

The Impact of Presidential Appointment of Judges: Montesquieu or the Federalists?[†]

By SULTAN MEHMOOD*

A central question in development economics is whether there are adequate checks and balances on the executive. This paper provides causal evidence of how increasing constraints on the executive—via removal of presidential discretion in judicial appointments—promotes the rule of law. The age structure of judges at the time of the reform and the mandatory retirement age law provide us with an exogenous source of variation in the termination of presidential discretion in judicial appointments. Overall, the results indicate that presidential appointment of judges deteriorates the rule of law. Even one degree of separation between the judiciary and the president matters. (JEL D72, K40, O17)

There is no liberty if the power of judging is not separated from the legislative and executive power.

—Montesquieu, *L'Esprit des Loix* (1748)

A judiciary's job is to interpret the law, not to challenge the administration.

—Muhammad Zia-ul-Haq, *Amnesty International Report* (1982)

In 70 percent of countries worldwide, it is the president who appoints judges to the courts (CIA World Factbook 2021). This makes it one of the most widespread institutions in the world. Proponents of presidential appointment argue that

*New Economic School, Moscow, Russian Federation (email: smehmood@nes.ru). Benjamin Olken was coeditor for this article. I would like to thank Ekaterina Zhuravskaya, Eric Brousseau, Daron Acemoglu, Esther Duflo, Abhijit Banerjee, Andrei Shleifer, Daniel Chen, Ruben Enikolopov, James Robinson, Henrik Sigstad, Ruixue Jia, Saad Gulzar, Yasir Khan, Jean-Philippe Platteau, Thierry Verdier, Sergei Guriev, William Howell, Georg Vanberg, Anandi Mani, Adam Szeidl, Monika Nalepa, Nico Voigtlaender, Christian Dippel, Paola Giuliano, Thomas Piketty, Eric Maskin, Prashant Bharadwaj, Dany Bahar, Claire Lim, William Hubbard, Dina Pomeranz, Marta Troya-Martinez, Federica Carugati, Lisa Bernstein, Monica Martinez-Bravo, Ikram Ul Haq, Abbas Askar, Sajwaar Khalid, Ali Shah, Osama Siddique, Roberto Galbati, Thomas Fujiwara, and participants at the ALEA 2019, MPSA 2019, SIOE 2019, WIP 2019, APSA 2019, RES 2019, NEUDC 2020, Econometric Society Winter Meeting 2020, CEPR STEG 2021 Annual Conference, AMIE 2021 Workshop, New Economic School, Paris School of Economics, Lund University, Stockholm School of Economics, University of Oxford, and University of Chicago for their comments and feedback. Financial support from the Chair Governance and Regulation Lab, DIAL, Paris Dauphine University is acknowledged. This work was also supported by French National Research Agency Grant ANR-17-EUR-0010. I also thank the anonymous referees for their thoughtful comments. Bakhtawar Ali provided excellent research assistance.

[†]Go to <https://doi.org/10.1257/app.20210176> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

this institution secures the independence of the judiciary, especially when combined with the security of tenure for judges (Madison, Hamilton, and Jay [1788] 1961). However, others, like Montesquieu (1778), worry that presidential appointment may result in judicial subversion by the government's executive branch, undermining the separation of powers and the rule of law.¹

This paper contributes to this debate by investigating the effect of removing presidential discretion in judicial appointments on progovernment rulings and decision quality. In 2010, a sudden and unique institutional reform in Pakistan dramatically changed the judicial selection system from presidential appointment with lifetime tenure (like Singapore or Brazil) to judicial, commission-based selection, which involves appointment by peer judges (like Sweden or the United Kingdom). This natural experiment provides an opportunity to understand how the removal of presidential discretion in the appointment of judges impacts judicial decision-making. We ask whether this switch from presidential appointment to appointment by judge peers affected judicial independence and, if so, which mechanisms link presidential appointment of judges to judicial decision-making?

To systematically examine the influence of this reform on judicial decision-making, we randomly sample the universe of cases in Pakistan's district high courts and obtain information on about 8,500 cases from 1986 to 2019. Our measure of executive influence over the judiciary is a judicial dependence dummy variable "state wins," taking a value of one for state victories and zero for state losses in cases where the state is a party. Following the literature, we asked legal experts at a law firm to code this variable as in Djankov et al. (2003) and La Porta, López-de-Silanes, and Shleifer (2008). Coding was performed by two independent teams of legal experts, each supervised by a senior lawyer specializing in cases involving the government.

The Pakistani government is a party in a wide range of judicial cases, from tax disputes to blasphemy, suppression of political rights, and the constitutionality of military rule. However, the most common cases concern land expropriation by the government, which total 40 percent of all petitions filed in the high courts.² Indeed, in many developing countries, property is vulnerable to outright seizure by the government (Shleifer and Vishny 2002; Somin 2015; Behrer et al. 2021). According to the World Justice Project, about 60 percent of respondents in developing countries say that it is "unlikely" or "very unlikely" that homeowners will "be fairly compensated by the government" if the "government decides to expropriate their property."

When the government expropriates land, courts are generally the only recourse for citizens seeking to recover their property (La Porta, López-de-Silanes, and Shleifer 2008). On November 29, 2017, a Pakistani court presided over by judges who were peer appointees ordered the Karachi Development Authority to return 35,000 "public encroachments" to their owners (*The News* 2017). Likewise, in

¹This debate has also attracted considerable attention in economics. North (1990) defines institutions as "humanly devised *constraints* that shape human interaction" and Acemoglu, Johnson, and Robinson (2001) cite this definition in their influential study to motivate the use of "constraints on the executive" as their central measure of institutions. (See also discussion of the importance of checks and balances as key for development in La Porta et al. 2004; Rodrik, Subramanian, and Trebbi 2004; and Acemoglu et al. 2020.)

²By "government," we mean all levels of the administration with executive authority (i.e., local, provincial, and federal government and public agencies, e.g., the various land development authorities in Pakistan).

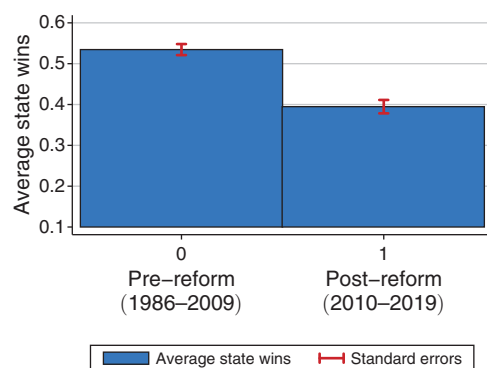
another judgment where all three judges were peer appointed, the Islamabad High Court's bench unanimously ruled that the foreign minister be removed from office for having "deliberately and willfully not disclosed his status as an employee of the foreign company, nor receiving of the salary per month." This is in stark contrast to presidential appointees' rulings involving individuals holding executive office.

Figure 1 generalizes the anecdotal accounts of less favorable rulings for the state following the 2010 reform to about 8,500 cases. Prior to the selection reform, around 50 percent of cases were decided in favor of the state, as opposed to about 40 percent thereafter (panel A). These differences are both qualitatively and statistically significant. A similar pattern emerges when we consider variation over time: there is a sharp fall in state wins precisely after presidential discretion in judicial appointments is removed (panel B). This, however, cannot be interpreted as conclusive evidence for a causal link between the change in judicial selection procedure and judicial outcomes, as a number of other changes occurred around the selection reform year. For instance, the transition from military to democratic rule took place in 2008. Likewise, a social movement by Pakistani lawyers in 2007 demanded President Musharraf's resignation, and in 2010, the president's power to unilaterally terminate the legislature was also removed from the constitution. The overall fall in progovernment rulings following the selection reform could be explained by any of these changes. Similarly, a simple difference-in-difference estimate of the fraction of judges appointed by the judicial commission in each district court may not yield the causal effect of the reform due to the potential reassignment of judges across districts or the strategic appointment of judges across districts. Indeed, Iyer and Mani (2012) show that the reassignment power of Indian politicians allows them to exert substantial control over bureaucrats. There is a legitimate concern here that an independent judge in Pakistan might be reassigned to a different district or that judicial-commission appointments might be made strategically.

We address these concerns by focusing on plausibly random cross-sectional variation in the implementation of the reform due to the age composition of near-retirees in the reform year. Both before and after the reform, Pakistani law made retirement at age 62 mandatory for judges.³ However, different high court benches had different numbers of vacancies arising from mandatory retirements in the 2010 reform year. For instance, the property bench of the Rawalpindi district had no vacancy arising from mandatory retirements in 2010, while the property bench of the Peshawar district did. Simply put, our design compares progovernment rulings at district benches where judges turn 62 pre-reform and are replaced by the president (control group) with district benches where judges turn 62 post reform and are replaced by judge peer appointees (treatment group). More precisely, our reduced-reform difference-in-difference framework compares pre- and postreform government victories at district benches experiencing low versus high mandatory retirements in 2010 as determined by the age structure of judges in 2010. The district benches

³The retirement of high court judges at age 62 is also common in other large developing countries—for instance, India. This is likely due to a combination of lower life expectancy and British colonial legacy (Siddique 2013). Taking just the case of Pakistan and India, where high court judges must retire at age 62, implies that this exact institution affects at least 20 percent of the world population, or about 1.5 billion people worldwide.

Panel A. Average state wins before and after the reform



Panel B. Average state wins over time

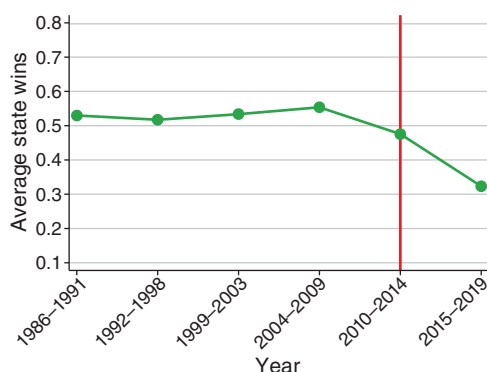


FIGURE 1. STATE WINS PRE- AND POST-REFORM

Notes: The figure compares state wins pre- and post reform. In panel A the bar chart shows average state wins in pre-reform and post-reform periods with 95 percent confidence intervals. In panel B we plot average state wins over time. The vertical line in panel B represents the 2010 reform.

with a lower fraction of mandatory retirements in 2010 serve as a counterfactual to district benches with a higher fraction of mandatory retirements in 2010.

We find that presidential appointment of judges substantially affects judicial decisions: a 10 percent rise in judges selected by the judge peers reduces state wins by about 2 percentage points. This is equivalent to a 4 percent reduction over the sample mean. We present evidence, consistent with qualitative accounts, that this reduction in state wins reflects improvement in decision quality: peer appointees have lower case delays, are more likely to rule on case facts relative to legal lacunae, and have higher ratings on due process followed than presidential appointees. The reduced progovernment rulings do not come at the expense of increased pandering to religious sentiments on the street or judicial subversion by large corporations. If anything, judges appointed by their peers rule in favor of the government more often when the state is opposing large corporations, which suggests that reduced progovernment rulings are also unlikely to result from corporate capture.

Two key threats to identification could still hinder causal interpretation of the selection reform's impact on judicial decision-making. First, we might be picking up a pure appointment effect. If, for instance, new appointments have an independent effect on judges' behavior, then we might be picking up the effect of new judicial appointments instead of the change in the judicial selection procedure. A falsification test, however, strongly suggests this is unlikely: we show that mandatory retirements before the selection reform—i.e., in 2007, 2008, or 2009—have no influence on government victories. Mandatory retirements in these years were also determined by judges' age structure and led to new judicial appointments, but by the president. In these instances new presidential appointments are not predictive of government victories. Second, we may be confounding the effect of the selection reform with differential prior trends among benches with high versus low mandatory retirements. However, we find no

evidence of the district benches with more vacancies being on different trajectories before the reform. This is consistent with qualitative accounts indicating that the selection reform was unanticipated, framed by a secret parliamentary committee and implemented in a hurry to prevent potential sabotage (Almeida 2018). Similarly, no effect of pretreatment mandatory retirements and a lack of evidence for differential pre-trends are found for other judicial outcomes proxying for decision quality.

We next explore the likely mechanism of the type of judges driving the results. In particular, we ask: What are peer appointments selecting for? We find that president-appointed and peer-appointed judges are similar in many ways, such as age at appointment, ethnicity, religion, and gender. That is, peer appointments are not selected based on seniority or ethnicity. Nevertheless, judicial commission appointees are about 35 percent less likely to have run for political office prior to their judicial appointments. This corroborates anecdotal accounts that the judicial commission chooses more meritocratic and bipartisan judges relative to presidential appointments (Zafar 2012). We do not find much evidence for spillover effects: peer-appointed judges do not impact the behavior of president-appointed judges, suggesting that incentives facing old-regime presidential appointees remained unchanged post reform.

Finally, we examine the heterogeneity by case type and find that politically salient constitutional cases involving land expropriations and political rights disputes with the state are essentially driving our results. These are cases where government expropriation of valuable tangible resources (such as land) and intangible resources (such as political rights) are at stake. The expropriation of private property by the government opens the opportunity to conduct a back-of-the-envelope calculation to approximate the value of total land expropriations avoided due to the selection reform. Based on judgment-order valuations of expropriated property and our point estimates, we compute that the selection reform likely prevents land expropriations worth 0.14 percent of GDP, or \$390 million, yearly. In other words, our computations suggest government would continue to expropriate an additional \$390 million of land every year if all judges were still presidential appointees. To put this amount into perspective, it is roughly equivalent to the entire federal government's expenditure on health care in 2019.

We test for and reject alternative explanations for the finding that judicial selection reform changed judicial decision-making in Pakistan. We show that the effect of selection reform is not a president- or chief-justice-specific effect. The results hold regardless of the president or chief justice in office. We also find no evidence for a strategic case filing mechanism whereby the government responded to the selection reform by strategically allocating cases, or that the chief justice differentially constituted judicial benches post reform. This is consistent with *de jure* random allocation of cases, substantial sanctions for "forum shopping" (i.e., litigants choosing a specific judge), and the unanticipated implementation of the reform. The results of balance tests and lack of evidence of change in the number of cases filed also corroborate this view. We conduct a number of additional sensitivity tests showing that the results are not confined to a particular district or a specific bench. The results are also robust to alternative coding of judicial outcomes, nonlinear estimations, or different levels of aggregation and clustering.

This paper relates to several strands of literature. First, it speaks to the literature on institutions and development, particularly the studies emphasizing the importance of checks and balances on executive power (North 1990; Acemoglu, Johnson, and Robinson 2001; Shleifer and Vishney 2002; La Porta et al. 2004; Rodrik, Subramanian, and Trebbi 2004; Acemoglu, Robinson, and Torvik 2013). We contribute to this literature by showing how executive control over the judiciary and expropriation of private property by the government sharply reduces when one of the world's most ubiquitous institutions, presidential appointment of judges, is discontinued. Hence, we provide empirical support for the theory and mechanisms behind many of these seminal studies.

Second, the paper speaks to the vibrant literature on state effectiveness and bureaucracies in developing countries, especially in fragile or weakly institutionalized settings (Besley and Persson 2009; Fujiwara and Wantchekon 2013; Jia, Kudamatsu, and Seim 2015; Colonelli, Prem, and Teso 2020; Acemoglu et al. 2020; Callen et al. 2020; Gulzar and Khan 2021).⁴ Our contribution lies in demonstrating how appointment by judge peers can increase judicial independence even in a weakly institutionalized setting. Very few studies have investigated the judiciary in developing countries; our work provides insights into how judicial independence can be fostered in a country where democratic institutions are weak to begin with. Third, we contribute to the extensive cross-country literature on courts (Djankov et al. 2003; La Porta et al. 2004; Voigt 2008; Palumbo et al. 2013; Boehm 2015; Bielen et al. 2018; Chemin 2020). By drawing on variation across judicial benches subject to the same national institutions, we overcome many of the common identification issues arising in work exploring differences between countries. We therefore complement the pioneering studies of courts at the subnational level (Ponticelli and Alencar 2016; Lambais and Sigstad 2021; Behrer et al. 2021). However, by comparing judicial cases across judicial benches within a district—via district-year fixed effects—we are able to account for both time-invariant and time-varying determinants of judicial decision-making, something hitherto not possible.

Last, our work is related to the literature on judges' behavior. Most of this literature has focused on judge behavior in criminal cases (Lim 2013; Chalfin and McCrary 2017; Silveira 2017; Cohen and Yang 2019; Ash et al. 2021), racial bias in criminal sentencing (Alesina and La Ferrara 2014; Rehavi and Starr 2014), and extraneous factors affecting sentencing such as lunch breaks (Danziger, Levav, and Avnaim-Pesso 2011), terrorism (Shayo and Zussman 2011), temperature (Heyes and Saberian 2019) and religion (Mehmood, Seror, and Chen 2022). We here reveal a political selection mechanism: judge behavior in politically salient cases is affected by the way in which judges are selected.

The remainder of the paper is organized as follows. Section I provides the background and institutional details. Section II describes the data, while Section III outlines the empirical methodology. Section IV presents and discusses the main results, and Section V explores the mechanisms behind them. Section VI presents a heterogeneity analysis of cases, while in Section VII, we rule out alternative explanations

⁴Other related works include Jones and Olken (2005); Lim (2013); Hessami (2018); and Ash and MacLeod (2019).

and detail a battery of robustness checks. Section VIII concludes. Further information on data construction, variable descriptions, and additional robustness checks is in the online Appendices.

I. Background

A. Contextual Details

Judicial Structure.—The judiciary in Pakistan has a three-tier hierarchical structure. At the lowest level are the civil and session courts hearing civil and criminal cases, respectively, whose rulings can be challenged in the high courts. In these high courts, an individual can file a case against the government in the form of a constitutional petition against the state. Cases with the state as respondent involve the federal government, provincial governments, local governments, government agencies, or any organ of the state with executive authority (such as the office of the president or the prime minister). Last is the final appellate court, the Supreme Court of Pakistan, located in the federal capital, which hears appeals from the high courts. The Supreme Court can have at most 16 judges, which greatly limits the number and scope of its cases; only a small fraction of cases end up being heard by the Supreme Court (Siddique 2013).

High Courts and Specialized Benches.—The focus in this study is on the 16 high court divisional benches that adjudicate cases involving the government. If the government expropriates land or violates a political right, the high court is the first—and in most cases, the only—platform offering remediation to individuals and firms. There are four provincial high courts and one federal high court, in Islamabad. Each high court has roughly 4 “divisional” benches, totaling 16 district or divisional high courts in Pakistan. Online Appendix C, Figure C1 shows their locations and respective jurisdictions. Judges within these 16 divisional courts are allocated according to a roster of sittings that designates the judge’s specialty within the divisional court. In cases involving the state, the focus of our study, the judges serve on one of four specialized judicial benches within each divisional court: a property bench, ruling on land or property disputes with the government; a tax bench, for tax disputes; a writ bench, for human rights petitions; and a criminal bench, for criminal cases. Cases are randomly assigned within the specialized bench according to caseload and judge specialty using computerized case flow management software. When a vacancy on a specific district bench arises, it is always filled by a judge with the same expertise or specialty. The specialty is determined a priori through consultation between the provincial chief justice and the judge in question and is officially announced via the “roster of sittings” (Kureshi 2020). From 1986 to 2019, about 70 percent of all cases filed in the high courts were “constitutional petitions,” the majority of which involved the government responding to land expropriation claims from the citizenry.

Judicial Selection Reform.—In April 2010, the ruling Pakistan People’s Party tabled a constitutional amendment before the Pakistani Parliament that would dramatically change the process of judicial appointment in Pakistan. The Eighteenth

Amendment to Pakistan's constitution was passed by parliament on April 15, 2010, and signed into law by the president on April 19, 2010, when it came into effect (Tavernise and Masood 2010). It removed the following clause from the constitution: "The Chief Justice and each of other Judges of a High Court shall be appointed by the President in accordance with Article 175A." This was replaced by: "There shall be a Judicial Commission of Pakistan, for appointment of Judges of the Supreme Court, High Courts and the Federal Shariat Court. The Commission by majority of its total-membership shall nominate for each vacancy of a Judge in the Supreme Court, a High Court or the Federal Shariat Court, as the case may be" (Constitution of Pakistan 2010).⁵ Article 195, instituting security of tenure for high court judges via mandatory retirement at age 62, and Article 209, stipulating that judges can only be removed by filing a reference to their peers, were the same as before the reform (Constitution of Pakistan 2010). That is, the constitutional amendment only changed the procedure to select judges; the judges' security of tenure, the mandatory retirement law, and the powers resting with the judges remained unchanged.

Peer Appointment System.—The peer appointment system selects judges through majority voting. The judicial commission considers senior lawyers' and lower court judges' candidacy for high or Supreme Court judgeships based on their "reputation for impeccable integrity" (Mushir Alam, interview with the author, 2021). When a vacancy for a judge arises, the Chief Justice of Pakistan, who heads the judicial commission, convenes the commission to deliberate on potential candidates for elevation to the bench. Supreme Court judges are selected by a commission consisting of the "Chief Justice of the Supreme Court and four senior-most judges, a former judge (nominated by the Chief Justice of Pakistan), federal law minister, and the attorney general of Pakistan, along with a senior advocate of the Supreme Court nominated by the Pakistan Bar Council for two years" (Constitution of Pakistan 2010). For the appointment of high court judges, the focus of our study, additional members from the court where the judge is to be appointed are included; this includes the province's chief justice, the province's law minister, the most senior judge of the high court (after the province's chief justice), and a lawyer nominated by the province's bar council. In addition to the formal procedure laid out in the constitution, there are some informal norms that govern the operation of judicial commission. A key norm that has been respected since the inception of the peer appointment system is that for the selection of high court judges, the chief justice of the state where vacancy arises recommends a list of potential candidates for the commission to vote on. This is based on the idea that the chief justice of the state is more likely to be privy to the institutional and contextual needs of the district high court. The judicial commission can, in principle, reject the provincial chief justice's list of candidates, although it is an informal norm to at least allow the vote for the suggested candidates, considering the stature of the provincial chief justice (Mushir Alam, interview with the author, 2021).

⁵The Eighteenth Amendment also aimed to increase the provinces' autonomy and weaken the overall power of the president. For instance, it also took away the president's power to unilaterally dismiss parliament.

The peer appointment system became operational in April 2010, when the Supreme Court and high court judges began to be appointed by the judicial commission consisting of peer judges and senior lawyers with no presidential involvement. The appointment power of the executive was likely curtailed by this reform, as judges constitute the overwhelming majority of the commission. The Eighteenth Amendment also created a parliamentary committee consisting of four members from the government and four from the opposition. Nominations by the judicial commission have to be debated within the parliamentary committee, although it was never able to exercise this authority, since the judicial commission could veto the parliamentary committee's objections without providing any explanation.⁶

B. Political Economy of the Selection Reform

Why would politicians willingly implement a judicial reform that entailed a loss of political power over judicial appointments? According to many political observers, the judicial selection reform introduced after a decade of military rule was intended to reduce the political power of the military. Pakistan's military leaders had long ruled as presidents and used the judiciary to obtain constitutional indemnity for their military coups. It was hoped that independent judges would uphold constitutional clauses barring military takeovers, thus shielding the country against extraconstitutional military takeovers (Kureshi 2020).⁷ Pakistani politicians therefore hoped to take shelter behind constitutional protections from military takeovers by ensuring that judges were more independent, even though this reduced their own effective control over the judiciary. Moreover, it was hoped that reducing presidential discretion over the selection of judges would prevent abuses of power by future autocratic rulers who might subvert the courts to imprison opposition politicians and violate fundamental rights (Sattar 2012; Zafar 2012).

The possibility of sabotage by Pakistan's politically powerful military also led to the reform being conceived and debated in complete secrecy. As one commentator observed, "It was debated and created in total secrecy by a small parliamentary committee" (Almeida 2018). This covertness meant that the selection reform came as a sudden and unanticipated shock to the judicial system of Pakistan. Further discussion of the reform, the political context, and the history of the courts in Pakistan can be found in online Appendix B.

II. Data

Our empirical analysis uses data on judicial cases from the central repository of cases in Pakistan used by lawyers to prepare their cases. We randomly sampled

⁶This was not in the original Eighteenth Amendment but was incorporated as the Nineteenth Amendment, which increased the number of judges in the judicial commission (judges now have the supermajority of eight out of eleven positions in the judicial commission as opposed to six out of nine, under the Eighteenth Amendment) and stated that the judicial commission now also had the power to overrule parliamentary committees' objections to appointments (Constitution of Pakistan 2010).

⁷Article 6 of the Constitution of Pakistan states that military takeover is "abrogating the constitution" and "high treason punishable by death." Yet, courts have never been able to enforce this clause in the constitution or convict military rulers.

8,500 cases—conditional on the state being one of the parties—from 1986 to 2019 for 64 high court benches (from the universe of all cases involving the government decided in this period). The case-level data was successfully matched with judge and district bench characteristics from judicial administrative data for 8,446 cases out of the 8,500. The random sample is conditional on the state being one of the defendants. Therefore, our sample only contains cases involving the state versus a citizen. This includes criminal cases and more politically salient constitutional cases. Based on judgment texts coded by two independent teams of legal experts, government victories were ascertained at the case level. Cases may be decided by a group of judges, i.e., a “double bench” (consisting of two judges), a “full bench” or “divisional bench” consisting of all judges within a division (about seven judges), or by a “single-member bench,” where one judge adjudicates alone. About 90 percent of the cases in the high courts are decided by a single-member bench, making case- and judge-level comparisons roughly equivalent.⁸ The government wins the case about 50 percent of the time, and there are about 7 judges in each of the 16 divisions shown in Figure C1. Table 1 shows the descriptive statistics of the variables used in the study. The key outcome and explanatory variables are detailed below. Further information on the variables, their sources, sampling, and data construction can be found in online Appendices A and B.

Outcome Variables.—The key outcome variable is state wins. This is a case-level measure of judicial independence constructed from the text of the judgment orders containing details of the case. Following Djankov et al. (2003) and La Porta, López-de-Silanes, and Shleifer (2008), we asked a law firm to code this variable. We entered into contract with two independent teams at a prominent law firm in Lahore, with each team consisting of five legal experts (four junior lawyers and a supervising senior attorney) coding the “state wins” dummy variable as one if the state, as a litigant, won a dispute and zero otherwise.

The state here includes all organs of the state yielding executive power, such as local, provincial, and federal governments, the Office of the Prime Minister, the Office of the President, and governmental agencies (in line with the conceptualizations of the state as an executive organ in Montesquieu [1748]). Figure 1 provides the evolution of state wins over time, while Figure C2 provides district-wide averages. In both instances, we observe a sharp fall in state wins in the reform year of 2010. For the analysis of the quality of judicial decisions, we use four additional outcome variables: case delay, merit, correct decisions, and process followed, where the unit of observation is also at the case level. These variables, too, are constructed from the information in the text of the judgment orders. “Case delay” is calculated as the difference between the case decision and filing years. This variable is most straightforward to code since it only requires reading filing and decision years off the judgment text. “Merit” is a dummy variable that takes a value of one if the decision is “based on evidence or case merits” and zero if it is based on a technicality. There are two reasons for constructing this variable. First, legal scholarship

⁸Our empirical method does not hinge on one-to-one case-to-judge matching, since we examine the impact of selection reform by comparing relative probabilities that a case was adjudicated by a peer appointee.

TABLE 1—DESCRIPTIVE STATISTICS

Variables	Observations	Mean	SD	Min	Max
<i>Panel A. Case characteristics (by case)</i>					
State wins	8,446	0.482	0.500	0	1
Case delay (years)	8,446	3.354	2.238	0	29
Merit	8,446	0.627	0.484	0	1
Process followed	8,446	3.314	1.496	1	5
Constitutional cases	8,446	0.722	0.448	0	1
Land cases	8,446	0.409	0.492	0	1
Human rights cases	8,446	0.314	0.464	0	1
Criminal cases	8,446	0.280	0.449	0	1
Pages of judgment order	8,446	8.878	7.706	1	81
Number of lawyers	8,446	4.124	1.807	2	32
Number of judges on a case	8,446	1.809	0.839	1	5
Chief justice on bench	8,446	0.065	0.246	0	1
<i>Panel B. Appointment and retirement in the reform year (by district bench)</i>					
Retirements in 2010/total	64	0.115	0.199	0	1
Appointments in 2010/total	64	0.080	0.168	0	1
<i>Panel C. Judge characteristics (by judges)</i>					
Tenure at decision	511	3.331	2.236	0	22
Gender	511	0.961	0.194	0	1
Promoted to SC	511	0.061	0.239	0	1
Former judge	511	0.115	0.320	0	1
Former officeholder, bar association	511	0.569	0.496	0	1
Ran for political office	511	0.186	0.389	0	1
Former lawyer	511	0.886	0.318	0	1
Post-reform judge	511	0.325	0.469	0	1
Punjabi ethnicity	511	0.196	0.397	0	1
Sindhi ethnicity	511	0.053	0.224	0	1
Balochi ethnicity	511	0.063	0.243	0	1
Pashtun ethnicity	511	0.143	0.350	0	1
Other ethnicities	511	0.483	0.500	0	1
<i>Panel D. Treatment variables and special bench characteristics (by district bench year)</i>					
Retirements in 2010 × post 2010	1,516	0.042	0.134	0	1
Appointments in 2010 × post 2010	1,516	0.031	0.111	0	1
Number of judges on bench	1,516	7.635	3.237	2	16
Number of criminal cases at bench	1,516	1.566	4.716	0	42
Number of land cases at bench	1,516	2.297	2.781	0	24
Number of human right cases at bench	1,516	1.749	2.229	0	16

Notes: This table reports the summary statistics for the baseline sample of 8,446 cases, with 511 judges covering the 64 district benches over the period 1986–2019.

in Pakistan suggests that ruling on technicalities is a “weapon of choice to rule unfairly” and that judges use decisions on technicalities or legal lacunas to “favor the state authorities” (Aziz 2001) and such rulings are “symptomatic of a biased decision” (Javed Arshad, interview with the author, 2017). Therefore, we proxy the “correctness” or unbiased nature of a judicial decision by this dummy variable. Second, this variable is consistent with common law jurisprudence, which aspires toward rulings on the merits, i.e., based on evidence and the spirit of the law rather than legal technicalities, an ideal Pound (1998) and Tidmarsh (2009) discuss in detail. We also cross-check this result by actually asking legal experts

to categorize judicial decisions as “correct” versus “incorrect.” “Process followed” is a discrete variable representing a quality rating on due process followed for each judicial case. Specifically, the legal experts were asked to rate the quality of the judgment order on a scale of one to five, taking into account the extent to which “all relevant jurisdictional, procedural, and evidential requirements” were accounted for. This variable approximates the “correctness” of the due process followed in reaching the judicial decision.⁹

Main Explanatory Variables.—A key explanatory variable used in the analysis, $\frac{\text{MandatoryRetirements}_{\text{in2010}}}{\text{TotalJudges}} \times \text{Post2010}$, is the fraction of judges reaching their mandatory retirement age in 2010 on a given district bench interacted with a post-reform dummy. Data on retirements, total judges, and other judge characteristics come from judicial administrative records obtained from the Registrar Offices of the High Courts. These two sources are also used to construct $\frac{\text{Appointments in 2010}}{\text{Total Judges}} \times \text{Post2010}$. This is the fraction of new judicial commission appointments on each district bench for 2010, the reform year, interacted with a post-reform dummy. In 2010, 11.5 percent of judges reached their mandatory retirement age and 8 percent of these vacancies were filled by peer-appointed judges (as can be seen in Table 1, panel B). Since there is a very strong correlation between mandatory retirements and new judicial appointments (both before and after the reform), we report as a baseline the results with mandatory retirements. Moreover, when a vacancy on a specific district bench arises, it is always filled by a judge with the same expertise or specialty—a pattern reflected in Figure 2, where we observe that mandatory retirements and new judicial commission appointments are almost perfectly correlated within district benches.¹⁰ This implies that new appointments instrumented with mandatory retirements will give us the local average treatment effect for those district benches where a judicial commission appointment and a mandatory retirement occurred in the same year.¹¹ Although, given the high correlation, the reduced-form or *Intention-to-Treat* estimate is, by construction, close to the 2SLS estimate.

Controls: Case, Judge, and District Bench Characteristics.—We rely on a combination of judgment texts and judicial administrative data to construct the case, judge, and district bench characteristics that we use as control variables. The case characteristics data, like the outcome variables, are obtained from the judgment

⁹Two independent teams coded each of these outcome variables, and the correlation coefficient between them appears in Table C1 of online Appendix C. Data from Team 1 is used in the paper, but the very high correlations between the two codings implies that using either dataset would give similar results. This checks out in Table C20 of online Appendix C, where we find the results to be essentially identical if we instead used Team 2’s coding.

¹⁰Ninety-one percent of judges in Pakistan serve out their full terms and only retire on their sixty-second birthdays, while the remaining 9 percent either die in office (6 percent) or are promoted to the Supreme Court (3 percent). Moreover, the fraction of retirements is always larger than the fraction of judicial commission appointments. This is because not all mandatory retirements from a judicial bench are accompanied by new appointments in the same year, although the correlation is very strong (0.9).

¹¹The compliers here would be those district benches where there are no strategic appointments or no transfers of judges across district benches and where vacancies arising from mandatory retirements were filled with judicial commission appointees in the same year.

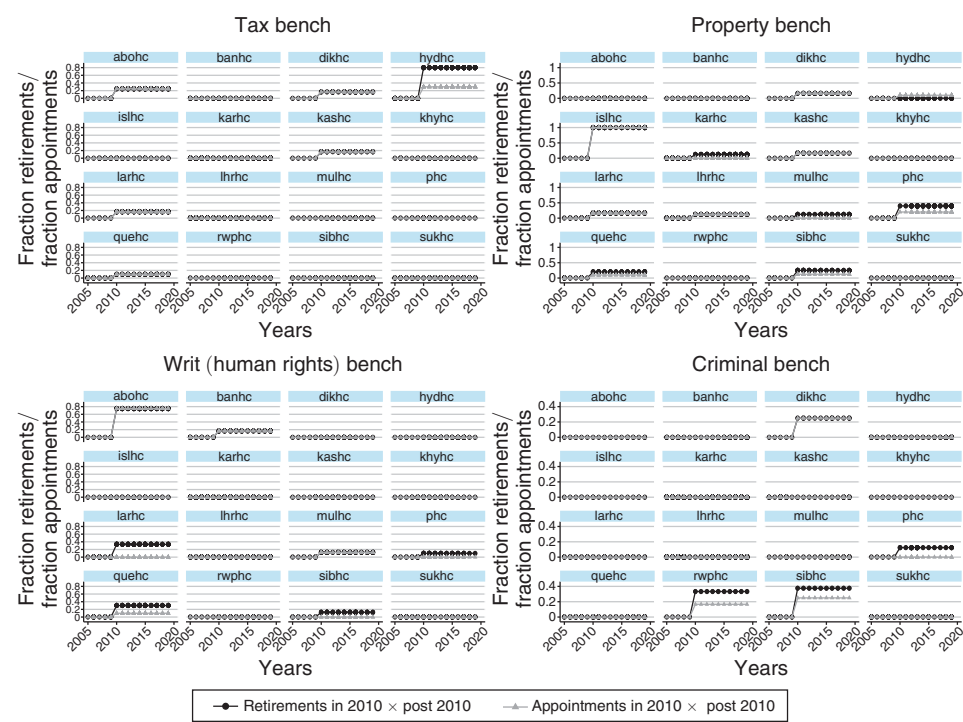


FIGURE 2. MANDATORY RETIREMENTS AND JUDICIAL COMMISSION APPOINTMENTS

Notes: The figure plots key explanatory variables, which vary at district-bench-year level. Each of the four panels shows a specialized judicial bench adjudicating cases involving tax, property, human rights, and crime, respectively. The dark line represents the fraction of judges reaching their mandatory retirement age of 62 on each district bench in 2010 interacted with the post-reform dummy. The light line represents the fraction of judges appointed by the judicial commission on each district bench in 2010 interacted with the post-reform time dummy. The correlation coefficient between these variables is 0.9. The regression-form representation of this figure with corresponding *F*-Statistics (first-stage) appears in Table C2 (panel B) of online Appendix C.

order texts. They include the district where the case was heard, the year when the case was filed, the decision year, the full name of the judge(s) adjudicating on the case, the number of lawyers and judges, the type of case, a dummy for land disputes with the government (land or eminent domain cases), and so on.

Table C1 in online Appendix C lists the means of the outcome variables, the case characteristics, and the corresponding correlation coefficients between these variables across the two teams of attorneys that coded them. Judge and bench characteristics are obtained from the judicial administrative records of the Registrar Offices of the High Courts of Pakistan (Table 1, panels C and D). This includes information on judges’ genders, ethnicities, religions, and previous employment. Holding office in the bar association and running for political office prior to judicial appointment are ascertained from biographical information in the judicial administrative records and bar association records. Combining the data from these sources gives us information on 8,446 cases and 511 judges across 64 district high court benches of Pakistan.

III. Empirical Method

A. First Specification

Our first empirical methodology uses district bench variation in mandatory retirements in reform year 2010, interacted with a post-reform dummy to estimate the effect of judicial selection reform on judicial outcomes at the case level. The corresponding specification is as follows:

$$(1) \quad Y_{cjdbt} = \theta + \alpha \left(\frac{\text{Mandatory Retirements in 2010}}{\text{Total Judges}} \times \text{Post 2010}_t \right)_{dbt} + \beta_{dt} + \mathbf{W}'_{cjdbt} \boldsymbol{\varphi} + \varepsilon_{cjdbt}.$$

The subscripts c , j , d , b , and t index cases, judges, districts, benches, and years, respectively. Y denotes state wins and judicial outcomes proxying decision quality. $\frac{\text{Mandatory Retirements in 2010}}{\text{Total Judges}} \times \text{Post 2010}$ is the fraction of judges on a given district bench reaching their mandatory retirement age in 2010, interacted with a post-reform dummy. This *Intention-to-Treat* effect can be interpreted as the effect of judicial commission appointments, since there is strong correlation between mandatory retirements and new judicial appointments. β_{dt} is district-year fixed effects, and \mathbf{W}_{cjdbt} is a vector of case and district bench controls as shown in Table 1.

Since our identifying variation comes from 64 district benches (16 districts \times 4 benches each), we cluster standard errors at the district bench level. The 64 clusters exceed the rule of thumb of 42 clusters given in Angrist and Pischke (2008, p. 219), where inference by asymptotic theory may be considered valid. Nevertheless, the results are robust to clustering by wild bootstrap for small numbers of clusters as suggested in Cameron, Gelbach, and Miller (2008) and clustering within each district bench separately pre- and post reform, as suggested in Bertrand, Duflo, and Mullainathan (2004).

The main identifying assumption in specification (1) is that mandatory retirements across district benches are as good as randomly assigned, conditional on controls—that is, exogenous to underlying factors that could have affected judicial decisions. We find this assumption plausible for several reasons. First, the number of mandatory retirements on judges' 62nd birthdays is determined by the age structure of near-retirees in 2010 and is predetermined at the time of judges' appointments. Second, anecdotal accounts suggest that the selection reform was unexpected and unrelated to specific district benches' dynamics (Almeida 2018). These two factors mean that the vacancies arising from mandatory retirements across district benches in 2010 are likely to be uncorrelated with determinants of judicial decision-making.

In the results section of the paper, we provide two additional pieces of evidence supporting this identification assumption. First, we show that pre-reform mandatory retirements leading to new judicial appointments but that were still being made by the president have no effect on government victories. Second, we

provide evidence against differential pre-trends and show that judicial outcomes on district benches that had more mandatory retirements were not on differential trajectories prior to the implementation of the selection reform. There is, however, a sharp effect of mandatory retirements on judicial outcomes precisely at the time of the reform. This strongly suggests that the pattern of judicial appointments due to mandatory retirements and rulings did not change in the years leading to the selection reform, and it is consistent with the reform's unexpected introduction.

B. Second Specification

The first specification uses cross-sectional variation across district benches—before and after the reform—to provide plausibly exogenous variation in the implementation of the selection reform. There is, however, a concern that this neglects potentially useful variation in the implementation of the reform arising from differential peer appointments occurring *after* the 2010 reform year. Therefore, we propose a second specification that utilizes all available cross-district bench and over-time variation in the implementation of the reform to estimate the effect of judicial selection reform on judicial outcomes. Specifically, we estimate the following linear probability model by OLS and 2SLS:

$$(2) \quad Y_{cjd\text{dt}} = \theta + \alpha \left(\frac{\text{Cumulative Commission Appointed Judges}}{\text{Total Judges}} \right)_{d\text{dt}} + \beta_{dt} + \mathbf{W}'_{cdt} \boldsymbol{\varphi} + \varepsilon_{cjd\text{dt}}.$$

The dependent variable, subscripts, and controls are identical to specification (1). $\frac{\text{Cumulative Commission Appointed Judges}}{\text{Total Judges}}$ is the cumulative fraction of judges appointed by the judicial commission from 2010 to 2019. This variable is instrumented by the cumulative fraction of judges expected to reach their mandatory retirement age of 62 in each district bench from 2010 to 2019, as determined by the predicted trajectory of mandatory retirements in 2010. (Figure C3 in online Appendix C shows how these variables evolve over time).

C. Discussion of the Two Specifications

This section presents the two key specifications estimated in the paper. The first one exploits variation due to mandatory retirements in only the 2010 reform year; however, it may be imprecisely estimated since it does not exploit post-2010 variation in the reform implementation. The second specification provides gains in precision by also utilizing variation in implementation of the reform post 2010. However, two comments are in order. First, the explanatory variable of interest in specification (2) necessarily follows an upward trend post reform, raising the possibility of confounding the effect of the selection reform with district-bench-specific, over-time trends. In contrast, specification (1) relies only on cross-sectional variation arising due to mandatory retirements in 2010. As a result, the explanatory

TABLE 2—IMPACT OF SELECTION REFORM ON STATE WINS

	State wins				
	(1)	(2)	(3)	(4)	(5)
Retirements in 2010 \times post 2010	−0.237 [0.0342]	−0.202 [0.0398]			
Retirements in 2009 \times post 2010			0.0640 [0.0610]		
Retirements in 2008 \times post 2010				0.126 [0.112]	
Retirements in 2007 \times post 2010					0.104 [0.0985]
District-year fixed effects	Yes	Yes	Yes	Yes	Yes
Case and bench controls	No	Yes	Yes	Yes	Yes
Observations	8,446	8,446	8,446	8,446	8,446
R^2	0.135	0.142	0.141	0.141	0.141
Mean of dependent variable	0.482	0.482	0.482	0.482	0.482

Notes: Robust standard errors appear in brackets (clustered at district bench level). The dependent variable is state wins, a dummy variable for the case being ruled in favor of the state. “Retirements in 2010” is the fraction of mandatory retirements on a given district bench in reform year 2010. “Post 2010” is a dummy for the post-reform period. See Table C2 for OLS and 2SLS estimates with appointments instead of retirements. The controls include all case and district bench characteristics shown in Table 1. The case controls also include case type fixed effects.

variable in equation (1) does not follow an upward trend post reform by construction (as can be observed when comparing Figure 2 and Figure C3). Second, specification (1) allows for a more transparent examination of pre-trends by allowing us to examine whether district benches that had more mandatory retirements in 2010 are on differential trajectories before the implementation of the reform. Importantly, however, the effects of selection reform are very similar across both specifications.

IV. Main Results

A. Effect of Judicial Selection Reform on State Wins

Table 2 (columns 1 and 2) presents the estimated effect of the judicial selection reform on state victories: there is strong and robust evidence of a substantial negative effect. The first column corresponds to the specification with only district-year fixed effects.¹² The second column adds all the available case and district bench characteristics and estimates specification (1). The estimates imply that if 10 percent of judges retired in 2010, state wins would be about 2 percentage points lower post reform. This is equivalent to a 4 percent decrease over the sample mean. In Table C2 of online Appendix C, instead of estimating the reduced-form relationship, we estimate the effect of judicial commission appointments in 2010 by OLS and

¹²The availability of case-level micro data is particularly helpful here since it allows us to flexibly account for district court and time effects that have not been possible in many other important works on courts (see, e.g., Ponticelli and Alencar 2016).

TABLE 3—IMPACT OF CUMULATIVE PEER APPOINTMENTS ON STATE WINS

	OLS		2SLS, second stage	
	(1)	(2)	(3)	(4)
<i>Panel A. Ordinary least squares and second-stage least squares results</i>				
	State wins			
Cumulative peer appointments from 2010 to 2019	−0.224 [0.0429]	−0.179 [0.0446]	−0.276 [0.0530]	−0.225 [0.0606]
District-year fixed effects	Yes	Yes	Yes	Yes
Bench and case controls	No	Yes	No	Yes
Observations	8,446	8,446	8,446	8,446
R^2	0.137	0.143	0.136	0.143
Mean of dependent variable	0.482	0.482	0.482	0.482
<i>Panel B. First-stage results</i>				
	Cumulative peer appointments from 2010			
Cumulative mandatory retirements from 2010 to 2019			0.666 [0.0613]	0.652 [0.0620]
District-year fixed effects			Yes	Yes
Case and bench controls			No	Yes
Observations			8,446	8,446
R^2			0.881	0.883
F -statistic (Montiel Olea and Pflueger 2013)			118.270	110.679

Notes: Robust standard errors appear in brackets (clustered at district bench level). The dependent variable is state wins, a dummy variable for the case being ruled in favor of the state. “Cumulative peer appointments from 2010 to 2019” is the fraction of cumulative appointments on a given district bench from 2010 onward. “Cumulative mandatory retirements from 2010 to 2019” is the fraction of mandatory retirements on a given district bench as predicted by age structure in 2010. These variables are plotted in Figure C3 of online Appendix C. The first-stage results corresponding to columns 3 and 4 appear in panel B. The F -Statistics on the first-stage results are well above both the rule of thumb of 10 and the threshold of 23 derived by Montiel Olea and Pflueger (2013) for 10 percent potential bias, 5 percent significance, and clustered standard errors. The controls include all case and district bench characteristics shown in Table 1. The case controls also include case type fixed effects.

2SLS, where we instrument appointments with mandatory retirements. The results are similar: the 2SLS estimates imply that a 10 percent rise in judicial commission appointments reduces state wins by about 3 percentage points.

We also obtain very similar results when we estimate specification (2), which exploits post-2010 variation in implementation of the reform. These results are reported in Table 3. There is evidence of a large and statistically significant negative effect of the selection reform on state wins. Panel A shows the OLS and IV (second-stage) results, while panel B presents the corresponding first stages. The first column of panel A corresponds to the OLS specification, with district-year fixed effects. Column 2 of panel A adds all the available case and district bench characteristics. In column 3, we instrument the fraction of commission-appointed judges by the fraction of the predicted trajectory of retirements based on mandatory retirements in 2010. Column 4 adds the available case and district bench controls to this IV specification. The OLS estimates are smaller than the 2SLS estimates, suggestive of a possible downward bias in OLS estimates. This would occur if new judicial commission appointments were made strategically in the most independent districts with relatively low state wins. By contrast, if a judicial commission judge

is randomly assigned to a typical district bench, she is more likely to decrease state wins than a judge assigned to an independent district bench, where most cases have low state wins to begin with.

We should, however, be cautious in this interpretation, since the effect size of 2SLS estimates is only about 25 percent larger than OLS estimates and new judicial appointments and mandatory retirements are very strongly correlated. The 2SLS estimates imply that a 10 percent rise in judges appointed by judge peers reduces state wins by about 2.5 percentage points, a 5 percent decrease over the mean dependent variable. The first stages of the 2SLS estimates are reported in panel B of Table 3. We find the instrument to be a strong predictor of the fraction of judges appointed by the judicial commission, with the F -statistic well above the critical value of 23 derived in Montiel Olea and Pflueger (2013) for 10 percent potential bias, 5 percent significance, and clustered standard errors. Overall, the estimates presented in Table 2 and Table 3 paint a consistent picture: the selection reform reduced state wins. A 10 percent increase in peer-appointed judges reduces state wins by about 4–5 percent over the sample mean.

B. Effect of Judicial Selection Reform on Decision Quality

We begin our investigation of decision quality by examining how the selection reform impacted case delay. Several influential studies argue that reduced delay in courts captures judicial efficiency and is associated with a reduction in court congestion (Djankov et al. 2003; Ponticelli and Alencar 2016). Column 1 of Table 4 reports the results on impact of selection reform on case delay. The estimated coefficient implies that retirement by 10 percent of the judges in 2010 would reduce case delay by about one month post reform. This is equivalent to about a 3 percent reduction over the sample mean and suggests that selection reform may have increased court efficiency.

It may be reasoned, however, that shorter case delay following the reform reflects reduced deliberation on cases, implying poorer-quality judicial decisions. Nevertheless, three additional pieces of evidence contradict this conjecture. First are the results for cases decided on the merits. In common law jurisprudence, rulings on merit imply that the judicial decision is “based on evidence rather than technical grounds” (Pound 1998). We constructed this variable so that it takes the value of one if the decision is “based on evidence or case merits” and zero if it is ruled on a technicality. This variable, also coded by the legal experts, follows the argument by Pakistani legal scholars that ruling on technicalities is a “weapon of choice to rule unfairly” (Siddique 2013) and “symptomatic of a biased decision” (Javed Arshad, interview with the author, 2017). Column 2 of Table 4 reports the results of estimating equation (1) with merit as the dependent variable. The estimates indicate that the selection reform increased decisions based on evidence: if 10 percent of judges retired from their respective district benches in 2010, merit decisions would increase by about 2 percentage points (Table 4, column 2). This is equal to a 3 percent increase over the sample mean. These results are consistent with scholarship as well as our discussions with legal experts (senior lawyers and judges) in Pakistan who suggested that decisions on the merits would approximate correct or unbiased

TABLE 4—SELECTION REFORM AND DECISION QUALITY

	Case delay (1)	Merit (2)	Correct decisions (3)	Process followed (4)
Retirements in 2010 × post 2010	−0.878 [0.387]	0.215 [0.0396]	0.191 [0.0515]	0.425 [0.127]
District-year fixed effects	Yes	Yes	Yes	Yes
Case and bench controls	Yes	Yes	Yes	Yes
Observations	8,446	8,446	8,446	8,446
R ²	0.218	0.141	0.098	0.080
Mean of dependent variable	3.354	0.627	0.469	3.314

Notes: Robust standard errors appear in brackets (clustered at district bench level). In column 1 the dependent variable is case delay, i.e., the difference between filing and decision year. In column 2 it is a dummy variable for the case being ruled on the merits or evidence. In column 3 the dependent variable is a dummy variable for the case being judged as correct by legal experts. In column 4 the dependent variable is a rating from one to five on decision quality. “Retirements in 2010” is the fraction of mandatory retirements on a given district bench in reform year 2010. “Post 2010” is a dummy for the post-reform period. The controls include all case and district bench characteristics shown in Table 1. The case controls also include case type fixed effects.

judicial decisions (Siddique 2013). This checks out quantitatively when we directly code the judicial decisions along this dimension, i.e., via a dummy that switches on for “correct” judicial decisions. Column 3 of Table 4 reports these results. We find that a 10 percent increase in judges retiring in the reform year increases “correct” decisions by 1.91 percentage points.

Finally, we ascertain whether the fall in state wins also coincides with better observance of due process. Several legal scholars argue that a higher-quality judicial decision should not only be unbiased or correct but also needs to follow due process of law (see, for instance, seminal treatment of this issue in Dahl 1957). To proxy for whether due legal process was followed, we leverage legal experts’ ratings for each judicial case on the observation of “relevant jurisdictional, procedural, and evidential requirements.” This is a proxy for the “correctness” of the legal process followed in reaching the judicial decision. A higher rating on “process followed” implies that higher jurisdictional, procedural, and evidential standards were met in reaching the judicial decision. Estimating equation (1) with this dependent variable indicates that if 10 percent of the judges retired from their respective district benches in 2010, the “process followed” rating would increase by about 0.04 points on a 5-point scale (Table 4, column 4). This is equivalent to about a 1.3 percent increase in rating over the sample mean. Taken together, the results from Table 4—of the selection reform reducing case delays, increasing following of due process of law, meritorious and correct decisions—all corroborate the view that peer-appointed judges issue higher-quality judicial decisions relative to rulings by presidential appointees.

C. Threats to Identification

We now examine two key threats to identification that could undermine the causal interpretation of these estimates. First, we may be confounding the effect of selection reform with differences in behavior toward the state by old and new

judges—a pure appointment effect. Table 2 (columns 3, 4, and 5) presents evidence against this hypothesis by showing that pre-reform, mandatory retirements have no effect on rulings in favor of the government. That is, the age distribution of judges across district benches in the years 2007, 2008, or 2009, when all new judicial appointments were still by the president, does not reduce state wins post reform. If anything, the coefficient estimates in all these instances are positive. This indicates that if there is a pure appointment effect, it is likely to be small and perhaps go in the opposite direction, with new appointees more likely to issue progovernment rulings. Similar results are obtained for the decision-quality variables with pretreatment retirements uncorrelated with all four of our available decision-quality outcomes: in Table C3 of online Appendix C, we show that pre-reform mandatory retirements have no impact on case delay, due process followed, merit, or correct decisions.

The findings of a significant effect of mandatory retirements in the selection reform year and no effect of pre-reform mandatory retirements are robust to different specifications (such as combining pretreatment and treatment variables in a single regression (Table C4 in online Appendix C). They are also insensitive to estimating the cumulative retirement specification, where we also find no effect of pre-reform retirements (Table C5 of online Appendix C). Only the reduced relationship between post-reform cumulative retirements and state wins is negative and statistically significant, while pre-reform cumulative retirements are not predictive of state wins.

A second threat to identification might come from confounding the effect of the reform with diverging trends prior to the reform. District benches with more vacancies arising from mandatory retirements could already have been following different trajectories before the reform. Therefore, to systematically examine pre-trends, we estimate the following specification:

$$(3) \quad Y_{cjd\text{b}t} = \beta_0 + \sum_{s=1986}^{2019} \alpha_s \left(\frac{\text{Mandatory Retirements in } 2010_{db} \times \delta_s}{\text{Total Judges}_{db}} \right)_{dbt} + \beta_{dt} + \mathbf{W}'_{cjd\text{b}t} \boldsymbol{\varphi} + \varepsilon_{cjd\text{b}t},$$

where Y is a given judicial outcome, δ_s is a dummy variable that takes the value of one in the year s , and $\frac{\text{Mandatory Retirements in } 2010}{\text{Total Judges}}$ is the fraction of judges reaching their mandatory retirement age of 62 in reform year of 2010. District-year fixed effects and controls are identical to the specifications (1) and (2). Equation (3) is conceptually identical to the specification used to test for pre-trends in Martinez-Bravo (2017), and it allows us to transparently assess systematic differences among district benches prior to the selection reform and to investigate whether the selection reform had a persistent effect. Figure 3 (and its corresponding table-form representation in Table C6 of online Appendix C) presents the results from estimating equation (3). We find that district benches with a higher fraction of mandatory retirements in 2010 show no change in state wins prior to 2010, although there is a sharp decrease in state wins post reform. The effect is likely persistent, as indicated by qualitatively and

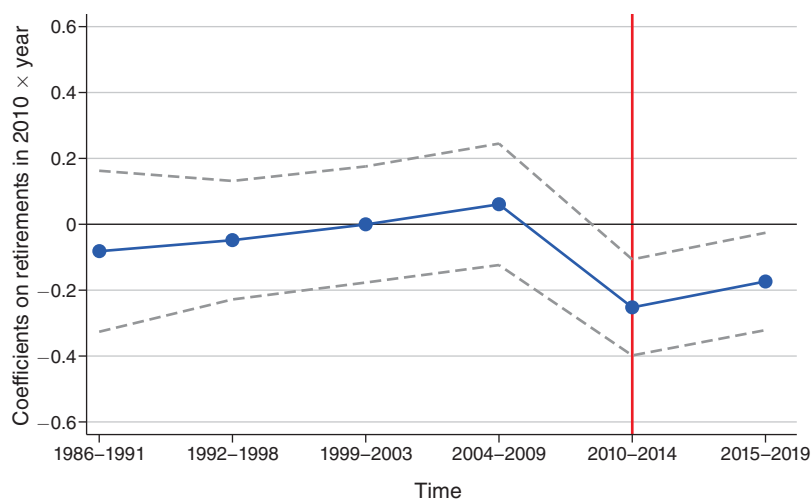


FIGURE 3. IMPACT OF MANDATORY RETIREMENTS IN 2010 ON STATE WINS OVER TIME

Notes: This figure presents the coefficients and their 95 percent confidence intervals when we estimate equation (3). The table-form representation of this figure appears in Table C6 of online Appendix C. District benches that had more mandatory retirements in 2010 display no change in state wins prior to the reform year of 2010, while there is a sharp fall in government victories post reform.

statistically significant effects of the selection reform observed post 2014. Similar evidence is found when we estimate equation (3) using the available decision quality measures. Figure 4 present these results for case delay (panel A), merit decisions (panel B), correct decisions (panel C), and due process followed (panel D). We observe a sharp break post reform and not much evidence of differential trends prior to the reform across all four of the outcomes. These results corroborate anecdotal accounts that the selection reform was an unanticipated shock to the judicial system of Pakistan (Almeida 2018) and suggest that differential pre-trends are unlikely to explain our results.

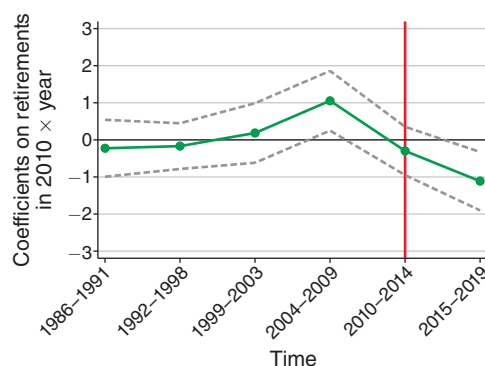
V. Mechanisms

This section is organized into two brief subsections. The first one provides supporting evidence consistent with the judicial selection mechanism by documenting the judge characteristics correlated with the peer appointees. The second subsection provides a lack of evidence for spillover or peer effects from judicial-commission-appointed judges on presidential appointees.

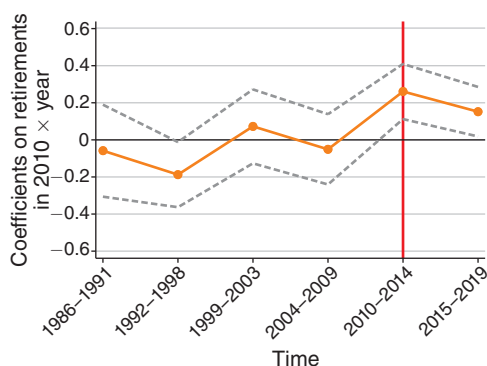
A. What Does Peer Appointment Select For?

First, we examine the key observed differences in the judges selected by judge peers relative to presidential appointees. The evidence suggests selection effects of judges with different characteristics selected under the two regimes as the likely mechanism behind the impact of selection reform on government victories. In

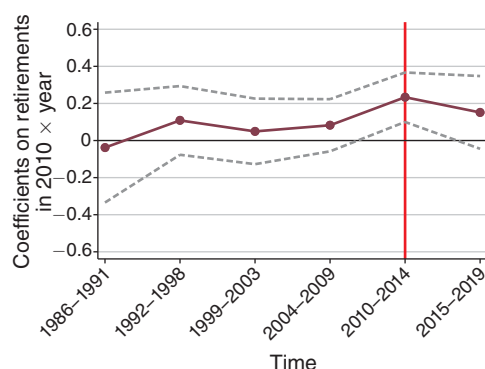
Panel A. Impact of selection reform on case delay



Panel B. Impact of selection reform on merit decisions



Panel C. Impact of selection reform on correct decisions



Panel D. Impact of selection reform on process followed

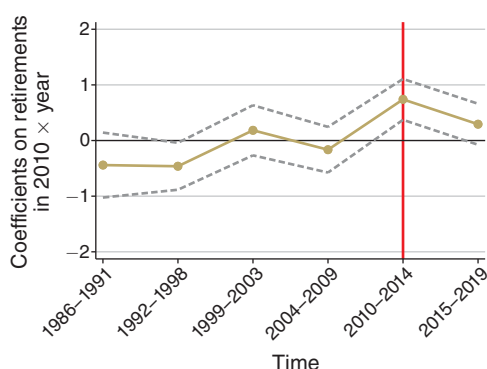


FIGURE 4. IMPACT OF MANDATORY RETIREMENTS IN 2010 ON DECISION QUALITY

Note: This figure presents the coefficients along with their 95 percent confidence intervals when we estimate equation (3) using the four available measures of decision quality as dependent variables: case delay, merit decisions, correct decisions, and process followed.

TABLE 5—JUDICIAL COMMISSION APPOINTEES AND JUDGE CHARACTERISTICS AT JUDGE LEVEL

	Gender (1)	Muslim (2)	Former lawyer (3)	Punjabi ethnicity (4)	Sindhi ethnicity (5)	Balochi ethnicity (6)	Pashtun ethnicity (7)	Former officeholder bar assoc. (8)	Ran for political office (9)
Post-reform judge	−0.0361 [0.0246]	−0.00636 [0.0142]	0.0148 [0.0403]	−0.0188 [0.0503]	0.0316 [0.0318]	−0.00357 [0.0326]	−0.0433 [0.0440]	−0.328 [0.0582]	−0.164 [0.0439]
Age control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District bench controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	511	511	511	511	511	511	511	511	511
R^2	0.025	0.009	0.040	0.045	0.020	0.037	0.017	0.174	0.066
Mean of dependent variable	0.961	0.990	0.886	0.195	0.053	0.063	0.143	0.569	0.186

Notes: Robust standard errors appear in brackets (clustered at the judge level). The dependent variable is state wins, judge-level average of the dummy for the case being ruled in favor of the state. “Post-reform judge” is a dummy variable that takes a value of one if the judge is appointed by the judicial commission and zero for presidential appointees. The controls include judge-level averages of case and district bench characteristics shown in Table 1. The unit of observation in this table is an individual judge; for analogous case-level regressions, see Table C7 in online Appendix C.

Table 5, we present correlates of peer appointees across a number of observed judicial characteristics. We present here results at the judge level, although identical results are found if we instead run case-level regressions (see Table C7 in online Appendix C). Peer appointments are uncorrelated with judges' gender, religion, previous employment, and ethnicity. Finding peer appointment uncorrelated with a judge's ethnicity is especially interesting, since it conflicts with concerns expressed in some quarters of Pakistan's legal fraternity that peer appointment may select judges from specific ethnic groups, i.e., from the politically powerful state of Punjab (Siddique 2013). We find no evidence of this in the data. Ethnic discrimination in the selection of judges is no greater under peer appointments for any major ethnic group in Pakistan.

Nevertheless, the two types of appointees differ in their observed political activity prior to their judicial appointments. Commission-appointed judges are about 35 percent less likely to have held political office in the lawyers' bar associations (Table 5, column 8). As candidates for office in Pakistani bar associations often run on a political party's platform, we consider this a plausible proxy for political activity prior to judicial appointment. We also provide more evidence that judicial commission appointees are significantly less likely to have run for political office in state or national elections before their judicial appointments.¹³ Specifically, we find that peer-appointed judges are also about 15 percent less likely to have run for elections in the provincial or national assembly prior to their judicial appointments (column 9, Table 5). We interpret these results in combination with anecdotal accounts that suggest that the peer appointment system is likely to be more meritocratic and averse to selecting political ideologues (Zafar 2012) or partisans. Together, they suggest that the selection of relatively apolitical judges by the judicial commission is likely important in explaining our results.

B. Do Peer Appointees Have Spillover Effects on Presidential Appointees?

In this subsection, we examine whether peer-appointed judges' arrival induces spillover effects on presidential appointees. To do this, we compare state wins at district benches with more mandatory retirements in 2010 with those at district benches with fewer mandatory retirements pre- and post 2010, but only for cases where all the judges are appointed by the president. Since these are the very places where peer-appointed judges are most likely to be appointed, given the strong correlation with new appointments, we can assess how presidential appointees react to the arrival of judicial commission appointees. Table 6 reports these results. Estimates from columns 1 to 4 paint a picture of no significant change in the behavior of old-regime judges with the arrival of peer-appointed judges. These results suggest that peer effects are not likely to be large (if present at all) and that the arrival of peer appointees is unlikely to have a large impact on incentives facing old-regime presidential appointees.

¹³ Once appointed, judges are barred from running for political office until two years after their retirement.

TABLE 6—IMPACT OF SELECTION REFORM ON PRESIDENTIAL APPOINTEES (SPILLOVER EFFECTS)

	State Wins			
	(1)	(2)	(3)	(4)
Retirements in 2010 × post 2010	−0.0718 [0.0594]	−0.121 [0.0774]	−0.102 [0.0788]	−0.0825 [0.0816]
District-year fixed effects	No	Yes	Yes	Yes
Bench controls	No	No	Yes	Yes
Case controls	No	No	No	Yes
Observations	6,629	6,629	6,629	6,629
R ²	0.0001	0.112	0.120	0.122
Mean of dependent variable	0.53	0.53	0.53	0.53

Notes: Robust standard errors appear in brackets (clustered at district bench level). The dependent variable is state wins, a dummy variable for the case being ruled in favor of the state. “Retirements in 2010” is a fraction of mandatory retirements on a given district bench in the reform year (2010). “Post 2010” is a dummy for the post-reform period. The case controls also include case type fixed effects. In this table we only consider the restricted sample of cases decided by presidential appointees (serving both pre- and post reform).

VI. Heterogeneity

This section is organized into three subsections. First, we discuss the heterogeneous effect of the reform by the political saliency of the case. Second, we present results on a back-of-the-envelope calculation of the likely value of land expropriations avoided every year due to the institution of the peer appointment system. Last, we present evidence that the peer appointment system does not have the unintended consequence of increasing pandering to religious sentiments on the street or facilitating corporate capture.

A. Heterogeneity by Politically Salient Cases

In this subsection we investigate the heterogeneous effects of selection reform by the political saliency of the cases. We find evidence that the judicial selection reform particularly affected politically salient constitutional cases. These are cases involving land expropriation and human rights abuses by the state. In columns 1 and 2 of Table 7 (panel A) these results are reported: a 10 percent rise in judicial commission appointees reduces state wins by about 2.5 percentage points in constitutional disputes with the state. Likewise, in columns 1 and 2 of Table 7 (panel B), we disaggregate constitutional cases into those involving political rights and land expropriation by state: a 10 percent rise in judicial commission appointees reduces state wins by 2.3 percentage points in human rights cases and by 3.1 percentage points in cases involving land expropriation by the state. These results are also unlikely to be driven by differential trends prior to the reform, since limiting our sample to constitutional cases or their constituent land or political rights cases also reveals no evidence of pre-trends, although there is an unsurprising decrease in precision due to the smaller sample size as a result of the sample restriction (Figure C4 of online Appendix C). Likewise, these results are robust to using the cumulative retirement specification (Table C8 of online Appendix C).

TABLE 7—IMPACT OF SELECTION REFORM ON STATE WINS (BY TYPE OF CASE)

<i>Panel A. Constitutional versus criminal cases</i>				
	State wins			
	Constitutional cases		Criminal cases	
	(1)	(2)	(3)	(4)
Retirements in 2010 \times post 2010	−0.270 [0.0304]	−0.258 [0.0271]	0.261 [1.365]	2.635 [2.714]
District-year fixed effects	Yes	Yes	Yes	Yes
Case and bench controls	No	Yes	No	Yes
Observations	6,094	6,094	2,368	2,368
R^2	0.158	0.159	0.277	0.281
Mean of dependent variable	0.456	0.456	0.548	0.548

<i>Panel B. Constitutional and criminal cases disaggregated</i>				
	State wins			
	Constitutional cases		Criminal cases	
	Human rights cases	Land cases	Non-Islamic case	Islamic case
	(1)	(2)	(3)	(4)
Retirements in 2010 \times post 2010	−0.230 [0.0460]	−0.307 [0.0471]	5.379 [3.387]	−1.895 [1.224]
District-year fixed effects	Yes	Yes	Yes	Yes
Case and bench controls	Yes	Yes	Yes	Yes
Observations	3,428	2,650	2,143	225
R^2	0.219	0.218	0.286	0.763
Mean of dependent variable	0.462	0.449	0.552	0.520

Notes: Robust standard errors appear in brackets (clustered at district bench level). The dependent variable is state wins, a dummy variable for the case being ruled in favor of the state. “Retirements in 2010” is the fraction of mandatory retirements on a given district bench in reform year 2010. “Post 2010” is a dummy for the post-reform period. Panel A reports the disaggregated results into constitutional and criminal cases. The constitutional and criminal cases do not add up to the 8,446-case sample because 16 criminal cases are also marked as constitutional. Panel B shows further disaggregation of constitutional cases into human rights and land cases, and of criminal cases into those judged under British common law and Islamic Hudood law. “Islamic cases” includes cases pertaining to consumption of alcohol, blasphemy, adultery, homosexuality, fornication, and false accusation of fornication. The controls include all case and district bench characteristics shown in Table 1. The case controls also include case type fixed effects.

This reduction in government victories in politically salient constitutional cases is corroborated by widespread qualitative accounts. For instance, expropriation of private property by government or “outright takings” is considered a major problem in Pakistan and many developing countries (World Justice Project 2020). Rulings in these ownership or expropriation disputes with the government are reported to be heavily influenced by political considerations (Abbasi 2017; Sattar 2017). Some legal scholars and lawyers in Pakistan go as far as to argue that land disputes involving the state are instances where the government is almost always in the wrong. For instance, “when you see [a government] housing agency involved in a land case, you know that justice is dead” (W. Sheikh, interview with the author, 2016), or “these housing development authorities [are] a mafia that operates with the full support of the highest level of the government ... some judges are part of it too” (Javed Arshad, interview with the author, 2017). This is consistent with survey evidence from across the world, with over 60 percent of respondents in developing countries

expressing fear of the expropriation of their private property by the state (World Justice Project 2020). Similarly, human rights or political rights cases involving the state are also considered highly political in nature. These constitutional cases are separately marked as “writ petitions.” These cases involve the violation of political rights, such as the persecution of political opponents, limiting freedom of movement, or discrimination based on gender, political affiliation, or caste. Typical examples include an opposition politician claiming that his fundamental right to freedom of movement within and outside Pakistan has been restricted by the government since he joined the opposition political party or an opposition leader pleading that his citizenship was cancelled a day before he was to lead a protest against the government (Naseer 2018).

Further evidence of case type heterogeneity by political saliency comes from a falsification test. As petty crime cases also involve the state (as prosecutor) but are politically less salient, we examine the impact of the selection reform on state wins in criminal cases.¹⁴ These criminal cases mostly involve petty crime, vandalism, minor fraud, theft, and burglary and, hence, are low stakes politically. Results from Table 7 (panel A, columns 3 and 4) show that the selection reform has no significant effect on state wins in criminal cases, and the point estimates are in fact positive. This suggests that judicial commission appointees do not rule against the government more than presidential appointees in less politically salient criminal cases. We interpret these results in light of the selection- or judge-heterogeneity mechanism, with peer-appointed—and relatively apolitical—judges making fewer rulings in the government’s favor in politically salient cases, while rulings in run-of-the-mill criminal cases are unaffected.

B. Land Expropriations Prevented under the Selection Reform

Our investigation of heterogeneous effects by type of cases highlighted reduced progovernment rulings in constitutional cases involving land expropriations and human rights disputes with the state. It is not straightforward to assess the economic value of the decrease in progovernment rulings in human rights abuse cases. However, it is possible to perform a back-of-the-envelope approximation on the economic value derived from avoiding land expropriations due to peer-appointed judges. This is similar to Mian and Khwaja’s (2005) computation of the economy-wide costs of political connections using minimum and maximum bounds. Using property values taken from judgment orders, our point estimates on impact of selection reform, and total land cases decided in this period, we suggestively infer that the selection reform may have been able to prevent land expropriations worth 0.07 to 0.3 percent of GDP every year.¹⁵ This averages out to about 0.14 percent of GDP

¹⁴ It should be noted that politically salient criminal cases—for instance, those involving corruption of political opponents—are adjudicated in national accountability courts, whereas our analysis concern the high courts that adjudicate over petty and violent crime cases in addition to the politically salient constitutional petitions against the government.

¹⁵ The calculation is made as follows: in 20 percent of our 8,500 randomly sampled cases, the government was successful in expropriating land. Since we randomly sampled 0.2 percent of the total population of cases, the total number of successful land expropriation cases is about 850,000. Basing computations on an average value of \$51,280 for the 57 expropriated properties whose market values are listed in judgment texts and assuming all

or \$390 million in land expropriations avoided every year through peer-appointed judges. To put this amount into perspective, it is equivalent to about 70 percent of Pakistan's federal budget for education or nearly the whole amount earmarked for health care in 2019.

C. Heterogeneity by Cases Involving Islamic Law and Large Corporations

We now present evidence that peer appointees are neither more likely to pander to religious sentiments on the street nor more amenable to corporate capture. First, we disaggregate criminal cases judged under Islamic law from those judged under secular British common law. The Islamic cases judged under Islamic limits or Hudood law pertain to cases involving consumption of alcohol, blasphemy, adultery, homosexuality, fornication, and false accusation of fornication that are criminal offenses under Pakistan's penal code (PPC). The state wins or convictions judged under Islamic law are reported in Table 7, panel B (column 4). The results are similar: peer-appointed judges are not significantly more likely to convict for violations of Islamic law than presidential appointees.

Second, we investigate whether the reduced rulings in favor of the state due to the selection reform come at the expense of increased rulings in favor of large corporations. Since about 15 percent of our sample involves the state opposing a firm rather than a citizen, we can test this hypothesis. We examine the effect of selection reform in the subsample of cases involving small or medium-sized enterprises as well as large firms. We consider a firm to be large if it is associated with any of the 12 big business groups of Pakistan (Tech 2020).¹⁶ Table 8 reports these results. Consistent with our main findings, we find that when peer appointments rise by about 10 percent, government victories against small and medium-sized enterprises fall by about 2 percentage points (columns 1 and 2). Nevertheless, in cases involving large corporations, the effect is reversed but marginally significant, possibly because of lower power due to the smaller number of large firms. Peer appointees are about 2.5 percentage points more likely to issue rulings in favor of the government when the state is opposing a large firm (Table 8, Column 4). This suggests a heterogeneous effect of the selection reform: peer appointees issue more progovernment rulings when the state is opposing large corporations and are less likely to succumb to corporate capture. It should be noted, however, that the evidence presented in the paper does not imply that peer appointment of judges is foolproof but rather that judges selected by their peers are likely to issue better decisions across a suit of measurements and are no more amenable to corporate capture or to pander to religious sentiments on the street than judges selected by the president.

judges are replaced by peer-appointed judges, government victories then fall by about 20 percentage points. We thus estimate the value of land expropriations prevented due to the selection reform as ranging from 0.07 to 0.3 percent of GDP during 2010–2019; however, we should be cautious in interpreting these computations, since they rely on all presidential appointees replaced by judge peers while we use partial equilibrium estimates in these computations. For more details on these computations, see Table C9 in online Appendix C.

¹⁶ This includes all cases involving firms owned by large conglomerates—for instance, Mian Muhammad Mansha's Nishat Group, Byram Avari's Avari Group, Razaq Dawood's Descon Group, Mian Muhammad Latif's Chenab Group, Mian Amir Mahmood's Dunya News Group, and Malik Riaz's Bahria Town Group (see Tech 2020 for the complete list).

TABLE 8—SMALL AND MEDIUM-SIZED FIRMS AND LARGE CORPORATIONS VERSUS THE STATE

	State wins			
	Small and medium-sized firms		Large corporations	
	(1)	(2)	(3)	(4)
Retirements in 2010 \times post 2010	−0.248 [0.144]	−0.203 [0.150]	0.189 [0.109]	0.237 [0.128]
District-year fixed effects	Yes	Yes	Yes	Yes
Case and bench controls	No	Yes	No	Yes
Observations	864	864	435	435
R^2	0.434	0.450	0.594	0.626
Mean of dependent variable	0.466	0.466	0.462	0.462

Notes: Robust standard errors appear in brackets (clustered at district bench level). The dependent variable is state wins, a dummy variable for the case being ruled in favor of the state. We consider a subsample of cases where the state is opposing a firm. We distinguish between small and medium-sized firms and large corporations, where we classify a firm as “large” if it is associated with the “Big 12” business groups of Pakistan (Tech 2020). The large corporations include all firms owned by Mian Muhammad Mansha’s Nishat Group, Byram Avari’s Avari Group, Razaq Dawood’s Descon Group, Mian Amir Mahmood’s Dunya News Group, Mian Muhammad Latif’s Chenab Group, and Malik Riaz’s Bahria Town Group. “Retirements in 2010” is a fraction of mandatory retirements from a given district bench in reform year 2010. The controls include all case and district bench characteristics shown in Table 1. The case controls also include case type fixed effects.

VII. Robustness

A. Robustness to Alternative Explanations

This subsection tests alternative explanations for the finding that the reform generated a change in judicial decision-making in Pakistan and investigates the robustness of our results to competing explanations. These include the possibilities that our results are driven by an idiosyncratic president or a specific chief justice or are simply due to changes in the types of cases adjudicated by the two types of judges.

President-Specific Effect.—First, the documented effect of the reform may be a president-specific effect. For instance, the fall in state wins post reform may simply reflect a correction from extremely high state wins during the tenure of an idiosyncratic president in the pre-reform period (say, President General Musharraf). Since judges appointed by five different presidents are included in the sample, we can examine this claim empirically. That is, we can compare rulings by judges appointed by the judicial commission with rulings by judges appointed by different presidents. We find no evidence that the effect of the selection reform is a president-specific effect (Table 9).

Chief-Justice-Specific Effect.—It could also be argued that the change associated with the reform is a chief-justice-specific effect: some chief justices in Pakistan are considered to be particularly antigovernment. As the chief justice of Pakistan is the head of the judicial commission, this alternative explanation is important to investigate. Since seven different chief justices headed the judicial commission during our

TABLE 9—IMPACT OF SELECTION REFORM ON STATE WINS (BY APPOINTING PRESIDENT)

	State Wins				
	Pres. Musharraf (1)	Pres. Tarar (2)	Pres. Leghari (3)	Pres. Khan (4)	Pres. Haq (5)
Retirements in 2010 × post 2010	−0.204 (0.0504)	−0.542 (0.0724)	−0.451 (0.0608)	−0.454 (0.0720)	−0.492 (0.0703)
District-bench and -year fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Observations	3,539	1,817	3,032	2,616	3,018
R ²	0.199	0.267	0.253	0.281	0.244
Mean of dependent variable	0.432	0.321	0.401	0.486	0.403

Notes: Robust standard errors appear in brackets (clustered at district bench level). The dependent variable is state wins, a dummy variable for the case being ruled in favor of the state. “Retirements in 2010” is a fraction of mandatory retirements from a given district bench in reform year 2010. “Post 2010” is a dummy for the post-reform period. The judicial outcomes of cases adjudicated by peer-appointed judges are compared to those of judges appointed by the last five presidents prior to selection reform. The sample size varies according to presidents’ length of time in office, which gave them differing opportunities to fill new judicial vacancies. The case controls also include case type fixed effects.

sample period, we therefore test and reject the hypothesis that the results are driven by an idiosyncratic chief justice. These results are reported as Table C10 in online Appendix C. The lack of evidence that the selection reform is a chief-justice-specific effect also suggests that the effect is unlikely to be confined to the tenure of the Pakistan People’s Party government that implemented the judicial selection reform. In fact, from Table C10, we can also observe that the selection reform is qualitatively and statistically significant at conventional levels after PPP is voted out of office in 2013 (Table C10, column 5).

Strategic Case Filings.—Finally, an alternative mechanism might involve litigants responding to the selection reform by changing their case-filing behavior (Klein and Priest 1984; Hubbard 2013). We conduct two tests to investigate this possibility. First, we examine the differences in the types of cases decided in court. If strategic case filings were driving the results, we would observe that cases adjudicated at district benches with more peer appointees differ from cases adjudicated at district benches with fewer peer appointees. Results from Table 10, where we observe that the selection reform is not correlated with a long list of case and district bench characteristics, indicate that this is highly unlikely, suggesting that districts with more versus fewer mandatory retirements are adjudicating similar types of cases. This “balance check” is robust to a cumulative retirements specification: the selection reform is similarly uncorrelated with a long list of case and bench characteristics (online Appendix C). These results also suggest that the chief justice endogenously constituting the bench or strategically assigning cases or specialties to judges is unlikely to explain the impact of selection reform on judicial decision-making. If the chief justice did strategically allocate cases, we would have observed the selection reform to be correlated with case characteristics. However, we find no evidence of this across either specification.

TABLE 10—IMPACT OF SELECTION REFORM ON CASE AND BENCH CHARACTERISTICS

	Constitutional case (1)	No. pages (2)	CJ on case (3)	No. lawyers on case (4)	No. judges on case (5)	No. judges on bench (6)	No. criminal cases on bench (7)	No. land cases on bench (8)	No. human rights cases on bench (9)
Retirements in 2010 × post 2010	−0.00303 (0.00644)	−0.112 (0.580)	−0.00964 (0.0171)	−0.145 (0.179)	0.0422 (0.0685)	−0.160 (0.175)	2.164 (1.177)	0.488 (0.683)	0.750 (0.668)
District bench fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,446	8,446	8,446	8,446	8,446	8,446	8,446	8,446	8,446
R^2	0.991	0.292	0.110	0.068	0.138	0.962	0.763	0.663	0.662
Mean of dependent variable	0.722	8.877	0.065	4.123	1.809	8.025	2.77	3.476	2.586

Notes: Robust standard errors appear in brackets (clustered at the district bench level). “Retirements in 2010” is the fraction of mandatory retirements from a given district bench in reform year 2010. “Post 2010” is a dummy for the post-reform period. The controls include case and bench characteristics outlined in Table 1 (except the dependent variable used in each respective column). The case controls include case type fixed effects.

Second, we directly test for strategic case-filing behavior by investigating whether litigants change their filing behavior in response to the reform. We find no evidence that total case filing or case filing in politically salient constitutional cases is affected by the selection reform (online Appendix C, Table C12). We interpret these results in light of recent evidence from Acemoglu et al. (2020) that documents litigants in Pakistani courts to be unaware of reductions in case delay when they occur and that informing litigants about falls in delays increases their likelihood of using the formal courts. Together, these results suggest that selection reform did not induce a change in case-filing behavior.

B. Additional Sensitivity Tests

We carry out a number of additional robustness checks, the results of which are presented in Tables C13 to Table C20 of online Appendix C. First, we show that our results are not sensitive to excluding cases decided in federal and provincial capitals (Table C13). Courts in these capital judicial districts are called “principal benches” and some qualitative accounts suggest that the most politically salient cases are decided there (Ikram Haq, interview with the author, 2018). Nevertheless, our point estimates are very similar when we exclude cases decided in these political capitals. This supports the idea that the effect of the reform is not geographically confined to cases decided in political capitals. Second, we show that our results are not driven by a specific judicial bench (Tables C14 and C15). Third, we show that the results are robust to different starting years. We choose to go back as far as 1986 in order to use all available data that we accessed. However, while going back in time allows us to gain precision, it may raise concerns regarding omitted variables. Therefore, we show that the results are robust to different starting years (Table C16). Fourth, our results are not sensitive to the aggregation of variables to the level of variation of the explanatory variable or nonlinear models such as

Probit or Logit (Tables C17 and C18, respectively). Fifth, we show in Table C19 that the results are robust to different levels of clustering: within each district bench, separately, before and after the reform (Bertrand, Duflo, and Mullainathan 2004); district-level clustering; or using the wild bootstrap method for a small number of clusters, as per Cameron, Gelbach, and Miller (2008).¹⁷ Finally, we show that essentially identical results are obtained across all our available judicial outcome variables when we use data coded by Team 2 (Table C20).

VIII. Conclusion

The removal of presidential appointment of judges was described by the prominent legal historian and former judge S.M. Zafar as follows:

Our judiciary or any other judiciary in the world should be independent to make correct decisions. We had provisions for this in the 1973 Constitution. Judges' tenure was secure. But this strong judicial dispensation was suffering from a key ailment or a flaw of the induction of the judges into the judiciary. The previous procedure was very arbitrary[;] it was not only arbitrary, but it was noninstitutional, discretionary and as a result bad judges, sometimes political ideologues, sometimes crazies got into the judiciary. So, we thought that the judicial appointments [should] be through a process, we institutionalized it. We took away the power, the discretionary power, the arbitrary power (Zafar 2012).

He may have been right. This paper has shown that removing the power of judicial appointment from the president and transferring it to peer judges substantially reduces government victories and likely increases decision quality. The identification strategy we propose allows us to obtain plausibly causal effects. We present evidence against a number of threats to identification and alternative explanations for our findings. The evidence suggests that these results are likely driven by judge peers selecting more meritocratic and bipartisan judges relative to judicial selection by the president. Last, we conduct a back-of-the-envelope computation showing that the selection reform likely prevents land expropriations amounting to about 0.14 percent of GDP or USD 390 million every year.

Research examining the selection of public officials has largely focused on politicians. Our work focuses on the judiciary and advances the long-standing debate regarding presidential appointment of judges. The results of this study highlight the potential to reform the judiciary in weakly institutionalized (low state capacity) settings and are consistent with recent evidence that tenure security for judges does not necessarily yield a *de facto* increase in judicial independence (Hayo and Voigt 2019). Although many countries have instituted "security of tenure" with presidential appointment of judges, as supported by the Federalists (Madison, Hamilton, and Jay [1788] 1961), our results suggest that this may not promote rule of law relative to other possible institutional arrangements. Some countries have therefore

¹⁷ The standard error concept does not formally apply for small numbers of clusters when implementing wild bootstrap. Therefore, we compute *p*-values and confidence intervals instead, as suggested in Roodman et al. (2019). The results are still significant at the 5 percent level and are presented in Figure C5 of online Appendix C.

innovated with new institutional arrangements, such as peer appointment of judges. The evidence presented in this paper corroborates the view that the innovation of peer appointment of judges—as introduced in Sweden, Greece, and the United Kingdom—may be an improvement over presidential appointment with security of tenure. More research on the judiciary, particularly on other judicial selection mechanisms, will further clarify the counterfactual policy choices available to policymakers as they grapple with institutional reform in vulnerable democracies. This may also provide a deeper understanding of the conditions necessary for the establishment of the rule of law.

REFERENCES

- Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. 2017. “When Should You Adjust Standard Errors for Clustering?” NBER Working Paper No. 24003.
- Abbasi, Ansar. 2017. “Plots Allotment to Judges a Favour, Misconduct: Ex-CJ.” *News*, February 24. <https://www.thenews.com.pk/print/200999-Plots-allotment-to-judges-a-favour-misconduct-ex-CJ>.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. “The Colonial Origins of Comparative Development: An Empirical Investigation.” *American Economic Review* 91 (5): 1369–1401.
- Acemoglu, Daron, James A. Robinson, and Ragnar Torvik. 2013. “Why Do Voters Dismantle Checks and Balances?” *Review of Economic Studies* 80 (3): 845–75.
- Acemoglu, Daron, Ali Cheema, Asim I. Khwaja, and James A. Robinson. 2020. “Trust in State and Nonstate Actors: Evidence from Dispute Resolution in Pakistan.” *Journal of Political Economy* 128 (8): 3090–147.
- Almeida, Cyril. 2018. “Re-Centralisation.” *Dawn*, March 25. <https://www.dawn.com/news/1397406/re-centralisation>.
- Amnesty International. 1982. *Pakistan Human Rights Violations and Decline of Rule: An Amnesty International Report*. London: Amnesty International Publications.
- Angrist, Joshua, and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton, NJ: Princeton University Press.
- Ash, Elliott, and W. Bentley MacLeod. 2019. “The Performance of Elected Officials: Evidence from State Supreme Courts.” NBER Working Paper No. 22071.
- Ash, Elliott, Sam Asher, Aditi Bhowmick, Sandeep Bhupatiraju, Daniel Chen, Tanaya Devi, Christoph Goessmann, Paul Novosad, and Bilal Siddiqi. 2021. “Measuring Gender and Religious Bias in the Indian Judiciary.” Unpublished.
- Aziz, Khursheed Kamal. 2001. *Religion, Land, and Politics in Pakistan: A Study of Piri-Muridi*. Lahore: Vanguard Books.
- Basu, Partha Pratim. 2008. “Post-Benazir Pakistan: Quo Vadis?” *Jadavpur Journal of International Relations* 11–12 (1): 317–23.
- Behrer, A. Patrick, Edward L. Glaeser, Giacomo A. M. Ponzetto, and Andrei Shleifer. 2021. “Securing Property Rights.” *Journal of Political Economy* 129 (4): 1157–92.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. “How Much Should We Trust Differences-in-Differences Estimates?” *Quarterly Journal of Economics* 119 (1): 249–75.
- Besley, Timothy, and Torsten Persson. 2009. “The Origins of State Capacity: Property Rights, Taxation, and Politics.” *American Economic Review* 99 (4): 1218–44.
- Bielen, Samantha, Ludo Peeters, Wim Marneffe, and Lode Vereeck. 2018. “Backlogs and Litigation Rates: Testing Congestion Equilibrium across European Judiciaries.” *International Review of Law and Economics* 53: 9–22.
- Boehm, Johannes. 2015. “The Impact of Contract Enforcement Costs on Outsourcing and Aggregate Productivity.” Unpublished.
- Bose, Sugata, and Ayesha Jalal. 2017. *Modern South Asia: History, Culture, Political Economy*. London: Routledge.
- Callen, Michael, Saad Gulzar, Ali Hasanain, Muhammad Yasir Khan, and Arman Rezaee. 2020. “Data and Policy Decisions: Experimental Evidence from Pakistan.” *Journal of Development Economics* 146: 102523.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. “Bootstrap-Based Improvements for Inference with Clustered Errors.” *Review of Economics and Statistics* 90 (3): 414–27.

- Chalfin, Aaron, and Justin McCrary.** 2017. "Criminal Deterrence: A Review of the Literature." *Journal of Economic Literature* 55 (1): 5–48.
- Chemin, Matthieu.** 2020. "Judicial Efficiency and Firm Productivity: Evidence from a World Database of Judicial Reforms." *Review of Economics and Statistics* 102 (1): 49–64.
- Cohen, Alma, and Crystal S. Yang.** 2019. "Judicial Politics and Sentencing Decisions." *American Economic Journal: Economic Policy* 11 (1): 160–91.
- Colonnelli, Emanuele, Mounu Prem, and Edoardo Teso.** 2020. "Patronage and Selection in Public Sector Organizations." *American Economic Review* 110 (10): 3071–99.
- Constitution of Pakistan.** 2010. "Judicial Selection Procedure." September 4. <https://pakistanconstitutionlaw.com/18th-amendment-2010/> (accessed July 18, 2019).
- Coulson, Noel J.** 2017. *A History of Islamic Law*. London: Routledge.
- Dahl, Robert.** 1957. "Decision-Making in a Democracy: The Supreme Court as a National Policy-Maker." *Journal of Public Law* 6 (1): 6–279.
- Danziger, Shai, Jonathan Levav, and Liora Avnaim-Pesso.** 2011. "Extraneous Factors in Judicial Decisions." *Proceedings of the National Academy of Sciences* 108 (17): 6889–92.
- De Montesquieu, Charles.** 1748. *Esprit des Lois*, Vol. 1. Paris: Firmin Didot Frères.
- Djankov, Simeon, Rafael La Porta, Florencio Lopez-De-Silanes, and Andrei Shleifer.** 2003. "Courts." *Quarterly Journal of Economics* 118 (2): 453–517.
- Fujiwara, Thomas, and Leonard Wantchekon.** 2013. "Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin." *American Economic Journal: Applied Economics* 5 (4): 241–55.
- Gadugah, N.** 2017. "Rawlings, 415 Past Government Officials, Institutions Grabbed State Lands." *Joy Online*, November 7. <https://www.myjoyonline.com/rawlings-415-past-govt-officials-institutions-grabbed-state-lands/>.
- Gulzar, Saad, and Muhammad Yasir Khan.** 2021. "'Good Politicians': Experimental Evidence on Motivations for Political Candidacy and Government Performance." Unpublished.
- Hayek, F.A.** 1960. *The Constitution of Liberty*. Chicago, IL: University of Chicago Press.
- Hayo, Bernd, and Stefan Voigt.** 2019. "The Long-Term Relationship between de jure and de facto Judicial Independence." *Economics Letters* 183: 108603.
- Hessami, Zohal.** 2018. "Accountability and Incentives of Appointed and Elected Public Officials." *Review of Economics and Statistics* 100 (1): 51–64.
- Heyes, Anthony, and Soodeh Saberian.** 2019. "Temperature and Decisions: Evidence from 207,000 Court Cases." *American Economic Journal: Applied Economics* 11 (2): 238–65.
- Hubbard, William H.J.** 2013. "Testing for Change in Procedural Standards, with Application to Bell Atlantic versus Twombly." *Journal of Legal Studies* 42 (1): 35–68.
- Jia, Ruixue, Masayuki Kudamatsu, and David Seim.** 2015. "Political Selection in China: The Complementary Roles of Connections and Performance." *Journal of the European Economic Association* 13 (4): 631–68.
- Jiang, Junyan.** 2018. "Making Bureaucracy Work: Patronage Networks, Performance Incentives, and Economic Development in China." *American Journal of Political Science* 62 (4): 982–99.
- Jones, Benjamin F., and Benjamin A. Olken.** 2005. "Do Leaders Matter? National Leadership and Growth since World War II." *Quarterly Journal of Economics* 120 (3): 835–64.
- Khwaja, Asim Ijaz, and Atif Mian.** 2005. "Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market." *Quarterly Journal of Economics* 120 (4): 1371–1411.
- Kureshi, Yasser.** 2021. "Selective Assertiveness and Strategic Deference: Explaining Judicial Contestation of Military Prerogatives in Pakistan." *Democratization* 28 (3): 604–24.
- La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer.** 2008. "The Economic Consequences of Legal Origins." *Journal of Economic Literature* 46 (2): 285–332.
- La Porta, Rafael, Florencio López-de-Silanes, Cristian Pop-Eleches, and Andrei Shleifer.** 2004. "Judicial Checks and Balances." *Journal of Political Economy* 112 (2): 445–70.
- Lambais, Guilherme, and Henrik Sigstad.** 2018. "Judicial Subversion: Evidence from Brazil." Unpublished.
- Lim, Claire S.H.** 2013. "Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges." *American Economic Review* 103 (4): 1360–97.
- Madison, James, Alexander Hamilton, and John Jay.** (1788) 1961. *The Federalist Papers*, edited by Clinton L. Rossiter. New York: New American Library.
- Martinez-Bravo, Monica.** 2017. "The Local Political Economy Effects of School Construction in Indonesia." *American Economic Journal: Applied Economics* 9 (2): 256–89.

- Martinez-Bravo, Monica, Priya Mukherjee, and Andreas Stegmann.** 2017. "The Non-Democratic Roots of Elite Capture: Evidence from Soeharto Mayors in Indonesia." *Econometrica* 85 (6): 1991–2010.
- Mehmood, Sultan.** 2022. "Replication Data for: The Impact of Presidential Appointment of Judges: Montesquieu or the Federalists?" American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E148001V1>.
- Mehmood, Sultan, Avner Seror, and Daniel L. Chen.** 2022. "Religious Rituals: Evidence from Ramadan." Unpublished.
- Montiel Olea, José Luis, and Carolin Pflueger.** 2013. "A Robust Test for Weak Instruments." *Journal of Business and Economic Statistics* 31 (3): 358–69.
- Naseer, Tahir.** 2019. "IHC Suspends Decision to Cancel JUI-F Leader Hafiz Hamid Ullah's Citizenship." *Dawn*, October 29. <https://www.dawn.com/news/1513576>.
- News.** 2018. "Complete IHC Order of Khawaja Asif Disqualification Case." *News*, April 26. <https://www.thenews.com.pk/latest/309427-complete-ihc-order-of-khawaja-asif-disqualification-case>.
- Ng, Edmond S.-W., Richard Grieve, and James R. Carpenter.** 2013. "Two-Stage Nonparametric Bootstrap Sampling with Shrinkage Correction for Clustered Data." *Stata Journal* 13 (1): 141–64.
- North, Douglas Cecil.** 1990. *Institutions, Institutional Change, and Economic Performance*. New York: Cambridge University Press.
- Palumbo, Giuliana, Giulia Giupponi, Luca Nunziata, and Juan S. Mora Sanguinetti.** 2013. "The Economics of Civil Justice: New Cross-Country Data and Empirics." Unpublished.
- Perlez, Jane, and Carlotta Gall.** 2008. "In Pakistan, Musharraf's Party Accepts Defeat." *New York Times*, February 20. <https://www.nytimes.com/2008/02/20/world/asia/20pakistan.html>.
- Platteau, Jean-Philippe.** 2017. *Islam Instrumentalized: Religion and Politics in Historical Perspective*. New York: Cambridge University Press.
- Ponticelli, Jacopo, and Leonardo S. Alencar.** 2016. "Court Enforcement, Bank Loans, and Firm Investment: Evidence from a Bankruptcy Reform in Brazil." *Quarterly Journal of Economics* 131 (3): 1365–1413.
- Pound, Roscoe.** 1998. *The Spirit of the Common Law*. New Brunswick: Transaction Publishers.
- Priest, George L., and Benjamin Klein.** 1984. "The Selection of Disputes for Litigation." *Journal of Legal Studies* 13 (1): 1–55.
- Registrar Office.** 2017. "List of Judges." Supreme Court of Pakistan Record Room.
- Registrar Office eCourt Case Data Supreme Court of Pakistan.** 2020. "Case Archives and eCourts Case Data." Lahore Registry.
- Registrar Offices, High Courts of Pakistan.** 2020. "Jurisdictional Maps of High Court Principal and Non-Principal Benches." Islamabad.
- Rehavi, M. Marit, and Sonja B. Starr.** 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122 (6): 1320–54.
- Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi.** 2004. "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." *Journal of Economic Growth* 9 (2): 131–65.
- Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb.** 2019. "Fast and Wild: Bootstrap Inference in Stata Using Boottest." *Stata Journal* 19 (1): 4–60.
- Sattar, Babar.** 2012. "18th Constitutional Amendment and Need for Passage of the 19th Constitutional Amendment." In *Eighteenth Amendment Revisited*, edited by Maqsoodul Hasan Nuri, Muhammad Hanif, and Muhammad Nawaz Khan, 74–87. Islamabad Policy Research Institute.
- Sattar, Babar.** 2017. "Legal Eye: Daylight Robbery." *News*, October 28. <https://www.thenews.com.pk/print/240307-Legal-eye-Daylight-robbery>.
- Shayo, Moses, and Asaf Zussman.** 2011. "Judicial Ingroup Bias in the Shadow of Terrorism." *Quarterly Journal of Economics* 126 (3): 1447–84.
- Shleifer, Andrei, and Robert W. Vishny.** 2002. *The Grabbing Hand: Government Pathologies and Their Cures*. Cambridge, MA: Harvard University Press.
- Siddique, Osama.** 2013. *Pakistan's Experience with Formal Law: An Alien Justice*. Cambridge, MA: Cambridge University Press.
- Silveira, Bernardo S.** 2017. "Bargaining with Asymmetric Information: An Empirical Study of Plea Negotiations." *Econometrica* 85 (2): 419–52.
- Somin, Ilya.** 2016. *The Grasping Hand: "Kelo versus City of New London" and the Limits of Eminent Domain*. Chicago, IL: University of Chicago Press.

- Stock, James H., and Motohiro Yogo.** 2005. "Asymptotic Distributions of Instrumental Variables Statistics with Many Instruments." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Vol. 6, edited by Donald W.K. Andrews and James H. Stock, 109–120. New York: Cambridge University Press.
- Tavernis, Sabrina, and Salman Masood.** 2010. "Pakistan Weighs Changes to Revise Constitution." *New York Times*, April 6. <https://www.nytimes.com/2010/04/07/world/asia/07pstan.html>.
- TechJuice.** 2020. "These Are the 12 Richest and Biggest Business Owners in Pakistan." *TechJuice*, September 26. <https://www.techjuice.pk/12-richest-and-biggest-business-owners-in-pakistan/>.
- Tidmarsh, Jay.** 2009. "Resolving Cases on the Merits." *Denver University Law Review* 87 (2): 407–36.
- Times of India.** 2018. "Union Minister Giriraj Singh Booked in Land-Grab Case." *Times of India*, February 9. <https://timesofindia.indiatimes.com/city/patna/giriraj-booked-in-land-grab-case/articleshow/62841673.cms>.
- Vannutelli, Silvia.** 2021. "From Lapdogs to Watchdogs: Random Auditor Assignment and Municipal Fiscal Performance in Italy." Unpublished.
- World Justice Project.** 2020. *World Justice Project Rule of Law Index 2020*. Washington, DC: World Justice Project.
- Xu, Guo.** 2018. "The Costs of Patronage: Evidence from the British Empire." *American Economic Review* 108 (11): 3170–98.
- Zafar, Ali.** 2012. "How True Judicial Independence Can Be Achieved." *Express Tribune*, April 30. <https://tribune.com.pk/story/372183/how-true-judicial-independence-can-be-achieved>.
- Zafar, Syed Mohammad.** 2012. "Inaugural Address on 18th Constitutional." *IPRI Policy Brief*: 4.

This article has been cited by:

1. Hantao Hao, Linyi Zheng. 2024. Does land expropriation to neighbors affect the enrollment of bystanders in pension programs?. *Economic Analysis and Policy* **84**, 576-588. [[Crossref](#)]
2. Abhay Aneja, Guo Xu. 2024. Strengthening State Capacity: Civil Service Reform and Public Sector Performance during the Gilded Age. *American Economic Review* **114**:8, 2352-2387. [[Abstract](#)] [[View PDF article](#)] [[PDF with links](#)]
3. Da Zhao, Jingyuan Guo, Shule Yu, Litian Yu. 2024. Tradeoff between local protection and public sector performance: Lessons from judicial fiscal centralization. *Journal of Economic Behavior & Organization* **220**, 254-278. [[Crossref](#)]
4. Peng Zhang. 2023. Anti-corruption campaign, political connections, and court bias: Evidence from Chinese corporate lawsuits. *Journal of Public Economics* **222**, 104861. [[Crossref](#)]