *International Studies Review* 11(2):358–364.

- Whittington, Keith E. 2005. "Interpose your friendly hand'O: Political Supports for the Exercise of Judicial Review by the United States Supreme Court." *American Political Science Review* 99(4):583–596.
- Wilson, B.M. and J.C. Rodr'iguez Cordero. 2006. "Legal opportunity structures and social movements." *Comparative Political Studies* 39(3):325–351.
- Wilson, Bruce M. 2011. Costa Rica Health Rights Litigation: Causes and Consequences. In *Litigating Health Rights: Can Courts Bring More Justice?* Harvard University Press pp. 132–154.