The Costa Rican Supreme Court’s Compliance Monitoring System

Varun Gauri, World Bank
Jeffrey K. Staton, Emory University
Jorge Vargas Cullell, Programa Estado de la Nación

In the summer of 2009, the Constitutional Chamber of the Supreme Court of Costa Rica began monitoring compliance with its direct orders in amparo and habeas corpus cases. The court announced the early results from its analysis at a well-attended March 2010 press conference. The president of the court promised to continue monitoring and publishing the results for the foreseeable future. We use a unique data set on compliance derived from this monitoring system to evaluate theoretical claims about the relationship between the transparency of judicial orders and compliance. We observe that vague orders, and orders issued without definite time frames for compliance, were associated with delayed implementation. We also find that orders issued after the press conference were implemented roughly two months earlier than orders issued just prior to the press conference.

A core element of the rule of law is that courts should be capable of remedying violations of legal obligations (Raz 1997, 218). To do so, relevant parties must comply with direct judicial orders. Judges themselves value compliance (e.g., Huneeus 2010; Widner 2001) in part because compliance is a key component of judicial power (Cameron 2002). Important factors that promote powerful courts rest largely beyond judicial control. Most obviously, judges are unlikely to have an immediate and strong influence on the degree to which political power is fragmented (Chávez, Ferejohn, and Weingast 2011; Ríos-Figueroa 2007) or on the drafting of formal rules that insulate themselves from external pressure (Pozas-Loyo and Ríos-Figueroa 2010). But some factors may be subject to judicial influence. Compliance, and judicial power more generally, depends on public support, which in turn is related to the transparency of the conflicts courts resolve because without at least the possibility of informing people about noncompliance, public support does not matter (Vanberg 2005; Yadav and Mukherjee 2014); and transparency is something that judges can influence.

In June 2009, we began a discussion with the Constitutional Chamber of the Supreme Court of Costa Rica (colloquially, the “Sala Cuarta” or “Sala IV”) concerning potential experimental research designs aimed at understanding better their compliance process. Instead of conducting experimental research, however, the Sala IV built its own system for monitoring compliance with all direct orders to public officials in its amparo and habeas corpus jurisdictions. In October, it began to quietly track reactions to its orders. Six months later, the Sala IV held a press conference to announce its preliminary results, which called into question the compliance record of major arms of the Costa Rican state. The press conference, advertised one week in advance, was well-attended and received careful coverage in the media.1
The Sala IV promised to continue monitoring cases and announced a plan to post compliance rates on its website.

The Sala’s compliance system and its decision to go public with early results presents a unique opportunity to study an aspect of the constitutional review process—enforcement—that is often either assumed or overlooked. Although there are numerous studies of judicial behavior in comparative politics, studies of the compliance process are largely limited to case studies (Gauri and Brinks 2008; Rodriguez-Garavito 2011; Vanberg 2005). The Sala’s surprise press conference suggests a way of evaluating the consequences of a one-time increase in transparency. The process by which the system was announced complicates causal inference, but we believe that there are nevertheless important lessons to learn from the Sala’s experience if researchers are transparent about their assumptions and the threats to causal inference they encounter. Viewed within the general research program on constitutional politics and the judicial role in the enforcement of rights, the findings are consistent with one mechanism by which compliance can be promoted—transparency. They also suggest a number of opportunities and challenges for future studies in which stronger designs are possible.

We divide the remainder of our article as follows. The next section of our article develops the theoretical structure of the study and discusses why the setting of our study is particularly useful. We then present a study that speaks to the effect of the press conference, as well as several theoretically plausible alternatives. We conclude our article by discussing some more general questions that the study raises.

TRANSPARENCY AND COMPLIANCE
Our study departs from a simple theoretical claim. Given the judiciary’s lack of coercive powers, its ability to induce compliance in many instances of constitutional review depends on the reputational incentives that public officials encounter. These incentives are strong when courts enjoy considerable public support or legitimacy (for recent expressions of this idea, see Carrubba, Gabel, and Hankla 2008; Clark 2010; Rodriguez-Garavito 2011). In such environments, officials believe that there are consequences for noncompliance, whether in the form of ballots, public protests, or reputation risk. For this to work, however, the public must be aware of the cases judges resolve, or at the very least, information must be available such that the public could be made aware of them. And for that reason, transparency, by which we mean the extent to which actors immediately involved in the case believe that third parties can learn about the case and understand what is required by the decision, is thought to promote compliance (Vanberg 2005). Indeed, because transparency affects judicial power by leveraging public support, judges have strong incentives in many cases to promote it (Krehbiel 2013; Staton 2010).

The Sala IV’s public announcement of its compliance system itself constitutes evidence consistent with this logic, but our study focuses on the compliance behavior of public officials. We evaluate whether compliance on the part of public officials, another implication of this logic, improved following the Sala’s press conference. In evaluating this claim, it is nevertheless important to recall that compliance takes place within a broader context of interbranch relations, which can influence the decision-making process in a variety of ways (Helmke 2005; Kapiszewski 2011). Our understanding of this context influences where to test our expectations and what threats to inference to consider. Notably, judges maintain control over features of the legal process, which affect the kinds of orders we observe. Clearly, judges control the choice over whether to find a constitutional violation. They also maintain control over the ways in which they ask for remediation in the event of a violation. The question is, what does judicial control over these elements of a case mean for our effort to learn about the effect on compliance of announcing the monitoring system? Our view is that this question is best answered within the context of a particular model of interbranch politics, which deals not only with strategic constitutional review and transparency but with two distinct types of transparency. We thus structure our discussion through the lens of the model developed in Staton and Vanberg (2008).

Two aspects of transparency
A judicial order can be viewed as transparent because people know about it or can be informed about it. This kind of transparency, what we refer to as “awareness,” is directly connected to public support itself by making it possible to leverage whatever support there happens to be. Empirically, this is the type of transparency that the press conference and the Sala’s monitoring system could have plausibly influenced. It is the kind of transparency addressed by Drezhsel (1986), Krehbiel (2013), and Staton (2010). Importantly though, judges also have control over a second type of transparency—the “clarity” or “vagueness”

2. Formal prosecution of public officials for failure to comply with judicial orders does occur in some jurisdictions, but for reasons we discuss later, this is rare.

3. For a closely related argument, see Davis (1994).

4. On this interpretation, the press conference could have served to increase the public backlash parameter, which reflects the consequences to the public official of noncompliance.
with which they express their order. By expressing clearly what is expected of an official in order to comply with the decision, a judge makes it difficult for an official to argue that an order has been implemented when in fact it has not. The press conference per se would not have influenced this type of transparency. Clarity varies at the level of orders. It captures whether an order is interpretable, irrespective of the level of public awareness. The critical question is whether an increase in awareness would have been likely to influence the clarity with which judges craft orders and, via that process, compliance.

The core idea of the model is that judges can use order vagueness for two distinct purposes (Staton and Vanberg 2008, 506–8). The first is to manage a means-ends problem in the process of judicial policy implementation. Given fundamental uncertainties about the policy-making process, judges are not always sure about how to translate a policy request into a desirable outcome. Order vagueness permits agencies to use their expertise to achieve the ends desired by judges. On the other hand, vagueness invites noncompliance by lowering the costs of delay or outright defiance, allowing agents to claim that they have implemented an order without actually doing so. Noncompliance in this context might be costly, when agents use their discretion to pursue divergent goals, but compliance with a poorly constructed remedy might be costly as well. In such cases, greater discretion permits agents to save judges from themselves. Thus, in the context of the means-ends problem, drafting orders in real language involves a trade-off between maintaining control over how orders are precisely implemented and capturing the informational gains associated with delegating power to actors with greater expertise. The second use of vagueness is to mask overt noncompliance when overt noncompliance would be particularly costly. Here, the idea is that courts can find it costly when they are observed being ignored. In this context, vagueness serves simply to hide such outcomes, largely because what has been asked for is ambiguous. Naturally, in order to take advantage of this effect, a judge would have to trade off control over policy implementation.

Several results of the model are useful for our purposes. First, in order to observe variation in compliance, we should be focusing on insufficiently salient policies. The reason is that for sufficiently salient conflicts, courts are expected to be maximally vague, making it impossible to observe noncompliance, since what is expected is entirely unclear (Staton and Vanberg 2008, 513–14). An alternative interpretation of maximum vagueness in the model is that where noncompliance is perceived to be particularly problematic, judges simply do not find constitutional violations. Either way, the point is that to observe an effect of the Sala’s publicity strategy, we ought to be looking at insufficiently salient cases, precisely where judges are willing to accept some level of noncompliance, as they are primarily focused on dealing with the means-ends problem rather than simply hiding likely noncompliance. Second, for a given level of order clarity (awareness), awareness (order clarity) never decreases compliance (Staton and Vanberg 2008, 510). Third, increasing awareness has competing effects on judicial incentives to be clear, and for that reason may have no influence on order clarity at all. 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. For example, as public awareness increases, courts can leverage it to ask more clearly for policy changes that would have not been feasible at lower levels of awareness. Yet as awareness increases, the same level of compliance can be achieved with a slightly lower level of clarity. The bottom line from the model is that 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.

### The Costa Rican context

The Sala IV’s monitoring system, and the press conference announcing it, should have raised awareness with the Sala’s decisions, and that change ought to have improved compliance outcomes. The plausibility of this claim depends in part on assumptions about the Costa Rican context. First, the types of cases the Sala IV monitored were ones in which its judges were willing to accept occasional noncompliance as a part of addressing the means-ends problem. Relevant actors must have believed that the Sala IV had made it possible to actually track noncompliance in practice. Public officials must have believed that the Sala’s intention to continue reporting on compliance behavior was credible. Public officials also must have believed that there were consequences for noncompliance with judicial orders. We discuss these assumptions.

### Constitutional review in Costa Rica

The latter half of the twentieth century witnessed a massive global increase in

---

5. This is most clearly visible by considering fig. 1B in Staton and Vanberg (2008, 512). As $b$ increases, the region where optimal vagueness is an interior solution expands upward in the space, rendering previously vague decisions more clear; however, at the same time, the region of complete specificity moves out to the right (as $b$ increases, so does $\sqrt{\alpha/b/3}$) so that orders that used to be entirely specific are now more vague. Likewise, on the interior, an increase in $b$ can increase or decrease optimal vagueness.
constitutional review powers (Ginsburg 2003, 2008), especially in states with civil law legal traditions. Civil law states commonly either created separate constitutional courts in the tradition of Kelsen or expanded the constitutional powers of their supreme courts (Navia and Ríos-Figueroa 2005). The Costa Rican approach reflects a combination of these two patterns. A constitutional amendment in 1989 essentially placed a constitutional court within the supreme court, as we have mentioned, adding a fourth chamber. The enabling law eased barriers to access so that anyone in the country, even a foreign national, could petition the court 24 hours a day, without the need for legal representation, and without charge (Wilson 2011). Reformers with ties to the Inter-American Court of Human Rights influenced the early decision making of the newly created Sala IV and encouraged the enforcement of social and economic rights that, while incorporated into the 1949 Constitution, had previously been understood to be nonjusticiable and merely aspirational.

The consequence of the Sala IV’s progressive social and economic rights jurisprudence is that it created a relatively large demand for its services. Our study focuses on the Sala’s amparo jurisdiction, and habeas corpus jurisdictions. The overwhelming majority of cases (98% in our sample) involve the amparo, a constitutional tool found throughout Latin America designed for the swift protection of fundamental rights and freedoms (Brewer-Carías 2009), and through which the Sala’s role in the management of social and economic rights has been most salient. Generally speaking, the complainant in an amparo suit in Costa Rica is an individual seeking a judicial order commanding a public official to take some action (or refrain from taking some action) in fulfillment of a constitutional obligation. Typical orders in amparo involve commands issued to public sector workers over conflicts with relatively low political salience. Consider order no. 13941 of 2011. Here the Sala found that the claimant was constitutionally entitled to swifter medical care. The court writes:

It is ordered that [Name], in her capacity as Director General, and [Name], in his capacity as Chief of Urology, both of the hospital [Name], 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 [Name] receives the transurethral resection of the prostate that he needs, within two months of the date of this communication.

This was not a case of high politics. It was also far from unique. The Sala receives roughly twenty thousand amparo petitions each year, the vast majority of which involve cases like this. The Sala’s amparo jurisdiction thus creates a form of high-volume, low-salience constitutional review, in which judges are in direct and repeated contact with the bureaucracy. In this context, the means-ends challenge, rather than the challenge of hiding likely noncompliance, is the paramount concern.

The compliance monitoring system. The Sala IV’s monitoring system was an initiative of its president, Ana Virginia Calzada. The system collects the data necessary to report meaningfully on patterns of noncompliance. The process 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 which a direct order is granted. The team records the time frame or deadline for compliance (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 noncompliance, the team calls the responsible authority for an explanation. Calls are repeated until the CJC has a specific answer about the case’s status or until five calls are made, after which the case is registered as an instance of noncompliance. 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 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, in instances of noncompliance, to each of the justices for follow-up.

We have already noted that the Sala IV’s press conference was well attended and that various media outlets reported on it. The Sala is a constant source of news in Costa Rica, and there is considerable media freedom.


7. Although this style of judicial review exists in common law settings (e.g., Shankar and Mehta 2009), the amparo is a feature of civil law and mixed systems. This is not to say that the theoretical challenges addressed by the model are only relevant to civil law systems. It does suggest an empirical study focusing on a peak court with constitutional review powers is particularly appropriate in a civil law context.

8. Information about the system can be found at http://sitios.poder-judicial.go.cr/salaconstitucional/seguimiento.htm. Details of the protocol for tracking compliance can be requested from the authors.

9. Freedom House (http://www.freedomhouse.org/) has considered Costa Rica to have a free press in every year it has conducted its rating, since 1980.
controversial that a serious, continuing noncompliance problem might eventually become newsworthy and that, therefore, a public announcement about this system would have increased agencies’ beliefs that court orders were becoming more transparent. We have also conducted in-depth interviews, based on a standard protocol, with legal staff from a variety of public entities, including the Caja Constarricense del Seguro Social, the Municipalidad de Alajuelita, the Dirección General del Servicio Civil, and the Ministerio de Educación Pública. All but one interviewee reporting being aware of the Sala’s monitoring system. The exception was aware that the Sala was actively tracking compliance but did not realize that there was a particular group in charge of the effort.10

The credibility of the Sala’s promise. Relevant parties also must have believed that the Sala’s commitment to monitor and publicize compliance rates in the future was credible. In retrospect, we can judge credibility clearly, as the Sala has indeed followed through.11 But what we need to evaluate is whether the promise would have been perceived to have been credible at the time. An important consideration, one consistent with the models of judicial politics that animate our study, concerns the independence of the court we are considering. A court that can be easily cowed may lose its nerve or otherwise be convinced to take down unfavorable information about powerful officials. The credibility of a promise like the one the Sala IV made is more likely to affect relevant parties in environments in which judicial independence is relatively high. Costa Rica is such an environment (Bumin, Randazzo, and Walker 2009; Wilson and Rodríguez Cordero 2006).

Public officials could have held somewhat more sophisticated expectations regarding the Sala. For example, they could have believed that although the Sala could not have been cowed into hiding information, it could have been embarrassed by what it found as monitoring continued, and in order to avoid public embarrassment, it might have hidden instances of noncompliance. If officials believed this, then we would not see a change in compliance outcomes associated with the press conference, given that officials would not then have expected a change in transparency.

The Sala’s own presentation at the press conference suggests that this more sophisticated belief on the part of officials was unlikely. Figure 1 shows a PowerPoint slide from that press conference.12 Consider the final column, which displays the percentage of cases being tracked in which the CJC could affirm that there had been compliance. The numbers are fairly low in many cases. The actual, overall compliance rate was close to 94% in the period captured in the slide, but at the time of the press conference, tracking had not been completed. This can be seen from the column labeled “Ignorado,” which shows the cases for which information was still unknown. The final column at least suggests that there might have been a serious noncompliance problem, particularly in certain agencies. At a minimum, the Sala would have been perceived to be either extremely risk accepting or not particularly concerned about revealing a noncompliance problem.13

Consequences of noncompliance. For awareness to work as we propose, officials also must have believed that there were consequences for noncompliance. It is important to note that if there were consequences, they were unlikely to come from the formal legal system. Article 53 of the Constitutional Jurisdiction Law creates a legal obligation to comply with amparo orders, and Article 71 provides that noncompliance be punished by a prison term between three months and two years. But it is quite difficult 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.14 As a consequence, prosecutors in the country rarely bring a successful prosecution. If public officials prefer not to be observed defying the Sala IV, it must be as a result of the underlying political and reputational costs that drive public support models of judicial behavior.

We assume that, at a minimum, senior public officials care about their reputations. For political appointees, reputations likely affect the chances of being reelected and/or
influence their ability to enact important policies and programs while in office. For senior civil servants, reputations often affect career prospects and, occasionally, their bargaining power in negotiations over agency budgets and policy outcomes. 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. All of the officials we interviewed reported a concern with compliance. Some entities (e.g., the Caja Costarricense del Seguro Social) have explicit processes designed for ensuring that orders are carried out.

Another way to ask the question is this: Would relevant parties have perceived that the public would attach significance to and be dissatisfied with noncompliant patterns? Costa Rica is characterized by a high commitment to many elements of the rule of law, including the principle that judicial orders ought to be obeyed. On multiple cross-national indicators of the rule of law, Costa Rica receives consistently high marks. While noncompliance in any particular case might be excusable due to circumstances beyond the control of public officials, it is unlikely that the demonstration of repeated and routine noncompliance would not have been publicly acceptable.

Another consideration is that the Costa Rican judiciary is a highly respected institution (Walker 2009). The Sala IV is generally understood as a major force in Costa Rican politics and essentially revolutionized rights protections (Wilson 2011). The Sala’s importance and popularity are evident in the use of the colloquial term for a legal proceeding designed to vindicate a right—“salacuartazo!” Indeed, a simple search on Twitter for “#salacuartazo” reveals multiple uses across many tweets, linked to news articles and blogs about amparo actions. The term “salacuartazo” is even used comically in Costa Rican Spanish, as in “I’m going to file a salacuartazo against mosquitoes for physical and moral damages!” The Sala IV has been used to solve major social rights issues. It also has been used to solve myriad, relatively small, bureaucratic problems that impinge upon rights enjoyment, some of which are perceived to be trivial. For these reasons, it seems plausible to assume that observed noncompliance in Costa Rica is likely to be costly.

### EMPIRICAL ANALYSIS

Our analysis makes use of the CJC monitoring data, which begins in October 2009 and ends in December 2011. Given enough time, the vast majority of amparos are implemented

<table>
<thead>
<tr>
<th>INSTITUCION</th>
<th>Cumplimiento</th>
<th>Cumplimiento parcial</th>
<th>Incumplimiento</th>
<th>Ignorado</th>
<th>Total</th>
<th>% certeza cumplimiento</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ministerio de Educación Pública</td>
<td>24</td>
<td>2</td>
<td>115</td>
<td>141</td>
<td></td>
<td>17%</td>
</tr>
<tr>
<td>Caja Costarricense del Seguro Social</td>
<td>70</td>
<td>2</td>
<td>3</td>
<td>27</td>
<td>104</td>
<td>67%</td>
</tr>
<tr>
<td>Ministerio de Gobernación, Policía y Seguridad Pública</td>
<td>7</td>
<td></td>
<td>6</td>
<td>13</td>
<td></td>
<td>54%</td>
</tr>
<tr>
<td>Ministerio de Justicia</td>
<td>6</td>
<td></td>
<td>7</td>
<td>13</td>
<td></td>
<td>46%</td>
</tr>
<tr>
<td>Poder Judicial</td>
<td>9</td>
<td></td>
<td>1</td>
<td>10</td>
<td></td>
<td>90%</td>
</tr>
<tr>
<td>Ministerio de Obras Públicas y Transportes</td>
<td>1</td>
<td></td>
<td>7</td>
<td>8</td>
<td></td>
<td>13%</td>
</tr>
<tr>
<td>Autoridad Reguladora de los Servicios Públicos</td>
<td>2</td>
<td></td>
<td>4</td>
<td>6</td>
<td></td>
<td>33%</td>
</tr>
<tr>
<td>Ministerio de Ambiente, Energía y Telecomunicaciones</td>
<td>1</td>
<td></td>
<td>5</td>
<td>6</td>
<td></td>
<td>17%</td>
</tr>
<tr>
<td>Ministerio de Trabajo y Seguridad Social</td>
<td>1</td>
<td></td>
<td>1</td>
<td>4</td>
<td>6</td>
<td>17%</td>
</tr>
<tr>
<td>Resto</td>
<td>20</td>
<td></td>
<td>0</td>
<td>3</td>
<td>92</td>
<td>63% 100%</td>
</tr>
</tbody>
</table>

| Total                                    | 141          | 4                    | 7              | 269      | 424   | 33%                     |

16. For a simple example, consider the results described in the BTI Transformation Index (http://www.bti-project.org/country-reports/lac/). For a general review, see Rodriguez, McCubbins, and Weingast (2009).
per the Sala IV’s requirements; however, there is considerable variation in the time it takes to implement an order. Insofar as delayed implementation is a form of noncompliance (Kapiszewski and Taylor 2013; Staton 2010), we focus our analysis on estimating the duration of noncompliance via an event history approach.

The CJC compliance team recorded the month and year of the Sala IV’s vote in each case and the date of compliance. Figure 2 displays a histogram of the number of months that orders spent in the CJC system prior to a registration of certain compliance. The minimum number of months is one (when compliance is documented within a month of the Sala IVs decision), and the maximum is 25. The median duration is seven 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.

We estimate Cox proportional hazards models of compliance, 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:

$$h_i(t) = h_0(t) \exp \left( \alpha \text{ April 2010} + \sum_{k=1}^{K} \beta_k x_{ki} \right),$$

where $h_0$ is the baseline hazard function, and the $x_{ki}$ denote control variables in models that have them. 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 0 otherwise. Orders issued after the press conference should have higher hazards of compliance than those issued before the press conference. Clearly, this historical control approach looks to use information on the compliance process prior to the press conference to learn about its effect. A first issue to consider concerns what we hope to compare. Units in the model defined above are orders, and thus we are comparing orders issued prior to the conference to orders issued after. An alternative approach, which we explore, is to conceive of the agencies that are defendants in these cases as the appropriate units of comparison. Under this approach, we would consider changes in agency behavior before and after the event.

A second, relatively obvious concern in the study involves potential sample selection bias, deriving from three potential sources. We believe it is important to consider carefully the theoretical bases for this concern and what they would mean for our analysis. First, we might be concerned that President Calzada, who controlled when the press conference was held, chose to hold the event on a day selected so as to generate the impression that going public was particularly effective. For this to work, President Calzada would have needed to have known the Sala’s record of compliance and been able to predict that the set of cases that were to be decided after the press conference were likely to be implemented. The evidence she presented at the event, as summarized in figure 1, suggests that some agencies had potentially problematic compliance records, so perhaps all she needed to do was look for case loads in which particular agencies with likely high hazards for compliance were likely to be present. Without some rather elaborate assumptions about the predictive ability of the president, this would have been relatively difficult for dockets deep into 2010 or beyond; however, it might have been possible in the period immediately following the press conference. Still, although possible, we should recall that the overwhelming number of cases summarized in her presentation fell within the “unknown” category, suggesting that she would have been highly uncertain about the extent of compliance with the Sala’s orders when she held the press conference. Further, given the decentralized way in which cases are allocated to different members of the Sala, it is unclear how she would have known that a particular month was likely to be better than another.

Another concern is that the judges of the Sala IV began drafting different types of orders in the wake of the press

---

17. As 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. The CJC was able to verify compliance in 9,391 instances (85%) among the orders that were tracked.

18. The press conference was held on March 2, but since we do not know the exact date of each decision, it was not clear how to code cases in March. We use April, since all votes in April happened after the press conference.
conference, taking advantage of the increase in awareness. We addressed this possibility in the theoretical section. Although there are incentives to increase clarity as awareness increases, there are corresponding incentives to decrease clarity. Thus, a judge who cares about the means-ends problem will not necessarily be more clear when there is a potentially greater degree of pressure for compliance. So, even if the judges of the Sala IV believed that the press conference would have generated greater awareness, it is unclear theoretically that this would have influenced the clarity with which they drafted orders.

A third possibility is that the types of litigants (and their corresponding constitutional questions) changed in response to the press conference. Suppose that complainants only bring claims when, conditional on winning, they believe that agencies are sufficiently likely to provide the ordered relief. As awareness increased, litigants who were skeptical about the probability of compliance, even if they won, perhaps because the relative difficulty of implementation of their request, might have been more likely to file a complaint. If this is true, then the pool of orders observed after the press conference would likely include policy challenges that were "more difficult" to implement. Yet, the consequence of this kind of process is not that a compliance hazard should be higher after the press conference. The prediction is that the hazard of compliance would be weakly lower, reflecting the increase in unobserved difficulty of forcing implementation.

The most straightforward way to address sample selection bias in our study is to directly model the selection process (Boehmke, Morey, and Shannon 2006); however, absent a sensible instrument to be used in an equation estimating an order’s probability of being resolved after the event, this kind of solution can make matters worse (Brandt and Schneider 2007). Our personal intervention is the most likely source of exogenous variation, since it was strongly related to there being a press conference, and extremely unlikely to have influenced amparo compliance except through its effect on the press conference. Unfortunately, it is not clear how to use this information in the selection stage. Absent a modeling fix, we address the concern in three ways. First and foremost, we keep in mind the theoretical arguments for sample selection bias and use those arguments to interpret what we find. Second, we consider a sample of cases that were resolved on dates very close to the date of the press conference (a small window of observation). If the source of bias derives from either litigant or judge behavior, it should be less likely to manifest in this kind of sample. The reasons are twofold: (1) given the short time frame, all actors in the system might have still have been learning about whether they believed that awareness among bureaucrats was really likely to have been influenced, and (2) many of the complaints the Sala was addressing would have been filed prior to the press conference itself. Third, we also consider a sample in which we exclude the decisions resolved in the period near the actual press conference (an "open" window of observation). The reason is that, if the source of the bias is actually President Calzada, by doing so we eliminate from the study the likely window of time over which the president could have possibly had information sufficient to a careful date selection.19

Results

We begin by considering models that do not condition on potential confounding variables.20 Table 1 displays the hazard ratio for the April month threshold for three different samples. The first column shows the estimate for the entire sample of orders, whereas the second column restricts the sample to cases not resolved in the four-month window around the press conference (i.e., from February through May of 2010). The third column restricts the sample to orders issued only during four months surrounding the press conference. The hazard ratios are greater than one for all of these models, though notably, there is a clear reduction in the size of the estimate in the model fit to the small window of orders around the press conference date. The fourth column shows the hazard ratio for the April month threshold in a model that includes fixed effects for the 202 distinct public agency defendants, fit to the entire sample. In that specification the coefficient on the April month threshold reflects the hazard ratio for the press conference date variable averaged across all agencies; that is, it focuses on within agency changes before and after the press conference.21 There is preliminary evidence suggesting increasing hazards of compliance in the period following the Sala’s press conference.

Setting aside the possibility that the press conference itself caused a change in the type of cases the court resolved or in the types of orders judges wrote, we might nevertheless observe differences in the Sala’s case load over time with respect to these issues, and we do observe features of the cases that are likely indicators of the ease of implementing and order, as well as its clarity. Some types of or-

19. This assumes that the window of time surrounding the press conference included the dates over which she was deliberating.
20. All models were fit in Stata 13.1. A replication file is available via the link given in the second unnumbered footnote.
21. Clustering standard errors at the case or agency level has no material effect on the results. The results are displayed in the appendix.
Finally, we have dummy variables indicating 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

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 commonly coded orders. For the commonly coded orders, we randomly chose the codes of one of the two coders. To illustrate, the order described on page 8 was considered clear. It specifies the procedure to be performed (a transurethreal resection of the prostate), as well as a precise location (a particular hospital). By contrast, consider order no. 16059 of 2009, which was not considered clear and definite. It reads:

The last clauses of the sentence introduce elements of vagueness, by asking for “necessary” actions and adjustments, taking “reasonable” account of family and personal needs. What is necessary and what it means to be reasonable are classically vague judicial commands (see Kaplow 1992; Lax 2012). Overall, the coders were able to identify the clarity of 72% of the orders in our data set. 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 3,121 orders that were coded for clarity, 2,556 (82%) entailed actions that were “clear and definite,” and 565 (18%) were not clear.

Figure 3 shows hazard ratios for the April month threshold, as well as the agency, plazo, and clarity control variables. We cluster the standard errors within cases. Note, first, that across all samples, the April 2010 threshold variable has a hazard ratio between 1 and 2, as would be true if

Table 1. Hazard Ratios for the April Month Threshold, with 95% Confidence Intervals

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th>Open Window</th>
<th>Small Window</th>
<th>Agency FE</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1.88</td>
<td>1.98</td>
<td>1.22</td>
<td>2.0</td>
</tr>
<tr>
<td>95% CI</td>
<td>(1.73, 2.04)</td>
<td>(1.81, 2.16)</td>
<td>(1.04, 1.42)</td>
<td>(1.85, 2.16)</td>
</tr>
</tbody>
</table>

Note. The first column is for the full sample of orders. The second restricts the sample to orders not issued between February 2010 and May 2010. The third column restricts the sample to orders only issued between February 2010 and May 2010. The final column includes fixed effects for agencies.
orders voted on after April 1, 2010, have hazard rates that are higher than those voted on prior to April 1, 2010. 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. The baseline \( \text{plazo} \) is an order where the term is immediate. Thus, observing hazard ratios below 1 in these models suggests that every \( \text{plazo} \) longer than an immediate \( \text{plazo} \) is likely to decrease the hazard of compliance. The strongest effect is for \( \text{plazos} \) 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.

The results of these models are consistent with theoretical expectations. The substantive effects are meaningful. Using the models fit to orders issued in 2010 or before (the upper right panel of fig. 3), and setting all covariates to their modal values, median expected survival time is 11 months. It is nine months for orders in cases voted on after the press conference. Consistent with common understandings about the professionalism of the Caja, the median survival time for an order being responded to by the Caja is only six months. Unsurprisingly, the \( \text{plazos} \) themselves have strong substantive effects. Indeed, there is a five-month difference in estimated survival times for a case where the Sala IV has ordered immediate compliance and a case in which the \( \text{plazo} \) is greater than six months.

**Placebo tests.** 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, insofar as there was one, occurred much earlier or much later, and had nothing to do with the press conference. Figure 4 shows estimated hazard ratios for four models. The first column shows models for 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, nor at the end of 2010, well after it.

**Summary**
We have found evidence suggesting that hazard of compliance with orders of the Sala IV likely increased follow-
ing the Sala’s press conference announcing its monitoring system. Specifically, orders issued after the press conference were estimated to be implemented roughly two months earlier than orders issued before the press conference. We also observe increasing hazards of compliance associated with increasing order clarity and decreasing hazards of compliance as time frames for implementation increase or are left totally unspecified.

CONCLUSION

The Costa Rican compliance monitoring experience reminds us that judges do care a great deal about compliance with their orders. It also suggests that judges may be able to influence compliance outcomes by making it easier to track what happens to litigant claims after orders are issued and judges turn to other business. The Sala’s experience raises questions about the typical way in which judges monitor compliance. 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-alarm 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 noncompliance 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. In such cases, a more direct, “police patrol” approach might be useful to consider.

A direct monitoring system 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 noncompliance, 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.

Although our article has considered the impact of the Sala IV’s program with respect to compliance in amparo cases, a natural, broader question deals with the problem of providing social and economic rights through a bureaucratic framework. Latin American laws and constitutions typically provide access to goods and services, as a matter of right, through the bureaucracy. Litigants commonly seek amparo protection or assistance when bureaucracies are slow or broken. An amparo order may bump a person up in a bureaucratic queue. What does this mean for the remaining members of the queue? A careful model of queuing and ordering is essential to identify precise expectations (e.g., Lui 1985), but it is possible to sketch out the contours of the problem.

We might begin by considering the underlying cause in delays in the bureaucratic process. One possibility is that the bureaucracy is simply not working hard enough—that there is simply slack in the system. An amparo system that re-shuffles the order of service is unlikely to influence overall work rate, and so while the successful litigant is moved to the front of the line, social welfare is unlikely to be affected. An increase in awareness, one that raises the costs of defying judicial orders, might have an effect, but much would depend on the way the bureaucracy reacts to this stimulus. For example, agents might increase their work rate with respect to successful litigants but decrease the rate at which they work on cases not litigated, again resulting in no welfare gains. Yet again, this supposes that amparo relief is sought more or less by any litigant who perceives there to be a problem. If instead, only a small group of people use the amparo route, then there may be important distributional consequences associated with solving a bureaucratic failure through the constitutional review system (Brinks and Gauri 2014). Another possibility is that the time an individual spends in the queue is systematically related to some personal or group feature. Constitutionalizing the process might at least restore some equity to the queue, but again much would depend on the process by which litigants come to court. Finally, we might imagine that delays result from some kind of budgetary (or other political) constraint. A model of this process is well beyond the scope of our study, but we believe that in order to answer questions about the aggregate effects of the Latin American amparo system and a judicial-led program of promoting more rapid compliance, it is essential that such a model or set of models be developed.

ACKNOWLEDGMENTS
For comments on our article, we would like to thank Tom Clark, Adam Glynn, Lisa Hilbink, Christoph Hönnigle, Mariana Llanos, Fernando Montenegro, Julio Rios Figueroa, Peter Rosendorff, Rachel Sieder, Adam Wagstaff, Fabrizio Zircon, as well as participants at the conference “Courts under Pressure: Formal and Informal Dynamics of Political Intervention in New Democracies,” Hamburg, Germany. We would like to thank Nancy Marin Espinoza for research support. We thank Sonia Villegas and Ana Virginia Calzada for support at the Sala Constitucional, Corte Suprema de Justicia de Costa Rica. The views and findings in this article do not necessarily represent those of the World Bank or its Executive Directors.

REFERENCES


