Reform Multiplier and Elite Entrenchment: Evidence from Pakistan

BY SULTAN MEHMOOD AND BAKHTAWAR ALI[1*]

June 2024

How and when does a reform trigger a cascade? How does this reshape the composition of the elite? Addressing these questions, this study documents the dynamic effect of a judicial reform in Pakistan that transitioned judge appointments from the executive to a judicial committee, introducing a novel conceptual perspective on reform durability. Our findings reveal that the sustainability of judicial reforms is linked to the selection mechanisms of appointing authorities rather than the appointment of individual officials. The evidence suggests a multiplier effect on judicial independence as committee-appointed judges are succeeded by their appointees, highlighting the importance of the selection of selectors in maintaining an institutional equilibrium. While the reform generates a cascading effect of increased judicial autonomy and a more concentrated judicial elite, it also aligns with enhanced decision-making quality, with increased recruitment of second-generation judges from premier law schools and increased rulings on case merits. Our findings contribute to understanding why some reforms, such as Pakistan's judicial appointment reform, become embedded while others remain transient, a phenomenon that persists despite increasing elite concentration.

Keywords: reform compounding, elite concentration.

[1*] Mehmood: New Economic School, Harvard Law School, e-mail: smehmood@nes.ru. Ali: Aix Marseille School of Economics, email: bakhtawar.ali@univ-amu.fr. We thank seminar participants at Northwestern, the Federal Reserve in Chicago, Northwestern, Manchester, the CEPR Annual Conference, the New Economic School, the University of Oxford, and the ETH Zurich for their comments and feedback. Financial support from the Chair Governance and Regulation Lab, DIAL, University of Paris Dauphine is acknowledged. This work was also supported by the French National Research Agency Grants ANR-17-EUR-0010.

1. Introduction

The landscape of institutional reforms is marked by a blend of permanence and transience. While certain reforms embed themselves into the fabric of the State, becoming virtually irreversible, others are ephemeral, succumbing to the forces of reversal (Acemoglu and Robinson, 2019). This dichotomy raises critical questions about the nature of institutional change: Why are some reforms sustained while others are prone to reversal? How do they alter the structure of the elite?

At the core of institutional transformations are judicial reforms, a linchpin for the efficacy of many other reforms, ranging from financial regulations to the enforcement of property rights and civil service reforms (Ponticelli and Alencar, 2016; Boehm and Oberfield, 2020; Liu et al., 2023). Yet, judicial reforms, often intended to mitigate governmental excesses, may paradoxically entrench elite dominance, reinforcing the power structures they seek to dismantle or pave the way for a new elite cadre to emerge.

This paper provides evidence of a ripple effect stemming from a watershed judicial reform in Pakistan, which introduced a judicial committee to appoint judges. Before the 2010 reform, the executive branch (President and the Prime Minister) appointed judges in Pakistan. The judge selection reform in 2010 marked a significant departure as it entrusted the task of making judicial appointments to a committee of senior judges (Mehmood, 2022). The selection reform, however, unfolded in two stages. Initially, first-generation judges are appointed by a committee of executive-appointed judges. Subsequently, first-generation committee-appointed judges secure a majority in the judicial selection committee, effectively controlling second-generation appointments. Consequently, there is a two-degree separation from the government in the judicial appointment process. We assess the multiplier effect of transitioning from first to second-generation committee-appointed judges, focusing on its repercussions for judicial autonomy and the implications of judges taking the reins in appointing their peers on the composition and concentration of the judicial elite. We identify this multiplier effect "at scale" (Muralidharan and Niehaus, 2017) as the reform propagates nationwide across Pakistan's constitutional courts.

The judge selection reform, which we focus on, provides an ideal context for studying the reform cascade effect on judicial independence and elite concentration. Firstly, the reform underwent a natural mutation, establishing a larger demarcation from executive appointments in two stages. In the first stage, the succession of judges unfolds with the first-generation judges (first-degree separation from the government) succeeding judges appointed by the executive branch, followed by the second stage, where the second-generation committee appointees (second-degree separation from the government) replace the first-generation appointees. This separation both widens the gap between executive-appointed judges and escalates the potential for judicial self-entrenchment. Secondly, our access to

case-level microdata over a decade following the reform's implementation enables us to effectively address a persistent measurement challenge in assessing the longer-term impact of judicial reforms (Chemin, 2021). Thirdly, the nationwide jurisdiction of Pakistan's constitutional courts, offering comparable measures across diverse regions and periods, enables an examination of judicial autonomy in the context of a population exceeding 240 million, providing insights into the governance of the world's fifth most populous country.[2]

Underpinning our study is a large-scale digitization of court records in Pakistan. To begin with, we construct outcome variables using judgment texts. These case-level microdata are utilized to build measures of anti-government rulings, merit-based decisions, and adherence to due process of the law. Following the influential work of Djankov et al. (2003), we contract out legal experts (two independent panels of attorneys) to code these variables.[3] We further harness judicial administrative data from the 'registrar offices' of High Courts and judicial records at the Law and Justice Commission of Pakistan (LJCP) to construct measures of concentration in the judiciary, enabling us to map out a network of the judicial elite.[4] We integrate these data with details on judges' promotions, both prior to and after their service, from Judicial Gazettes of Pakistan, which we also digitized. This integration empowers us to trace the entire career trajectory of judges and facilitate connecting judges to their previous employment at law firms when they served as attorneys, identifying the law school they attended, and tracking their post-employment positions at the Supreme Court if they were promoted—something that eluded prior scholarly work.[5]

One challenge in studying public sector reforms, such as the judge selection reform in Pakistan, is that such reforms are not randomly assigned but reflect deliberate policy choices (Besley and Case, 2000; Ujhelyi, 2014; Liu et al., 2022). This non-random assignment raises concerns regarding the potential correlation between the adoption of judge selection reform and court-specific factors that could impact judicial decisions. Our study, however, leverages a unique setting. Despite the selection reform being federally mandated through a constitutional amendment, the arrival of first and second-generation committee appointees depends on the availability of vacancies at the regional high court's

[2] Pakistan's judiciary, mirroring those in India and Bangladesh, reflects a shared colonial legacy in its procedural laws, statutes, legal hierarchy, and protocols. Together, these nations comprise over a quarter of the world's population, rendering the insights from this study pertinent to a substantial demographic expanse.

[3] A comparable database, coded by legal professionals, is available for the United States Supreme Court decisions, known as The Supreme Court Database (http://scdb.wustl.edu). This resource has been extensively utilized in scholarly work, see, for instance, Spaeth et al. (2014).

[4] Recent scholarship, such as Mehmood (2022) and Liu et al. (2023), has begun to explore how judicial reforms influence decision-making through the analysis of judgment texts. However, there appears to be a stark gap in the literature regarding the effects of these reforms on the elite power structure, as the reforms themselves can alter the composition and structure of the judicial elite.

[5] The career progression of High Court judges in Pakistan follows a well-defined trajectory: a lawyer with at least ten years of experience may be elevated to a High Court judgeship, and approximately 5% of these judges are subsequently elevated to the Supreme Court upon reaching the mandatory retirement age of 62. The remaining judges are required to observe a two-year moratorium on employment in all sectors following the conclusion of their judicial service.

judicial benches. These vacancies arise and get filled one-to-one in a near-deterministic fashion, primarily driven by the mandatory retirement age law that allows us to leverage appointment variation within a regional court at the judicial bench level.[6] This approach enhances the difference-in-differences design, as it permits us to incorporate court-by-bench fixed effects, year-fixed effects, and court-by-bench-specific linear trends.[7] In this demanding design, we can explicitly account for some time-varying omitted variables by conducting within-court off-trend comparisons.[8]

Under the parallel trends assumption, this generalized difference-in-differences design allows us to address several empirical issues that could otherwise hinder our ability to interpret the results as causal. For instance, by incorporating court-by-bench (or district-by-bench) fixed effects, we can account for time-invariant court-by-bench specific effects such as the district's customary laws. Concurrently, including year-fixed effects allows us to address the impact of shared factors that influence all courts over time, such as national elections or the tenure of a specific Chief Justice. We also incorporate court-by-bench-specific linear trends to capture some time-varying omitted factors, such as the evolution of legal precedents or the fluctuating levels of local court congestion. Since case precedents and court congestion evolve slowly over time, they may be factored into our analysis by including these court-by-bench-specific trends.

While this difference-in-differences methodology with court-by-bench-specific linear trends helps address some time-variant sources of endogeneity, we further leverage our institutional knowledge of Pakistani High Courts to exploit plausibly random variation in new judicial appointments. First, we leverage the additional effect of new judicial appointments that emerges from changes in the composition of the judge hiring panel, bringing the committee appointees in the hiring panel close to a majority with a 3:4 ratio in the judge hiring panel versus the generational shift occurring when the committee-appointed judges attain a 4:3 majority within the judge hiring panel. In this research design, we can compare the judicial decisions of newly appointed judges chosen by the judge hiring committee as it approaches versus attains a majority of committee-appointed judges. By assessing the judicial decisions made by judges appointed during this pivotal change, we pinpoint what we denote as the "reform multiplier effect." Second, we utilize Pakistan's law mandating retirement for High Court judges at age 62. By using the mandatory retirements on their 62nd birthdays as a source of quasi-random variation in new appointments, we demonstrate that the reform multiplier effect emerges only when these retirements coincide with the generational transition from first to second-generation

[6] Within each region, a judicial bench constitutes a fixed panel of judges who are arranged according to their specialized fields, for example, the tax bench at the Lahore High Court, with an average of 10 judges per judicial bench.

[7] In our study, court-by-bench fixed effects are synonymous with district-by-bench fixed effects since each region hosts a single "divisional" High Court.

[8] The similarity of our coefficient estimates, both with and without the inclusion of court-by-bench specific linear trends, nevertheless, points to the likelihood that our treatment is likely independent of these district-specific linear trends.

committee appointments (specifically, when the retirement shifts the composition from a 3:4 minority to a 4:3 majority of committee-appointed judges on the judge hiring panel).

Pro-Government Rulings. We begin to ascertain the reform multiplier effect by comparing pro-government rulings between the first and second generations of committee appointees. We analyze whether there is a statistical difference in pro-government rulings when the judge hiring committee gains a majority of committee-appointed judges versus when it is in the minority. The key finding from this exercise is that there is a notable reform multiplier effect. Specifically, the first degree of separation in judicial appointments from the government leads to a decrease of 10 percentage points in pro-government rulings, while the impact intensifies further with the second degree of separation from the government, resulting in a decline of approximately 30 percentage points in pro-government rulings, relative to decisions by government-appointed judges. The difference in the coefficient estimates is both statistically and qualitatively significant, amounting to a 20-percentage point reduction, which represents the reform multiplier effect. This fall corresponds to a 50% decrease in pro-government rulings compared to the sample mean. To put these estimates into perspective, the magnitude of this reform's compounding effect alone is roughly equal to the consequences of a 2011 judicial reform in the United States, as analyzed by Rose (2021), and it is nearly five times the effect of a 2014 judicial reform in China, as documented in important new work by Liu et al. (2023).[9]

Elite Entrenchment. The reform multiplier effect, nevertheless, comes with a substantial rise in concentration among the judicial elite. As the second generation of committee appointees replaces the first generation of committee-appointed judges, there is a 10% increase in the appointment of ex-colleagues from law firms that once employed members of the judge-hiring panel as attorneys. We use the Herfindahl-Hirschman Index (HHI) to benchmark the magnitude of the concentration of judges appointed from the same law firms as the judge appointment committee. According to the United States Department of Justice, HHI values below 1500 indicate low concentration, while values between 1500 and 2500 suggest moderate concentration; HHI values above 2500 indicate high concentration. The Herfindahl-Hirschman Index (HHI) for the first-generation committee appointees stands at approximately 1400, or 0.14, on a scale from 0 to 1, signifying a low level of concentration. In contrast, the second-generation committee appointees have an HHI close to 4800 (0.48 on a 0-1 scale), reflecting a high concentration of judges with appointments predominantly from a select few law firms. This suggests that the increased independence of the judiciary from the government may have come at the expense of a highly concentrated judicial elite. Judges might be consolidating their influence as the

[9] Subsequently, we will investigate whether these findings stem from incentive effects, with the rulings of first-generation judges aligning more with their second-generation peers, or from judge selection effects, reflecting differences in characteristics between the two generations of judges. The evidence is consistent with the latter selection effects mechanism.

separation from executive appointments increases, hiring more assertively from their established network of ex-colleagues.

Decision Quality. The effect of this increased concentration of judicial elites on the quality of judicial decisions is, a priori, ambiguous. On the one hand, a greater concentration of judges from the same law firm as the judge hiring panel may result in the selection of less-qualified or 'patronage' judges, leading to a deterioration in decision quality. On the other hand, a higher concentration of judges from the same law firms could enhance the quality of judicial decisions if these hires are also more qualified. The evidence is consistent with the latter explanation. Despite increased elite concentration, judicial decisions by second-generation judges are perceived by legal experts to be of higher quality compared to those of their first-generation counterparts. Moreover, second-generation judges are more likely to hail from top-ranked law schools in Pakistan than first-generation committee-appointed judges, suggesting that judges from the same law firm could be better legal professionals with top-tier educational credentials. We assess decision quality using two measurements, with the first relying on a panel of attorneys' evaluations regarding whether a case is decided on merits and evidence rather than legal technicalities. This motivation stems from legal literature and common law jurisprudence, which holds that decisions on case merits encapsulate judgments based on the spirit of the law and are considered to exhibit superior legal reasoning (see, e.g., Pound, 1963; Tidmarsh, 2009; Siddique, 2013).[10] Furthermore, empirical legal research suggests that a greater emphasis on legal technicalities rather than the substantive merits of cases typically indicates a decline in decision quality, as rulings on technicalities tend to depend on arbitrary "discretionary rules" rather than on well-founded legal principles (Liu and Li, 2019; Wang et al., 2022). We also use a due process rating on judgments, scored on a one-to-five scale by lawyers, as our second measure of judicial quality. This rating evaluates the extent to which all jurisdictional, procedural, and evidential prerequisites are duly considered in the decision, aligning with the widely applied legal principle that justice should be perceived through an assessment of due process followed, not just the outcome of the judicial decision (often attributed to the landmark judgment of McCarthy v. Sussex Justices, 1924). The transition from the first to the second generation of committee appointees increased rulings based on case merits and a heightened observance of due process of law, suggesting a similar multiplier effect in the quality of judicial decisions. This holds irrespective of whether the judges shared the same law firm affiliation

[10] In particular, the "Merit" variable is a binary indicator that takes a value of one when the decision is determined to be "based on evidence or case merits" and zero if it is based on a technicality. There are two primary reasons for constructing this variable. Firstly, legal scholarship in Pakistan suggests that ruling on technicalities is often used as a means to favor state authorities unjustly, and such decisions are considered indicative of biased judgments (see, e.g., Siddique, 2013). Secondly, this variable aligns with the principles of common law jurisprudence, which aims to base rulings on the case's merits, considering evidence and the spirit of the law rather than relying on legal technicalities. This ideal is extensively used in common law jurisprudence and is discussed at length in Pound (1963) and summarized in Tidmarsh (2009).

with the judge hiring panel or not. A reform multiplier effect of improved decision quality is observed with each degree of separation from executive appointments.

Threats to Identification. A key threat to identification is the possibility that we are capturing the effect of the arrival of new judges per se rather than the reform compounding effect. In other words, the reform compounding effect marked by a significant decline in judicial decisions when the judge hiring committee achieves a majority of committee-appointed judges would manifest with the arrival of any new judge, irrespective of the composition of the judge hiring panel that selected them. We provide two pieces of evidence that undermine this conjecture. First, we show that nearing a majority with committee-appointed judges does not have a discernible influence. In particular, we find no evidence of reform compounding when a new committee-appointed judge joins the judge hiring panel, bringing, for example, the committee appointees in the judge hiring panel close to a majority with a 3:4 ratio. However, we observe a substantial and statistically significant reform compounding effect when the committee-appointed judges attain a 4:3 majority within the judge hiring panel, underscoring how judge selection fundamentally dictates their group conduct. Second, we use a mandatory retirement age law that creates vacancies depending on the age structure of judges on the bench. The reform multiplier effect is solely activated with mandatory retirements of judges in transitional years, when the first 'next generation' judge is appointed at the bench, in contrast to non-transitional years when the same selection regime persists.

Mechanisms. Having documented the reform multiplier effect as the reform mutates from the first to the second degree of separation from government appointments, we next tentatively explore the potential underlying mechanisms. As professional bureaucracies, judiciaries often feature fixed retirement ages, limited discretionary firing, and a lack of performance pay. These responses to past patronage systems in public bureaucracies imply that mechanisms such as career incentives and selection effects are likely more important in directing the conduct of public institutions like the judiciary (as suggested in Finan et al., 2017; Xu, 2018 or Leucht, 2023). Building on the insights from existing literature and novel data derived from the digitization of court records in Pakistan, we can explore seven potential mechanisms that may explain the reform compounding. These possible mechanisms are (1) selection effects (Besley, 2005), (2) judge ideology (Bonica and Sen, 2021), (3) career incentives (Bertrand et al., 2020), (4) spillover or peer effects (Sunstein et al., 2007), (5) precedent compliance (Chen, 2023), (6) strategic response of litigants (Yang, 2016; Acemoglu et al., 2020) and (7) corporate or lawyers' capture (Viktorov, 2019; Khan, 2019). All, any, or none of these mechanisms may explain the reform compounding effect. The evidence supports the selection mechanism, revealing that second-generation committee appointees are more experienced attorneys, educated in top-ranked law schools, compared to their first-generation counterparts. More formally, we utilize causal mediation analysis to document that 40% of the reform multiplier effect is mediated

by hiring judges from elite law schools (Heckman and Pinto, 2015). Alternative channels of judge ideology, peer effects of the arrival of second-generation judges on first-generation judges, greater compliance with case precedents, changes in case filings, or corporate and lawyers' capture do not statistically appear to mediate the reform multiplier effect.

Case Composition. A common concern in studies using judicial data is "judge shopping," where litigants may select judges (Klien and Priest, 1984). However, our study offers several pieces of evidence that temper such apprehensions. Firstly, the average seven-year congestion in Pakistani courts minimizes the immediate effects of litigant-driven actions on judicial decisions, as judges' case dockets are slow to respond to changes in case filings (LJPC, 2022; Dunya News, 2023). Secondly, leveraging the random case assignment mechanism within the High Court facilitates an inter-regime comparison of judicial decisions. Forum shopping is prohibited in Pakistan, as established by settled jurisprudence, for example, in "Sharaf Faridi v. Federation of Pakistan" (PLD 1989 Karachi 404). The results from baseline case-level and judge-level regressions, where the latter directly leverages the random assignment, are essentially identical. Lastly, a balance test over a long list of case and bench characteristics indicates that first- and second-generation committee appointees adjudicate over similar cases, consistent with random assignment. Together, the evidence strongly suggests that the changes in case composition do not elucidate the observed patterns in the data.

Parallel Trends and Heterogeneous Treatment. The difference-in-differences identification relies on the parallel trends assumption, and standard methods to estimate the two-way fixed effect model assume homogeneous treatment effects. We discuss and analyze each of these issues in turn. First, we estimate a fully dynamic two-way fixed effect model: the year-by-year event study estimates reveal no evidence of differential pre-trends and suggest that we are not confounding the reform multiplier effect with slower-moving omitted factors. In other words, in the years leading up to the generational shift from the first to the second generation of committee appointees, we do not observe any reform multiplier effect. Second, we investigate differential trends prior to the reform using alternative estimators introduced by Borusyak, Jaravel, and Spiess (2021), Callaway and Sant'Anna (2021), and Sun and Abraham (2021) that do not rely on the relatively restrictive homogeneous treatment effects assumption. Results from these estimators also do not contradict differential pre-trends. Third, to the extent that the differential trends may confound our results at the court-by-year level, the results are robust when we include court-by-year fixed effects. This allows us to account for the omitted factors that change in the district over time, such as local governments or local court congestion. Lastly, we speak to negative weighting bias in TWFE models arising due to potentially heterogeneous treatment effects across units and time by demonstrating the robustness of our results to estimators that do not hinge on the homogeneity assumption. These estimators provide consistent estimates by comparing newly treated and already treated units using 2x2 difference-in-differences

comparisons, as proposed by Borusyak, Jaravel, and Spiess (2021), Callaway and Sant'Anna (2021), Goodman-Bacon (2021), and Sun and Abraham (2021). Across these more robust estimators, the main results remain qualitatively and statistically similar.

1.1. Related Literature

Our study makes contributions to multiple strands of literature. Firstly, this paper advances the scholarly dialogue on the political economy of reforms, drawing upon the foundational contributions of Alesina et al. (2006) and subsequent literature, including Pastor and Veronesi (2012), Acemoglu et al. (2015), and Rodrik (2021). Whereas much of the current literature focuses on the political constraints to policy implementation, as explored by Mian et al. (2014) and Besley and Persson (2014), our study extends the investigation to reforms that limit political influence in bureaucratic decision-making and also offers empirical evidence into the direct impacts and inherent trade-offs associated with such institutional reforms. This focus aligns with studies on conditions that "constrain executive power" (Acemoglu and Robinson, 2020; Acemoglu and Wolitzky, 2021), facilitate the emergence of the "rule of law" (North, Weingast, and Wallis, 2009), and promote "legal order" (Hadfield and Weingast, 2012).[11] Despite the significant theoretical strides made in this literature, it is surprising to note the scarcity of systematic empirical evidence supporting reforms that could bolster the rule of law, as well as studies of the trade-offs that such reforms entail. It is even rarer to empirically investigate the trade-offs associated with institutional reforms. This paper endeavors to fill this gap by introducing novel evidence that demonstrates a ripple effect originating from a significant institutional reform in Pakistan, which has implications for both judicial independence and the concentration of elite power. Specifically, we document how a judge selection reform, which has been investigated for its short-run effects on pro-government rulings in Mehmood (2022) has a deep impact on concentration of the judicial elite. This research diverges from previous work by posing a novel question—when may a reform stick?—and provide a new conceptual answer regarding the sustainability of reforms: the enduring nature of judicial reforms hinges not so much on the appointment of individual officials but rather on the process by which these appointing authorities, or 'selectors,' are chosen. The selection process of these 'selectors' may be crucial in maintaining and sustaining institutional equilibrium.

[11] This theoretical literature in question places a strong emphasis on the central importance of judicial independence, as well as the need for a clear division of powers between the judiciary and the government. (Smith, 1776; Hamilton, 1788; Hayek, 1960; Buchanan, 1974; North, 1986; Glaeser and Shleifer, 2002; La Porta et al., 2004; Gennaioli, and Shleifer, 2007; Baker and Mezzetti, 2012; Acemoglu et al. 2013), there remains a dearth of rigorous empirical evidence on how we can effectively increase judicial independence, especially in weakly institutionalized settings. Notable exceptions include Liu, Lu, Peng, and Wang (2023) studying China and Ponticelli and Alencar (2016) exploring Brazil. Chemin's (2021) important study is also relevant to our examination, as it investigates 183 judicial reforms across 36 countries over the last fifteen years. The findings indicate that judicial reforms that bolstered the separation of powers between the government and the judiciary have increased citizen confidence in the State institutions, and there is a need for more well-identified studies studying the long-run impact of the judicial reforms (also related to this work is Chemin (2020)'s cross-country exploration of the effect of judicial reforms with firm productivity across 74 countries).

Moreover, unlike prior work, the novel administrative data on judges' previous employment, the law schools they attended, and their career trajectories, including whether they were promoted to the Supreme Court, allows us to extend prior scholarship to (1) assess the longer-term impact of the judicial reform and statistically test for the reform multiplier effect, answering a pivotal question whether reforms can amplify (2) showcase the trade-off between greater judicial independence and increased concentration of appointments within judge networks. To the best of our knowledge, this paper provides the first causal evidence of how judicial reforms may have multifaceted effects, simultaneously enhancing judicial independence from the government, creating a cascade that makes it stick and reshaping the composition of the judicial elite. We find, contrary to our priors, that the amplification of elite concentration does not compromise the quality of decisions rendered. A reform multiplier effect is observed both for decision quality and elite concentration.

Secondly, our paper contributes to the state capacity literature (Besley and Persson, 2009; Acemoglu et al., 2015). While there has been research on the effects of state capacity across countries and regions (Dincecco and Katz, 2016; Dell et al., 2018; Besley et al., 2021) and the role of civil service reforms in enhancing state capacity (Moreira and Pérez, 2021; Vannutelli, 2022; Aneja and Xu, 2023; Leucht, 2023; Bandiera et al., 2023), there is relatively limited evidence of the link between judicial reforms and state capacity. Indeed, a key component of state capacity is "legal capacity" (Besley and Persson, 2009, p. 1218), with an independent judiciary facilitating the emergence of an "effective state" (Bandiera et al., 2021, p. 49). Thus, rigorous research on reforms that bolster judicial effectiveness is of first-order importance. However, judicial reforms are often implemented across entire administrations, so a significant empirical challenge is the lack of counterfactuals (Aberra and Chemin, 2018; Chemin, 2021; Aney et al., 2021; Liu et al., 2022). Most of existing research has relied on cross-country comparisons (Djankov et al., 2003; Voigt et al., 2015), with well-identified studies in this area predominantly focusing on the roles of civil servants (Bertrand et al., 2020; Fenizia, 2022) and fiscal capacity (Moreira and Perez, 2021; Mocanu, 2022; Khan et al., 2016), often at the expense of exploring judicial capacity. A fundamental feature of the judiciary is its intended independence, setting it apart from other government bodies such as the civil service or the legislature, which, to varying degrees, remain answerable to the executive branch, even within well-established democracies. Unlike the legislature, which often remains linked to the executive, the judiciary is supposed to function with a considerable degree of autonomy (Boston and Silveira, 2019; Lambais and Sigstad, 2023). The autonomy of the judiciary positions it as an essential counterbalance to the other government branches, ensuring its pivotal role in the execution of policy reforms. Due to this distinctive role, the study of the judiciary holds significant importance. We contribute to this literature by empirically investigating how increasing the separation of judges from government appointments may enhance judicial capacity by making the selection process more autonomous and merit-based.

Our results and analysis of mechanisms underscore the longer-run significance of "selection and recruitment," as emphasized in the comprehensive review by Finan et al. (2017, p. 469). Our study's central contribution to this literature lies in revealing the potential for judge selection to trigger a multiplier effect on legal capacity.

Finally, the paper speaks to the broader literature on institutions and development, with a particular focus on studies that underscore the significance of maintaining checks and balances on executive power (Acemoglu et al., 2001; Shleifer and Vishney, 2002; La Porta et al., 2003) and role of policy in changing the political equilibrium (Gennaioli, and Shleifer, 2007; Acemoglu et al., 2008; Alesina and Romer, 2008; Jain et al., 2014). Notably, North (1990) offers a concise definition of institutions as "humanly devised constraints that shape human interaction," a concept that Acemoglu, Johnson, and Robinson (2001) draw upon in their influential study to motivate the use of "constraints on the executive" as their central measure of long-run institutional development. Our contribution to this influential literature is in demonstrating how judicial reforms increasing "constraints on the executive" (Acemoglu et al., 2001, p. 1394) can exert a cascading influence on institutional development. In doing so, our study empirically validates the theories and mechanisms elucidated in many of these seminal investigations.

The remainder of this paper is organized as follows: Section II provides an overview of Pakistan's judicial structure and the institutional context pertinent to the selection reform. Section III delineates the data employed in our analysis. In Section IV, we explicate the empirical methodology. Section V articulates the main findings and addresses concerns regarding internal validity. Section VI probes into the mechanisms that underlie our main findings, examining both corroborative and alternative explanations. Section VIII deliberates on the implications of our main results and proffers concluding remarks. Appendices A and B, available in the online appendix, contain additional details on data construction, variable descriptions, alternative explanations, robustness checks, and supplementary figures and tables.

2. Institutional Background

Pakistan's Judicial Structure and the High Courts. Pakistan's court system is structured into three tiers. The first tier includes Civil and Sessions Courts, which preside over civil and criminal cases. The second tier includes the constitutional High Courts, with benches across all regions of the country, possessing broad constitutional authority, including adjudicating cases against government abuses of power and infringements of the constitutional rights of citizens. Last, the Supreme Court of Pakistan, located in Islamabad, holds the highest appellate jurisdiction. This apex court can have a maximum of 17 judges and handles a limited number of cases. Our study focuses on the High Courts of Pakistan. These constitutional courts serve as pivotal institutions within Pakistan's legal framework, encompassing an array of functions crucial to upholding the rule of law and administering justice. The

decisions in these courts have the power of law (Siddique, 2013). These constitutional courts are tasked with preventing the misuse of governmental authority (under "writ jurisdiction") and safeguarding political rights (under "original jurisdiction"), it is imperative that these courts maintain a robust independence from governmental influence to execute their duties effectively. In cases where the government infringes upon citizens' property or violates political freedoms, these High Courts serve as the primary, and often sole, recourse for citizens seeking redress (Siddique, 2013). Typical cases in these courts revolve around the government's persecution of opposition politicians, redress for abuses of power by law enforcement authorities, election irregularities, harassment of journalists, and the expropriation of private property by the government. The High Courts have the jurisdiction to reverse government decisions and to fine, remove, replace, or even imprison public officials who overstep their authority. There are four provincial high courts and one federal high court, with each High Court further subdivided into four regional or divisional High Courts. Each divisional high court has four specialized judicial benches (panels of judges), resulting in 64 divisional benches dispersed across Pakistan. In this manner, all the High Courts collectively exercise jurisdiction across Pakistan. The judges in these 64 judicial benches are assigned to cases according to a "roster of sittings," which designates their specialties within the regional courts. The Chief Justice of the State is responsible for the assignment of judges to specific benches by random assignment. While a random rotation schedule typically dictates judicial assignments, the Chief Justice of the High Court may, in 'exceptional' situations, reassign judges—a practice advised against except under extraordinary circumstances. This approach is akin to the protocols observed in Indian High Courts, as explored recently by Chandra et al. (2023).

Judicial Selection Reform in Pakistan. The judicial selection reform in Pakistan was implemented following the enactment of the 19th Amendment to the Constitution of Pakistan. Prior to this amendment, the appointment process was relatively opaque, with the President and the Prime Minister appointing judges, which resulted in a system where the executive branch had a dominant role in judicial appointments. The 19th Amendment established the Judicial Commission of Pakistan, comprising a committee of judges that includes the Chief Justice of Pakistan and other senior judges, who would appoint judges through majority voting. This shift transferred the power of judicial selection from the executive branch to a committee of serving judges.[12] Specifically, the 19th Amendment modified Article 175A of the Constitution of Pakistan, which governed the process of selecting judges. The modified Article 175A reads: "*(1) There shall be a Judicial Commission of*

[12] Originally, the 18th amendment to Pakistan's constitution was the first attempt to shift the authority for appointing judges from the executive branch to the Judicial Commission, led by senior judges and the Chief Justice of Pakistan (along with provisions to grant more provincial autonomy). However, it included a provision allowing a parliamentary committee to overrule the judicial committee's decisions. Following legal challenges to the 18th Amendment, the 19th Amendment was introduced, removing the legislative committee's veto power in the judicial selection process.

Pakistan, hereinafter in this Article referred to as the Commission, for appointment of Judges of the Supreme Court, High Courts and the Federal Shariat Court, as hereinafter provided. (2) For appointment of Judges of the Supreme Court, the Commission shall consist of (i) Chief Justice of Pakistan; (ii) four most senior Judges of the Supreme Court;

For appointment of Judges of a High Court, the Commission in clause (2) shall also include the following, namely: (i) Member Chief Justice of the High Court to which the appointment is being made; (ii) Member the most senior Judge of that High Court" (Supreme Court of Pakistan, 2019). High Court selection of judges is now primarily determined by a seven-member committee of senior-most serving judges via majority voting.[13]

Reform Mutation. Before 2010, the government directly appointed judges, denoting zero-degree separation from the executive branch in judicial appointments. After 2010, judge appointments have fallen under the purview of the Judicial Commission of Pakistan, where a committee of senior judges selects judges through majority voting. The judicial selection reform unfolded in two stages. The first-generation judges, one degree separated from the government, are appointed by a judicial committee consisting of executive-appointed judges. In the subsequent stage, as the executive-appointed judges retire from the judge hiring panel, the responsibility for judge appointments shifts to those appointed by the judicial committee, representing a second degree of separation from the government. The pivotal moment in the reform's evolution occurs when judges appointed by the Judicial Commission gain a majority on the judicial selection committee. The reform multiplier effect is investigated when the majority of first-generation committee-appointed judges appoint judges (when the executive-appointed judges in the judge hiring panel lose their majority to committee-appointed judges). We call these new appointees, appointed by the majority of committee-appointed judges, the second-generation judges. In this situation, committee-appointed judges effectively control the committee appointments process, resulting in a two-degree separation from the government for judicial appointments. The evolution and dynamics of this reform are illustrated in Figure 1, showcasing the transition from executive-appointed judges to the first and then the second generations of committee appointees.

Mandatory Retirements. High Court judges in Pakistan, similar to their counterparts in many countries with a history of British colonial rule, such as India, have constitutional provisions that

[13] In addition to serving judges, the Judicial Commission for the High Courts is also composed of several members from the legal community who are not serving on the bench. This includes the Federal Law Minister, the Attorney General, a lawyer nominated by the Pakistan Bar Council for a two-year term, and a judge appointed by the Chief Justice also for a two-year term. The Provincial Minister for Law and a lawyer nominated by the Provincial Bar Association are also members of the Commission. These six non-serving-judge members are subject to change simultaneously with shifts in government, while Bar Council and former judge nominees serve fixed two-year terms. Their tenure does not vary with the age structure or mandatory retirement age laws; however, their turnover represents a collective change that occurs in certain years, for instance, when their two-year term expires in the case of a former judge and a lawyer nominated by the Bar Association. Year-fixed effects in our regressions would account for such changes among the non-judge members of the Judicial Commission.

ensure tenure security. The High Courts in both countries, which collectively adjudicate cases for a fifth of the world's population, have a similar organization. For example, the retirement age for High Court judges is set at 62 in both India and Pakistan. Mandatory retirement age, laws governing civil, criminal, and constitutional procedures, and the process for removing judges carried out through the Supreme Judicial Council by senior judges themselves are some of the constitutional safeguards against political interference and provide a level of security of tenure in both these countries.[14]

The Political Economy of the Reform. Why would Pakistani politicians willingly implement a judicial selection reform that involved ceding appointment power to judges? The judicial selection reform, implemented after a decade of military rule, aimed to curtail the military's political influence (Almeida, 2018). Historically, the military had wielded power by seizing executive offices such as the President and Prime Minister. This reform established judge appointments as almost entirely independent of these government officials to foster a new generation of judges committed to upholding constitutional provisions against the abuse of executive power, including that by military rulers (Zafar, 2012). Given the potential for sabotage by Pakistan's military, the reform was drafted and debated with the utmost secrecy by a small parliamentary committee charged with drafting and swiftly passing the legislation (Almeida, 2018). This meant that the selection reform came as an abrupt shock to Pakistan's judicial system, and this paper explores its short- and long-term impacts.[15]

3. Data

At the core of this research lies a large-scale consolidation and digitization of court records in Pakistan, the fifth most populous country in the world. We initiate this process by creating outcome variables based on case-level judgment text data, serving as the foundation for generating measures related to pro-government rulings, merit-based decisions, and adherence to due process of the law. Following the influential work of Djankov et al. (2003), we employ a panel of attorneys to code these variables (more details on this are provided in the Online Appendix A1). This approach is akin to creating "The Supreme Court Database" for U.S. Supreme Court decisions by Washington University's Law School, a data repository also coded by legal experts and extensively utilized in scholarly work (e.g., Spaeth et al. 2014). We combine these case-level data with multiple administrative datasets in the Pakistani judiciary. This includes court records at the Law and Justice Commission of Pakistan (LJCP), archival records at the High Courts registrar offices and Judicial Gazettes of Pakistan. Our

[14] Despite experiencing a series of military coups, such as those in 1977 and 1999, Pakistan has maintained stability in its High Court judicial exits, with less than 2% of High Court judges leaving office prematurely at the time of the coups; the vast majority continue to serve until their mandatory retirement age, both during civilian rule and under military governments.

[15] The change in judicial reform was described by the distinguished legal historian S.M. Zafar, a framer of Pakistan's constitution in the following manner: "The previous procedure was very arbitrary, it was not only arbitrary, but it was noninstitutional, discretionary and as a result bad judges, sometimes *jiyalas* (political ideologues), sometimes *diwanas* (crazies) got into the judiciary. So, we thought that the judicial appointments be through a process, so we institutionalised it. We took away the power, the discretionary power, the arbitrary power (of the executive)" (Zafar, 2012).

dataset significantly advances the work initiated by Mehmood (2022), enhancing it in three pivotal ways. Firstly, it doubles the number of randomly sampled cases involving the State, increasing the statistical power of our analysis.[16] Secondly, it extends the time frame from 1986 to 2022, enabling the study of the reform multiplier effect over the longer term (see Figure 1 for how reform mutates over time). Thirdly and crucially, we enrich our dataset on judicial decisions by integrating it with details on judges' educational backgrounds, the law firms they were associated with as attorneys, and their career progressions after their high-court judgeships, allowing us to construct a comprehensive database on the network of judicial elites using mandatory but confidential reports submitted to the Law and Justice Commission of Pakistan (LJCP) by each High Court.[17] This enables us to trace the full career trajectories of these High Court judges, including their education in law schools, pre-judgeship roles as attorneys in law firms, their tenures as High Court justices, and their subsequent promotions as Supreme Court justices, a task that was not possible in prior work due to data limitations. We achieved near-universal data matching by successfully linking 18,578 out of 20,000 cases across all our datasets. Below, we offer explanations of the key outcome and explanatory variables.

State Wins. The primary outcome variable in our study is "State Wins," representing judicial independence from the government at the case level. This variable is constructed based on judgment orders, capturing details from each case. The State encompasses various executive entities, including local, provincial, and federal governments, the Office of the Prime Minister, the Office of the President, and government agencies, aligning with Montesquieu's (1748) conceptualization of the State as an executive body. How the State Wins vary over time is depicted in Figure 2, while the distribution of the variable can be visualized in Figure B2 of Online Appendix B. Figure 2 illustrates significant declines in State Wins as the judicial appointment regimes transition, initially decreasing from Executive appointments to the first wave of committee appointees, and further diminishing with the advent of the second wave of committee appointees. Our construction of this variable adheres to established literature with the coding outsourced to legal experts such as lawyers (e.g., Djankov et al., 2003; La Porta et al., 2008; Aney et al., 2021). Two independent teams at a prominent law firm led by senior attorneys were tasked with coding the "State Wins" variable, assigning a value of 1 if the State emerged victorious in the dispute and 0 otherwise. Notably, 97% of the judgments received consistent coding from both teams of attorneys.

[16] For details regarding the case sampling procedure relative to the total, please refer to Figure B1 in Appendix B and the corresponding discussion in Online Appendix A2.

[17] Partial biographical information about judges, particularly those currently in service, is accessible online through the High Court websites, as demonstrated by this link: https://data.lhc.gov.pk/judges/sitting_judges. However, we ensure completeness of data by acquiring the entire universe of High Court judges (serving and former from 1986-2022) records maintained by the Law and Justice Commission of Pakistan (LJCP).

Same Law Firm Hire. The second significant outcome variable pertains to the appointment of judges from the same law firm to the bench. This variable comes into effect when a newly appointed judge shares affiliation from the same law firm as at least one member of the judicial hiring committee (judicial commission). Utilizing employment history records of judges at the High Court registries and the Law and Justice Commission of Pakistan (LJCP) reports, which include mandatory but confidential disclosures on judges' prior employment, this variable enables exploration into the hiring practices involved in the committee selection. This enables an examination of whether the selection of judges by first- and second-generation committees strikes a balance between fostering judicial independence from the government and engaging former judges' colleagues from their professional networks. Figure B3 in Appendix B provides illustrative examples of judges who hail from the same law firms as members of the judge hiring panel (judicial committee), while Figure B4 in the same appendix depicts the variation in this variable at the court-by-bench-by-year level.

Herfindahl Index of Judge Concentration. This variable measures the market concentration of appointed judges from the same law firms as at least one of the selection committee members making judicial appointments. It provides a benchmark to assess the concentration of judges from the same law firms —as in the judge hiring panel— between first and second-generation committee appointees. The index takes into account the relative proportion of judges who join the judicial bench from law firms where at least one judge from the hiring panel had previously served as an attorney. In a situation where 100 judges are spread across 10 law firms, the Herfindahl index, when computed in terms of market share, can be calculated as $10 \times 10^2 = 1000$. Converting this to a proportion on a scale from zero to one would give the value of 0.1 (because of the division by the maximum potential value of 10,000). However, if the 100 judges are hired from a different distribution, such as 5 law firms providing ten judges each and 2 law firms providing the remaining 50 judges, the Herfindahl index would be higher because many judges came from only 2 law firms. In this case, the Herfindal index would equal $25^2 \times 2 + 10^2 \times 5 = 1750$. The Herfindahl index ranges from a low value close to 0, indicating a high diversity of law firms or an unsaturated market, to a maximum value of 10,000 (or 1 when measured in proportions) when all judges come from the same law firm. According to the U.S. Department of Justice, an index below 1,000 indicates an unsaturated market, while ranges between 1,500 and 2,500 points to moderate concentration. In comparison, markets with an index exceeding 2,500 points are typically considered highly concentrated (Farrel and Shapiro, 2021).

Decision Quality. To approximate “correctness” or analyze the quality of judicial decisions, we construct two outcome variables: Case Merits and Process Followed, also observed at the case level. These variables are generated using information extracted from the judgment texts. Specifically, the "Case Merits" variable is a binary indicator: it assumes a value of 1 when a decision is grounded in the legal merits of the case and 0 when it is based on legal technicalities. This characterization is consistent

with legal literature, as highlighted by Liu and Li (2019), indicating that decisions anchored in legal merits generally exhibit superior reasoning. Legal scholarship in Pakistan also suggests that decisions based on technicalities (as opposed to case merits) often favor State authorities, potentially leading to biased rulings (Siddique, 2013). The decisions on case merits also align with the aspirations of Common Law jurisprudence, which encourages decisions based on legal merits, considering the spirit of the law rather than relying solely on legal technicalities (Pound, 1963; Tidmarsh, 2009). Thus, we use this dummy variable as a proxy for the quality of judicial decisions. The second variable assessing decision quality is "Process Followed." This variable comprises ratings on the process followed in each judicial case by a panel of attorneys. Their task was to evaluate, on a scale of 1 to 5, the extent to which the pertinent jurisdictional, procedural, and evidential prerequisites were observed in the detailed judicial decisions within the judgment text. This variable approximates the degree to which due process was adhered to in reaching the judicial decision, capturing the quality of the due process followed when making the decision.[18] The process followed was rated identically by the two independent teams of lawyers in around 90% of the cases. Both these measures are becoming standard in empirical legal literature, as rulings on case legal merits are believed to rely less on arbitrary "discretionary rules" and more grounded in evidence and legal principles (Wang et al., 2022), while rulings observing due process of law proxies for the quality of the legal process itself, not just the final decision (Chemerinsky, 2019).[19]

Main Explanatory Variables. The key explanatory variables employed in the analysis are denoted as the "Executive Appointments Vs. 1st Degree" and "Executive Appointments Vs. 2nd Degree". As the first or second generation of committee appointees joins the judicial bench, these dummy variables switch from zero to one, offering identifying variation at the court-by-bench-by-year level. These variables chronicle the sequencing of judicial committee appointments. Specifically, they track the arrival of the first committee-appointed judge (Executive Appointments Vs. 1st Degree) and the arrival of the committee-appointed judge that tilts the committee-appointed judges to a majority in the judge hiring panel (Executive Appointments Vs. 2nd Degree). We also construct related variables that utilize both nearing and reaching of the majority –of committee-appointed judges– in the judicial committee (judge hiring panel). This allows us to run placebo tests to ascertain the effect of nearness but not the attainment of second-generation committee appointments. Among these variables, the key two key are dummy variables that activate when the arrival of a new judge achieves a 4:3 majority of committee appointees within the 7-panel hiring committee (Committee Appointees Majority 4-3), and

[18] These variables were chosen after extensive discussions with legal experts, lawyers, and judges in Pakistan. We specifically thank eminent Supreme Court of Pakistan's lawyer Javed Arshad for his extensive discussions and insights.

[19] The process followed variable is also inspired by the well-grounded legal principle: 'justice is not only done but also seen to be done' (McCarthy v. Sussex Justices, 1924).

the arrival of a judge when the committee appointees near a majority of 3:4 in the judge hiring panel (Committee Appointees Minority 3-4). Figure 1 illustrates the evolution of the judge hiring process from executive-controlled appointments to first- and second-generation committee appointments.[20]

Control Variables. Our regressions use all available time-varying case and judicial bench characteristics as control variables. We present our main results both without any controls and with the complete set of controls applied. These control variables include the number of pages of the judgment order, the presence of the chief justice on the case, the number of judges on the case, the number of lawyers involved in the case, and dummy variables for different case types (for instance, eminent domain vs. political rights "writ" petition). In our robustness checks section, we additionally estimate a more saturated model, incorporating the interaction of baseline (pre-reform) case and judge characteristics with year-fixed effects. Even in this demanding specification, our results remain robust and align closely with the baseline estimates. Table B1 reports the summary statistics for all the variables used in the study. Further details on variable definitions, sources, data construction details, and the sampling procedure are provided in Appendix A for Online Publication.

4. Empirical Methodology

Difference-in-Differences. In this study, we aim to estimate the differences in pro-government rulings as the degrees of separation increase from the executive appointment of judges. Nonetheless, a straightforward regression between pro-government rulings and the modes of judicial appointment can be challenged by significant endogeneity concerns. These concerns encompass issues like reverse causality (e.g., judges with higher pro-government rulings may be more likely to be appointed under the committee appointment system) and omitted variable bias (e.g., judges selected through the committee appointment system might face heightened public accountability compared to executive appointees, which could impact their propensity to issue pro-government rulings). To obtain estimates with causal interpretation, we leverage the first and second-generation committee appointees' arrival within the High Court's judicial bench. Specifically, we estimate the following generalized difference-in-differences model with court-by-bench fixed effects, year-fixed effects, and court-by-bench linear trends:

$$Y_{cjdbt} = \alpha \times Executive\ Appointments\ Vs\ 1st\ Degree_{dbt} + \beta \times Executive\ Appointments\ Vs\ 2nd\ Degree_{dbt} + \boldsymbol{X}_{cdbt} \times \gamma + \lambda_{db} + \delta_t + t \times \theta_{db} + \epsilon_{cjdbt} \quad (1)$$

The subscripts *c*, *j*, *d*, *b*, and *t* index cases, judges, courts or districts, benches, and decision years, respectively. Y represents government victories or other outcome variables, such as the arrival of the

[20] It is worth noting we do not have the information on the exact votes of individual judicial commission members which is restricted and classified information. Despite several attempts to request to declassify the 'minutes' from the commission meetings that contain this information, we were unsuccessful in obtaining this data.

same law firm judge or proxies for decision quality. α and β represent the coefficients of interest, indicating the impact of the first and second generation of committee appointees, respectively. The disparity between these coefficients, β - α, is of primary import and measures the reform multiplier effect. λ and δ are court-by-bench and year-fixed effects, respectively. $t\theta$ denote the 64 court-by-bench specific linear trends. Court-by-bench fixed effects (λ) allow us to account for time-invariant court-by-bench specific factors like the district court's customary laws. The inclusion of year-fixed effects (δ) addresses shared factors influencing all courts and benches over time, such as national elections, government changes, periods of intense media scrutiny, or a specific Chief Justice's tenure. Furthermore, incorporating court-by-bench-specific linear trends ($t\theta$) lets us account for some time-varying factors that might otherwise be omitted, like the changing landscape of local case precedents. Finally, the vector **X** denotes the time-varying case and bench controls. These controls include variables such as the type of case (e.g., political rights dispute), the number of pages in the judge's order, the presence of the state chief justice in the judicial bench, the number of judges and lawyers involved in the case and total judges in the judicial bench.[21] At the baseline, standard errors are clustered at the court-by-bench level, corresponding to the level of treatment assignment (Abadie et al., 2023).[22]

Identification. Ensuring the causal validity of the estimates in Equation (1) depends on effectively addressing the critical identification challenge associated with the potential non-random joining of the first- and second-generation judges on the bench. While including court-by-bench fixed effects helps mitigate concerns regarding level differences, differences across treated and untreated judicial benches could still affect the outcome of interest in a time-varying way. For example, case precedents in certain courts may evolve differently compared to others; therefore, a simple difference-in-differences can mistakenly attribute the effect of such evolution of case precedents to the arrival of new judges. There are several ways we can speak to this issue. Firstly, we can utilize the institutional knowledge about the composition of the judicial hiring committee to extract the treatment effect from the pure arrival of a new judge effect. In particular, we can analyze the judicial decisions following a new judicial appointment, distinguishing between scenarios where committee appointees constitute a minority of executive appointees in the judge hiring committee (3-4 committee-appointed members in the judge hiring panel) and situations where the arrival of a new judge tips the scale, giving committee appointees

[21] In the robustness section, we will demonstrate the stability of our results by employing alternative specifications. This encompasses incorporating district-by-year (which is identical to court-by-year) fixed effects and including interactions between baseline case, bench, and judge characteristics with year fixed effects.

[22] Our results remain robust at conventional significance levels when we use alternate clustering, for instance, when we construct standard errors using Wild Bootstrapping (Cameron et al., 2008), accounting for a potentially smaller number of clusters.

the majority (4-3 committee-appointed members in the judge hiring panel). This difference can help illuminate the reform's multiplier effect from the arrival of a new judge effect. Secondly, we can leverage knowledge about how judicial bench replacements occur, specifically through the mandatory retirement age law. We can evaluate whether the reform multiplier effect is only triggered by mandatory retirements on judges' 62nd birthdays during "transition years," which mark the transition from the first to the second generation of committee appointments. This enables us to isolate the additional effect of judges' plausibly exogenous exits during pivotal "transition years," comparing them to the effect of exits in less pivotal non-transitional years. Thirdly, we can leverage the annual frequency of judicial decisions and estimate an event-study version of the two-way fixed effect model, allowing us to discern treatment effects from slower-moving factors, bolstering our confidence in capturing the impact of the specific institutional change. Fourthly, baseline difference-in-differences can be extended to account for a broader range of time-interacted baseline case, judge, and bench characteristics and include province × year and court × year fixed effects, providing tighter treatment-control comparisons. Lastly, consistent with similar cases showing up in front of treated versus control judicial benches (which follows from the severe backlog in Pakistani courts and random allocation of cases within the bench), we can test whether our treatment variable changes the type of cases that appear in court.

5. Main Results

5.1. Reform Compounding and Government Victories

Difference-in-Differences Estimates. We begin the presentation of the results by reporting evidence on the reform's compounding effect on pro-government rulings. To do this, we estimate the parameters α and β using State Wins as the dependent variable in Equation (1) and report the results in Table 1. The first column of the table displays the results from the simplest specification, which does not incorporate fixed effects, bench-specific linear trends, or case and bench controls. Moving to the second column, we introduce the time-varying case and bench controls, detailed in Table B1 of Appendix B. In the third column, we incorporate court-by-bench and year-fixed effects, allowing us to account for time-invariant factors and common temporal factors impacting pro-government rulings. Finally, the fourth column incorporates court-by-bench specific linear trends, encompassing 64 bench-specific linear trends that capture some time-varying factors influencing pro-government rulings across district benches and over time. According to our preferred baseline specification (Column 4, Table 1), the first-generation committee-appointed judges are about 10 percentage points *less* likely to issue pro-government rulings than the executive-appointed judges. However, the second generation of committee appointees demonstrates an even more substantial decrease in pro-government rulings relative to executive-appointed judges, with a difference of around 30 percentage points. This 20

percentage-point difference, which we call the reform multiplier effect, is statistically significant (p-value < 0.01) and implies a 50% decrease in pro-government rulings compared to the sample mean. Judges' inclination to rule in favor of the government decreases as they move further from executive appointments with a sharp and statistically significant difference in pro-government rulings between the first- and second-generation judges.

Event Study Estimates. The critical assumption for interpreting the reform multiplier effect as causal is that conditional on covariates, there are no systematic differences in the trends of the outcomes of interest among judicial benches with different arrivals of first- and second-generation committee appointees. We estimate an event study to partially test for this parallel trend assumption and examine the dynamics of the reform multiplier effect. We focus on decisions made by first and second-generation committee appointees, using indicators for the time before and after the arrival of second-generation judges on the bench. Specifically, we estimate the following event-study version of the TWFE model using decisions made by the first and second generation of committee appointees:

$$Y_{cjdbt} = \Sigma\, \psi_k \times D_k + \lambda_{db} + \delta_t + \epsilon_{cjdbt} \quad (2)$$

where *Y* is government victories, and *D* is a set of indicator variables that take the value one with the arrival of the second-generation committee-appointed judge in the judicial bench and zero otherwise. The subscript "k" serves as a time index denoting the decision year, and it represents the number of years that have passed since the first second-generation committee appointee took office. For instance, when "k" is equal to +2, it corresponds to judicial decisions observed two years after the judge is hired by the judicial committee (judge hiring panel) with a majority of committee-appointed judges in the panel. λ and δ are court-by-bench and year-fixed effects, respectively. When estimating the model, we treat the year before the arrival of the second-generation committee-appointed judge as the omitted category and compare it with pro-government rulings in the other years. However, the OLS estimation of Equation (2) will give us inconsistent estimates in the presence of heterogeneous treatment effects. Sun and Abraham (2021) show that the fully dynamic version of the Two-Way Fixed Effects (TWFE) model, estimated using Ordinary Least Squares (OLS), provides consistent estimates only under strong assumptions regarding treatment effect homogeneity. To address this challenge, we examine the robustness of our results by estimating Equation (2) using estimators that do not require the homogeneity assumption. These alternative estimators, proposed by Borusyak, Jaravel, and Spiess (2021), Callaway and Sant'Anna (2021), Goodman-Bacon (2021), and Sun and Abraham (2021) shut down the 2x2 difference-in-differences comparisons between newly treated and already treated units to deliver consistent estimates. By applying these robust estimators along with the ordinary least squares (OLS) to estimate Equation (2), we graphically illustrate in Figure 3 Panel A the reform multiplier effect on pro-government rulings year by year (De Borusyak, Jaravel, and Spiess 2021;

Callaway and Sant'Anna 2021; Sun and Abraham 2021). The figure shows how pro-government rulings compare before and after the arrival of second-generation committee appointees across estimators. Estimates across different estimators are consistent with the parallel trends assumption, as evidenced by the coefficients on the years preceding the arrival of second-generation committee appointees being close to zero. Figure 3 Panel A further illuminates the dynamics of the reform multiplier effect, highlighting a persistent effect of the reform mutation across all estimators. The arrival of 2nd generation committee-appointed judges continues to maintain lower pro-government rulings for up to four years relative to rulings by first-generation committee appointees. The reform multiplier effect endures.[23]

Identification Checks. A key identification concern in interpreting our estimates as causal is the possibility that the observed effect is wholly driven by the arrival of new judges per se rather than the reform compounding effect. In other words, the reform multiplier effect—marked by a notable decrease in pro-government rulings from the first to second-generation judges—could arise whenever a new committee-appointed judge joins the judge-hiring panel, irrespective of whether the judge selection committee is dominated by committee or executive appointees. We present two pieces of evidence that counter this conjecture. Firstly, Table 2's Column 1 and 2 demonstrates that the arrival of new committee-appointed judges that bring the judge hiring panel closer to obtaining a majority of committee-appointed judges has no discernible impact on pro-government rulings. Specifically, we find no evidence of a reform compounding effect when the new committee-appointed judges are appointed, bringing committees versus executive appointees in the judge hiring panel close to a majority with a 2:5 or 3:4 ratio at the judge hiring panel. If anything, the point estimates are positive (Table 2, Columns 1 and 2). However, a substantial reform multiplier effect is observed when the committee-appointed judges actually reach a 4:3 majority within the judge hiring panel (Table 2, Columns 3 and 4). In other words, when the hiring committee making the judicial selection decisions is primarily composed of judges appointed by the executive branch, the arrival of committee-appointed judges in the hiring panel does not affect pro-government rulings. The statistically significant decline in pro-government rulings is solely observed when the judge-hiring panel attains a clear majority of committee-appointed judges. These findings corroborate the notion that the selection method substantially impacts the collective conduct of judges. While data on the precise voting within the judicial commission are unavailable, the evidence presented in Table 2 indicates a high level of unanimity among judges, making it improbable that variations in dissenting votes between committee-

[23] We estimated four pre- and post-periods so to ensure the estimation of all the new estimators of TWFE. Estimating additional post-periods would necessitate using already treated units as controls for newly treated units, potentially introducing bias due to heterogeneous dynamic treatment effects (Sun and Abraham 2021). Therefore, the baseline analysis includes up to four post-periods for the most reliable estimation. However, we can include more pre-periods which we do in Figure B5 of Appendix B, and find no evidence of differential pre-trends.

and executive-appointed judges explain the observed results. The judges appear to be casting their votes as a block. Secondly, we utilize a mandatory retirement age law in Pakistan that leads to judge vacancies arising due to their age structure. In Table B2 of Appendix B, we observe that the reform multiplier effect is solely triggered by mandatory retirements on judges' 62nd birthdays during 'transition years,' which mark the transition from executive appointments to the first generation of committee appointments and from the first generation to the second generation of committee appointments (we call them Transition Year = 0, signifying that mandatory retirements promptly follow the assumption of duties by a new generation judge). Mandatory retirements in other non-transition years, occurring, for instance, one or two years before the transition, show no impact on pro-government rulings despite leading to the replacement of judges (in these cases, either the retirement followed executive branch still appointing judges or retirement followed committee-appointed judges still having a minority in the judge hiring panel). Lastly, another challenge in our analysis lies in the non-random composition of cases that appear before the court during transitional versus non-transitional years. However, this non-randomness is unlikely for several reasons. First is the highly congested docket of cases for judges in Pakistan. For instance, cases in Pakistan typically experience substantial delays, averaging about *five to ten years* from filing to issuing a final decision (Dunya, 2023; LJPC, 2023). Therefore, during these brief transitional periods, there is limited opportunity for strategic case filing to impact the exact cases that show up in court. Second, the Supreme Court of Pakistan has established a clear stance against the practice of forum or judge shopping, as evidenced by well-settled jurisprudence. Notable cases such as "Sharaf Faridi v. Federation of Islamic Republic of Pakistan (PLD 1989 Karachi) and Government of Sindh v. Sharaf Faridi (PLD 1994) that made forum shopping practice illegal, making non-random case assignment costly and infrequent (Siddique, 2013). Third, this is corroborated quantitatively when the results remain robust in balance checks. Initially, we substitute the dependent variables with case and bench characteristics and re-estimate equation (1). Subsequently, we perform a more rigorous joint orthogonality test, as proposed in McKenzie (2017). The findings, detailed in Table B3 of Appendix B, indicate that the treatment variable does not correlate with the extensive range of case and bench characteristics, reinforcing the notion that comparable cases are presented at treated and untreated judicial benches. Furthermore, Figures B6 to B8 visually underscore a homogeneous allocation of cases across judges, which appears to produce a pattern one would expect from a random assignment of cases. These observations collectively affirm that the judge selection mechanism do not skew the distribution of cases that appear before the court. Additional robustness and identification checks are reported in an Online Appendix A5.

5.2. Reform Compounding and Judicial Concentration

Difference-in-Differences Estimates. The reform multiplier effect comes with the recomposition of the judicial elite. As the degree of separation from executive appointments increases, there is a noticeable increase in the hiring of judges from the same law firm as the committee responsible for judge selections. Table 3 provides evidence that a similar pattern of reform compounding is also evident in the hiring of judges from the same law firm as at least one member of the judge hiring panel. In Panel A of Table 3, we utilize data on all judges and observe a statistically significant ten percentage-point difference in hiring of judges from the same law firm for first versus second generation of committee-appointees. Panel B in Table 3 narrows its focus to comparing first and second-generation committee appointees to focus on judges who hire judges through the judicial committee. The estimates indicate that the second-generation judges, appointed by the committee-appointed judge hiring panel (who are in the majority), are also approximately ten percentage points more likely to be hired from the same law firm as the panel of judges responsible for making the hiring decisions, relative to the first-generation committee-appointed judges. These results suggest that judges may consolidate their power as the separation from executive appointments increases, assertively hiring from their network of former colleagues from affiliated law firms. The judge selection reform not only bolsters judicial independence but also contributes to the entrenchment of a judicial elite.

Event Study Estimates. The reform multiplier effect of increased recruitment of judges from the same law firm as the judge hiring committee appears to persist. We estimate Equation 2, with the same law firm hiring as the dependent variable, to assess the dynamic impact of the arrival of second-generation judges on hiring judges from the same law firm as the judicial selection committee. Figure 3 Panel B shows these results, with pre-arrival years of second-generation judges having coefficients near zero, indicating no apparent pretrends. Nonetheless, post-arrival coefficients reveal a gradual increase in judge concentration, indicating a rise in judges hiring their former colleagues from the law firms where they formerly worked. Regardless of the estimators used to estimate Equation (2), the influence of hiring judges from the same law firm as the judge hiring panel remains positive as the degree of separation from executive appointment increases. These findings imply that while reform compounding may bolster judicial autonomy from the government, it also results in an enduring reconstitution of the judicial elite. These results also pass the key identification checks we conducted in the case of the pro-government rulings variable. For instance, nearing the majority of committee-appointed judges on the judge hiring panel (2:5 or 3:4 ratio) does not prompt an increase in hiring judges from the same law firm. However, a statistically significant increase in hiring former colleagues from the same law firm by the judge hiring panel is observed only when committee-appointed judges attain the majority (Table B4 in Appendix B).

The Magnitude of Judge Concentration. How concentrated is hiring by the committee-appointed judges compared to executive-appointed judges on the judge hiring committee? To benchmark the magnitude of the concentration of judge hires from the same law firm as at least one of the members of the judge hiring panel, we construct a "Herfindahl–Hirschman Index" of market concentration of judges' law firms utilizing a unique database on judges' prior employment at law firms.[24] In competition policy, the Herfindhahl index serves as a rule of thumb to approximate the level of monopoly power in a market. Under the United States Department of Justice, when assessing market concentration, an HHI value below 1500 typically indicates a market with low concentration. A range of 1500-2500 is considered moderate concentration, while an HHI above 2500 is regarded as a highly concentrated market, often warranting further investigation by the Department of Justice to determine whether there is a need to address potential monopolistic abuse of power. When expressed in terms of market shares, the Herfindahl-Hirschman typically ranges from 0 to 10,000. However, for ease of interpretation, the HHI can be scaled to a range between 0 and 1 by dividing the original value by 10,000. In this scaled form, it is called the HHI in proportions. Appendix B's Figure B9 visually represents the market concentration of judges from the same law firm for those hired by the majority of the executive and committee-appointed judges. For first-generation judges, the index is, 1400 (0.14 in proportion) which suggests a relatively low level of market concentration. In contrast, second-generation committee appointees have an index of about 4800 (0.48 in proportions), reflecting a high concentration level (Farrel and Shapiro, 2021). If this were a market of U.S. firms, an HHI of this magnitude would likely trigger an investigation by the Department of Justice's Antitrust Division to determine if there is a need to dismantle a potential monopoly that might be abusing its power.

From State Capture to Judicial Capture? The ultimate impact of the increased concentration of judges from the same law firm on bureaucratic performance is theoretically ambiguous (Alger and Weibull, 2010; Bramoullé and Goyal, 2018; Chandrasekhar et al., 2020). On the one hand, increased concentration may result in the appointment of less-qualified 'patronage' judges, potentially leading to lower-quality judicial decisions (Brassiolo et al., 2021; Riaño, 2021). On the other hand, a bench with a higher representation of judges from leading law firms could potentially improve the quality of judgments rendered, particularly when these judges are selected based on superior qualifications, like graduating from highly selective law schools or possessing more substantial legal experience. To

[24] Our "Herfindahl-Hirschman" index measures the concentration of judges from the same law firm as the selection panel making judge selections. It considers the proportion of judges arriving at the bench from law firms where the committee members previously worked as attorneys. For example, in a scenario with 10 law firms and 100 judges evenly distributed, the index would be 1000. However, if the judges are hired in a more uneven distribution, such as 5 law firms providing ten judges each and 2 law firms providing the remaining 50 judges, the index would be higher (1750 in this case). This index helps provide a benchmark of the extent of judge concentration. A lower index value indicates a market with a larger number of law firms of relatively equal size. In contrast, a maximum value of 10,000 points is reached if all judges in the judge hiring come from a single law firm.

evaluate these theoretical possibilities, we undertake two approaches. Firstly, we examine how the recruitment of judges from the same law firm impacts our measures of decision quality: decisions on case merits and evaluations of due process followed in the judicial case. Secondly, we explore whether there is a shift in the type of judges selected as the judicial committee increasingly chooses candidates from the same law firms. Initially, to assess whether hiring judges from the same law firm mediates the reform multiplier effect, we estimate the interactive effect of the reform multiplier and the hiring of the same law firm judges. The results, presented in Table 4, indicate that appointing judges from the same law firm as the judicial committee members does not mediate the reform multiplier effect. We observe no statistically significant correlation between the arrival of judges from the same law firm and pro-government rulings (Table 4, Column 1) or decision quality measures, including case merits and adherence to due process of law (Table 4, Columns 2 and 3). Furthermore, no disproportionate effect of the reform is observed with the arrival of judges from the same law firms. Irrespective of whether or not a judge joins the bench from the same law firms as the judicial committee, we observe a similar reform compounding effect on pro-government rulings, case merits, and ratings of due process followed. These findings suggest that the compounding of the same-law firm hires does not compromise judicial independence or decision quality. In the subsequent section, we will delve into how the compounding impact of heightened anti-government rulings aligns with a simultaneous increase in measures of decision quality as degrees of separation from executive appointments rise.

5.3. Reform Compounding and Decision Quality

Difference-in-Differences Estimates. Is there also a similar reform compounding effect on decision quality? Firstly, we investigate the potential compounding effect of the reform on case merits. As discussed earlier, in common law jurisprudence, a ruling on the case's legal merits indicates that the judicial decision is based on evidence rather than legal lacunae (Pound, 1963; Tidmarsh, 2009). Legal scholars in South Asia (including Pakistan) also view rulings based on case merits as better reasoned and less biased (Siddique, 2013). To capture this concept and leverage the literature that uses the text of judgment orders to ascertain decision type (La Porta et al., 2008; Liu et al., 2023; Ash and Poyker, 2023), we construct, based on legal experts' evaluations, a variable that takes the value of one if the decision is based on case merits and zero otherwise. Columns 1 and 2 of Table 5 showcase the estimated results for the baseline Equation (1), with case merits as the dependent variable. These results align with those depicted in Table 4, indicating that the transition from first-generation to second-generation committee-appointed judges has a positive multiplier effect on rulings on the case's legal merits. Notably, comparing coefficient estimates between first- and second-generation judges reveals a statistically significant 20-percentage point increase in case merits-based rulings (Table 5, Column

2), indicating a 30% rise in merit-based decisions over the sample average (p-value < 0.01). Secondly, we investigate whether a reform compounding also exists for greater adherence to due process of the law. Scholars have long argued that judicial decisions should not only be based on case merits but also follow the due process of the law (see, e.g., Dahl, 1957). This aligns with a long-standing legal tradition that "justice should not only be done but should manifestly and undoubtedly be seen to be done," as famously observed by Chief Justice Hubert Parker of the United Kingdom in 'McCarthy v. Sussex Justices" case in 1924. We measure compliance with due legal process by considering attorneys' ratings of each judicial case on due process of law. These ratings serve as a proxy for ascertaining whether the due process of law was observed in determining the judicial decision, a topic almost always elaborated upon in the judgment text.[25] A higher rating for "process followed" signifies the fulfillment of elevated standards concerning jurisdiction, procedure, and evidence standards. Comparing due process ratings between first- and second-generation committee appointees (Table 5, Column 4) shows a significant improvement of around 0.5 points on a 1-5 scale. This indicates a 15% increase in the evaluation of due process compared to the average rating of around 3.5 out of 5. A positive multiplier effect on the due process followed, akin to case merits, is also observed. These findings also satisfy the identification check demonstrated in our analysis of 'State Wins' or 'Same Law Firm Hires' results. Specifically, a near-majority of committee-appointed judges on the judicial committee, with a ratio of 3:4, does not lead to a change in rulings based on merits or adherence to due process. Conversely, when committee-appointed judges attain a 4:3 majority on the judicial committee, we observe a statistically significant increase (see Table B5 in Appendix B for these results).

Event Study Estimates. Substituting case merits and due process as dependent variables in Equation 2 reinforces our findings. This analysis focuses more on the reform multiplier effect by examining cases decided by first-generation versus second-generation committee appointees (excluding cases adjudicated by executive appointees). The event study plots for case merits and due process followed are presented in Figure B10 and Figure B11 of Appendix B, respectively. Figure B10 illustrates a notable jump in rulings on case merits with the arrival of second-generation committee appointees, an increase of roughly 20 percentage points. Similarly, Figure B8 highlights that decisions made just after the arrival of second-generation judges receive ratings one point higher. Importantly, across different estimators of TWFE, we observe no evidence of differential pre-trends for both case merits and due process followed, indicating that before the arrival of the second-generation committee appointees at the bench, decisions based on case merits and adherence to due process remained unaffected. This is consistent with the fact that almost all judicial appointments are determined by the quasi-random rotation of the judicial committee into a majority of committee-appointed judges and mandatory

[25] The use of judgment text length as an indicator of decision quality is also often predicated on the view that lengthier judgments are suggestive of a more due process-complaint adjudication (Ash and MacLeod, 2015).

retirement age law.[26] Collectively, the results from Table 5 and Figures B10 and B11 substantiate the claim that an increase in degrees of separation from executive appointments corresponds to an improvement in the quality of judicial decisions. The multiplier effect in decision quality persists over time.

6. Mechanisms

The extensive collection and digitization of court records for our analysis allowed us to explore multiple mechanisms potentially driving the reform multiplier effect. Ultimately, we identify selection effects as a critical pathway explaining the reform multiplier effect. The judicial committee's selection of second-generation judges appears to show a preference for candidates with more substantial legal experience and education from top law schools, compared to the judicial committee's appointment of first-generation judges.

Reform Compounding and Selection Effects. Our primary findings reveal a reform multiplier effect characterized by a decrease in pro-government rulings and an increase in indicators of elite concentration and decision quality among first- and second-generation committee appointees. To explain these results, we examined whether the judge hiring committee, composed of a majority of executive versus committee-appointed judges, differed in the type of judges they selected. Table 6 presents the results. We observe that the second-generation committee appointees are about ten percentage points more likely to have been educated at an elite "Top 10" ranked law school in Pakistan than the first-generation committee appointees. This is equivalent to a 15% increase in judges hired from these select law schools over the sample mean. Furthermore, the second-generation committee appointees tend to be about 0.75 years or about nine months older at the time of appointment. Given that eligibility for judgeship mandates ten years of practice as an attorney, it suggests that second-generation committee-appointed judges may have more experience in the legal profession than first-generation appointees. The synergy of having more experienced attorneys, educated from top-ranked law schools, elevated to judgeships by the judicial committee with a majority of committee appointees may explain the multiplier effect contributing to improved decision quality relative to decisions by first-generation appointees. Earlier findings presented in Table 4 (Columns 2 and 3) also supported this interpretation, highlighting that an increased concentration of judges from the same law firm as the judicial committee members had no adverse impact on decision quality. Moreover, other observed characteristics, such as judges' tenure at the bench and gender composition (Table 6), are similar for judges selected by a majority of executive-appointed and committee-appointed judicial committee members. Consequently, the findings in Tables 4 and 6 suggest that the observed reform multiplier

[26] Figure B12 in the Appendix depicts a robust correlation between mandatory retirements at age 62 and the subsequent appointments of new judges.

effect can be attributed to a more meritocratic selection process characterized by the appointment of judges who are both more experienced and academically distinguished, leading to decisions that are assessed as higher in quality by legal experts. Heightened elitism does not necessarily undermine judicial independence or the quality of decisions, as the practice of hiring judges from within their networks may coincide with the selection of academically distinguished candidates.

Causal Mediation Analysis. We further utilize the causal mediation approach developed by Imai et al. (2010) and further popularized by the influential work of Pearl and Mackenzie (2018). Using this framework we show how selection effects are key in mediating the reform multiplier effect. Specifically, we employ the terminology suggested by Nguyen et al. (2021) to dissect the overall impact (average treatment effect) into two factors: the direct effect (NDE), which does not depend on the mediator, and the indirect effect (NIE) via a mediating factor. This method, also discussed by Heckman and Pinto (2015) in the context of economic research, allows us to quantify the contribution of selection effects to the reform multiplier effect. Our findings indicate that the educational pedigree of second-generation judges, especially their higher attendance at elite law schools, accounts for approximately 40% of the observed disparity in pro-government decisions or measures of decision-quality compared to their first-generation counterparts. Our findings, detailed in Panel B of Table 6 and further elaborated in Table B6 of Appendix B, reveal that among the mediating variables considered, only attendance at top law schools is qualitatively and statistically significant. This underscores the substantial influence that education at elite law institutions has on the reform multiplier effect, suggesting its significant mediating role in explaining the reform multiplier effect.

Alternative Mechanisms. Finally, in the Online Appendix A4, drawing upon recent academic literature, we investigate a spectrum of alternative explanations for the observed reform multiplier effect. Our analysis evaluates and tests a variety of underlying factors, such as the sway of judicial ideology, career motivations among judges, the impact of peer dynamics or spillover phenomena, fidelity to established case law precedents, strategic case filings, and the potential influence exerted by corporations or the legal profession as explaining the reform multiplier effect (Bonica and Sen, 2021; Bertrand et al., 2020; Sunstein et al., 2007; Chen, 2023; Yang, 2016; Acemoglu et al., 2020; Viktorov, 2019; Khan, 2019). The evidence presented in Appendix A4 undermines the plausibility of these alternative explanations for the reform multiplier effect. Collectively, the evidence aligns with the selection effect mechanism.

7. Conclusion

Over the past five decades, countries, including consolidated democracies, have pursued judge selection reforms to bolster judicial independence from the government (World Justice Report, 2023). The creation of judge selection bodies like the UK's Judicial Appointments Commission, Kenya’s

Judicial Service Commission, and India's Collegium System exemplifies efforts to shield the judiciary from political influence in appointments. These judicial reforms reflect a global push toward establishing impartiality and meritocracy in judicial appointments. However, institutional reforms may be permanent or transient (Acemoglu and Robinson, 2013). Some reforms may become ingrained in the state's fabric, resisting reversals, while others can prove short-lived. This dynamic prompt an essential inquiry: Why do certain reforms endure while others face backtracking?

This study advances a new conceptual perspective on the durability of reform initiatives: the resilience of judicial reforms is less a function of the appointment of discrete officials and more a reflection of the selection mechanisms governing the appointing authorities. The selection of selectors emerges as a pivotal factor in fostering and perpetuating the institutional equilibrium.

In particular, this study shows that the hiring of state officials like judges can generate a cascading effect on judicial independence and decision quality even when they increase concentration among the judicial elite. Specifically, we analyze a judge selection reform in Pakistan that shifted the power to appoint judges from the executive to a judicial committee in two stages. Initially, a committee with judges appointed by the president and prime minister selected the first-generation judges. Subsequently, these first-generation judges, chosen by the judicial committee, gained a majority in the judge hiring committee, overseeing the appointment of second-generation judges. This process resulted in a two-step separation from executive control in judicial appointments. This paper explores the impact of this separation of powers on judicial decision-making and the makeup of the judicial elite.

As the reform mutated from the first degree to the second degree of separation from the executive appointment, there is a substantial and enduring reform multiplier effect on pro-government rulings. The reform multiplier effect, however, is accompanied by a heightened concentration of hiring within a network of elite jurists, as judges appointed from the same law firms as the judicial selection committee substantially increase with each degree of separation from government-controlled appointments. Interestingly, this does not compromise but, instead, enhances decision quality: we observe a corresponding ripple effect of the reform on decision quality, resulting in more rulings on case merits and adherence to due process of law. In exploring potential channels mediating the reform multiplier effect, our results point towards "selection effects." The judge hiring committee, mainly consisting of committee appointees, tends to choose judges graduated from top law schools compared to the judicial committee comprising government-appointed judges. Causal mediation analysis suggests this channel explains approximately 40% of observed disparity in judicial rulings.

These findings bear implications for the study of judiciaries and democratic institutions, particularly given the distinctive nature of the judiciary in contrast to other government branches. One fundamental aspect of the judiciary is its presumed independence, distinguishing it from entities like

the civil bureaucracy or the legislature, which, to some extent, are accountable to the government. Given its exceptional role, studying the judiciary, which is supposed to serve as "constraints on the executive" (Acemoglu et al., 2001; p. 1394), becomes critical. This research presents one of the first systematic empirical investigations into the cascading effect of a judge selection reform that increases the separation of powers between the executive and the judiciary. Our results emphasize that efforts to strengthen judicial independence in developing nations can effectively boost both judicial autonomy and decision quality, even when they result in greater concentration among the judicial elite.

There is a need for further research that sheds light on the dynamics of judicial independence worldwide, including Russia, Poland, and Hungary, which have recently implemented judicial reforms that go in the opposite direction of decreasing the separation of powers between the government and the judiciary. Through a collective examination of these cases, we can glean valuable insights into the broader challenges and implications of executive control over the judiciary, unraveling the dynamics that shape the delicate balance between state power and judicial autonomy.

References

Abadie, A., Athey, S., Imbens, G.W. and Wooldridge, J.M., 2023. When should you adjust standard errors for clustering?. The Quarterly Journal of Economics, 138(1), pp.1-35.

Aberra, A. and Chemin, M., 2018. Does access to the legal system increase investment? Evidence from a randomized experiment in Kenya. Journal of Development Economics, 150.

Acemoglu, D., Johnson, S. and Robinson, J.A., 2001. The colonial origins of comparative development: An empirical investigation. American Economic Review, 91(5), pp.1369-1401.

Acemoglu, D., 2010. Theory, general equilibrium, and political economy in development economics. Journal of Economic Perspectives, 24(3), pp.17-32.

Acemoglu, D., Johnson, S., Querubin, P. and Robinson, J.A., 2008. When does policy reform work? The case of central bank independence (No. w14033). National Bureau of Economic Research.

Acemoglu, D., Cantoni, D., Johnson, S. and Robinson, J.A., 2011. The consequences of radical reform: The French Revolution. American Economic Review, 101(7), pp.3286-3307.

Acemoglu, Daron, and James A. Robinson. 2013. "Economics versus Politics: Pitfalls of Policy Advice." Journal of Economic Perspectives, 27 (2): 173-92.

Acemoglu, D., García-Jimeno, C. and Robinson, J.A., 2015. State capacity and economic development: A network approach. American Economic Review, 105(8), pp.2364-2409.

Acemoglu, D., Cheema, A., Khwaja, A.I. and Robinson, J.A., 2020. Trust in state and non-state actors: Evidence from dispute resolution in Pakistan. Journal of Political Economy, 128(8), pp.3090-3147.

Acemoglu, D. and Robinson, J.A., 2020. The narrow corridor: States, societies, and the fate of liberty. Penguin.

Acemoglu, D. and Wolitzky, A., 2021. A theory of equality before the law. The Economic Journal, 131(636), pp.1429-1465.

Alesina, A. and Romer, D., 2008. When Does Policy Reform Work? The Case of Central Bank Independence. Comments and Discussion. Brookings Papers on Economic Activity, 2008, pp.419-429.

Acemoglu, D., Egorov, G. and Sonin, K., 2015. Political economy in a changing world. Journal of political economy, 123(5), pp.1038-1086.

Mian, A., Sufi, A. and Trebbi, F., 2014. Resolving debt overhang: Political constraints in the aftermath of financial crises. American Economic Journal: Macroeconomics, 6(2), pp.1-28.

Alesina, A., Ardagna, S. and Trebbi, F., 2006. Who adjusts and when? The political economy of reforms. IMF staff papers, 53, pp.1-29.

Ali, B. 2023. Travesty of Justice: Unintended Consequences of Anti-Corruption Efforts. Mimeo.

Arshad, J. 2023. Lawyer Supreme Court of Pakistan. Interview with Author.

Ash, E., Asher, S., Bhowmick, A., Bhupatiraju, S., Chen, D.L., Devi, T., Goessmann, C., Novosad, P. and Siddiqi, B., 2022. Measuring gender and religious bias in the indian judiciary.

Ash, E. and Poyker, M., 2023. Conservative news media and criminal justice: Evidence from exposure to fox news channel. Forthcoming at the Economic Journal.

Ash, E. and MacLeod, W.B., 2015. Intrinsic motivation in public service: Theory and evidence from state supreme courts. The Journal of Law and Economics, 58(4), pp.863-913.

Alger, I. and Weibull, J.W., 2010. Kinship, incentives, and evolution. American Economic Review, 100(4), pp.1725-1758.

Aneja, A. and Xu, G. 2023. Strengthening State Capacity: Civil Service Reform and Public Sector Performance during the Gilded Age.

Aney, M.S., Dam, S. and Ko, G., 2021. Jobs for justice: Corruption in the Supreme Court of india. The Journal of Law and Economics, 64(3), pp.479-511.

Angrist, J.D. and Pischke, J.S., 2009. Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

Bandiera, O., Callen, M., Casey, K., La, E., Ferrara, C.L. and Teachout, M., 2021. State effectiveness. International Growth Centre Evidence Paper.

Bandiera, O., Burgess, R., Deserranno, E., Morel, R., Sulaiman, M. and Rasul, I., 2023. Social incentives, delivery agents, and the effectiveness of development interventions. Journal of Political Economy Microeconomics, 1(1), pp.162-224.

Besley, T., 2005. Political selection. Journal of Economic Perspectives, 19(3), pp.43-60.

Besley, T. and Case, A., 2000. Unnatural experiments? Estimating the incidence of endogenous policies. The Economic Journal, 110(467), pp.672-694.

Besley, T. and Persson, T., 2009. The origins of state capacity: Property rights, taxation, and politics. American Economic Review, 99(4), pp.1218-1244.

Besley, T. and Persson, T., 2014. Why do developing countries tax so little?. Journal of economic perspectives, 28(4), pp.99-120.

Bertrand, M., Burgess, R., Chawla, A. and Xu, G., 2020. The glittering prizes: Career incentives and bureaucrat performance. The Review of Economic Studies, 87(2), pp.626-655.

Boehm, J. and Oberfield, E., 2020. Misallocation in the Market for Inputs: Enforcement and the Organization of Production. The Quarterly Journal of Economics, 135(4), pp.2007-2058.

Bonica, A. and Sen, M., 2021. Estimating judicial ideology. Journal of Economic Perspectives, 35(1), pp.97-118.

Boston, J. and Silveira, B.S., 2019. The electoral connection in court: How sentencing responds to voter preferences. Journal of Law and Courts, pp.1-22.

Bramoullé, Y. and Goyal, S., 2016. Favoritism. Journal of Development Economics, 122, pp.16-2

Brassiolo, P., Estrada, R., Fajardo, G. and Martinez-Correa, J., 2021. Family Rules: Nepotism in the Mexican Judiciary.

Bruhn, M. and McKenzie, D., 2009. In pursuit of balance: Randomization in practice in development field experiments. American Economic Journal: Applied Economics, 1(4), pp.200-232.

Cameron, A.C., Gelbach, J.B. and Miller, D.L., 2008. Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90(3), pp.414-427.

Chandra, A., Kalantry, S. and Hubbard, W.H., 2023. Court on Trial: A Data-Driven Account of the Supreme Court of India.

Chandrasekhar, A.G., Morten, M. and Peter, A., 2020. Network-Based Hiring: Local Benefits; Global Costs (No. w26806). National Bureau of Economic Research.

Chemerinsky, E., 2019. Constitutional law: principles and policies. Aspen Publishing.

Chemin, M., 2020. Judicial efficiency and firm productivity: Evidence from a world database of judicial reforms. Review of Economics and Statistics, 102(1), pp.49-64.

Chemin, M., 2021. Can judiciaries constrain executive power? Evidence from judicial reforms. Journal of Public Economics, 199, p.104428.

Chen, D.L., 2023. Judicial compliance in district courts. International Review of Law and Economics, p.106122.

Choi, S.J., Gulati, M., and Posner, E.A. (2013). The law and policy of judicial retirement: An empirical study. The Journal of Legal Studies, 42(1): pp. 111–150.

De Montesquieu, C. 1748. Firmin Didot frères. Volume 1. Esprit des lois.

Dahl, R., 1957. "Decision-Making in a Democracy: The Supreme Court as a National Policy-Maker." Journal of Public Law 6 (1): 6–279.

Dincecco, M. and Katz, G., 2016. State capacity and long-run economic performance. The Economic Journal, 126(590), pp.189-218.

Dunya News. (2023) 'Shortage of judges causing huge backlog of cases in superior judiciary.' [Online] Available: https://dunyanews.tv/en/Pakistan/735636-3,80,539-cases-pending-in-superior-courts-of-Pakistan

Epstein, L. and Segal, J.A., 2000. Measuring issue salience. American Journal of Political Science, pp.66-83.

Farrell, J. and Shapiro, C., 2021. The 2010 Horizontal Merger Guidelines After 10 Years. Review of Industrial Organization, 58, pp.1-11.

Finan, F., Olken, B.A. and Pande, R., 2017. The personnel economics of the developing state. Fenizia, A., 2022. Managers and productivity in the public sector. Econometrica, 90(3), pp.1063-1084.

Hadfield, G.K. and Weingast, B.R., 2014. Microfoundations of the Rule of Law. Annual Review of Political Science, 17, pp.21-42.

Hamilton, A. (1988) "The Judiciary Department." The Federalists Paper 78, 14.

Handbook of economic field experiments, 2, pp.467-514.

Gennaioli, N. and Shleifer, A., 2007. The evolution of common law. Journal of Political Economy, 115(1), pp.43-68.

Gibbons, R. and Murphy, K.J., 1992. Optimal incentive contracts in the presence of career concerns: Theory and evidence. Journal of Political Economy, 100(3), pp.468-505.

Hadfield, G.K. and Weingast, B.R., 2012. What is law? A coordination model of the characteristics of legal order. Journal of Legal Analysis, 4(2), pp. 471-514.

Hayek, Friedrich August. 1960. The Constitution of Liberty: The Definitive Edition. Chicago: University of Chicago Press.

Heckman, J.J. and Pinto, R., 2015. Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. Econometric reviews, 34(1-2), pp.6-31.

Imai,K., L.Keele, andD. Tingley. 2010.Ageneral approach tocausalmediationanalysis. PsychologicalMethods 15: 309–334. https://psycnet.apa.org/doi/10.1037/a0020761.

Jain, S., Majumdar, S. and Mukand, S.W., 2014. Walk the line: Conflict, state capacity and the political dynamics of reform. Journal of Development Economics, 111, pp.150-166.

Khan, M.S. (2019) 'Empowerment without Accountability? The Lawyers' Movement in Pakistan and its Aftershocks', IDS Bulletin, 50(3). Available at: https://bulletin.ids.ac.uk/index.php/idsbo/article/view/3049/Online%20article

Khan, A.Q., Khwaja, A.I. and Olken, B.A., 2016. Tax farming redux: Experimental evidence on performance pay for tax collectors. The Quarterly Journal of Economics, 131(1), pp.219-271.

Kureshi, Y., 2022. Seeking Supremacy: The Pursuit of Judicial Power in Pakistan. Cambridge University Press.

Lambais, G. and Sigstad, H., 2023. Judicial subversion: The effects of political power on court outcomes. Journal of Public Economics, 217, p.104788.

Leucht, L. (2023) 'Jobs for Votes: Patronage and Performance in Tammany Hall's NYPD.' Available at: https://www.lukasleucht.com/pdf/JMP_Leucht.pdf (Accessed: 22 November 2023).

Liu, John Zhuang and Xueyao Li,. 2019, "Legal techniques for rationalizing biased judicial decisions: Evidence from experiments with real judges," Journal of Empirical Legal Studies, 2019, 16(3), 630–670.

Liu, E., Lu, Y., Peng, W., and Wang, S., 2023. Court capture, local protectionism, and economic integration: Evidence from China (No. w30432). National Bureau of Economic Research. Available at: http://www.sdwang.org/uploads/4/4/8/5/44856715/judicial_independence_draft__17_.pdf [Accessed 2 January 2024].].

Malik, A. (2022). 'Protection for Sale': The Political Economy of Trade Protection in Pakistan. Pakistan Institute of Development Economics (PIDE) Discussion Paper.

McKenzie, D., 2017. Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. American Economic Review, 107(8), pp.2278-2307.

Mehmood, S., 2022. The impact of Presidential appointment of judges: Montesquieu or the Federalists?. American Economic Journal: Applied Economics, 14(4), pp.411-45.

Mehmood, S., 2023. Judicial Independence and Development: Evidence from Pakistan. Mimeo.

Mehmood, S., and Ali, B., (2023). Judicial Capture. The Economic Journal. Accepted.

Moreira, D. and Pérez, S., 2021. Civil service exams and organizational performance: Evidence from the Pendleton Act (No. w28665). National Bureau of Economic Research.

Muralidharan, K., & Niehaus, P. (2017). Experimentation at Scale. Journal of Economic Perspectives, 31, 103-124.

Nguyen, T.Q., Schmid, I. and Stuart, E.A., 2021. Clarifying causal mediation analysis for the applied researcher: Defining effects based on what we want to learn. Psychological Methods, 26(2), p.255.

North, D. 1990. Institutions, Institutional Change, and Economic Performance. New York: Cambridge University Press.

North, D.C., Wallis, J.J. and Weingast, B.R., 2009. Violence and social orders: A conceptual framework for interpreting recorded human history. Cambridge University Press.

Pearl, J. and Mackenzie, D., 2018. The book of why: the new science of cause and effect. Basic books.

Pakistan Law Digest. PLD (1989). Sharaf Faridi v. Federation of Islamic Republic of Pakistan (PLD 1989 Karachi)

Pakistan Law Digest. PLD (1994) Government of Sindh v. Sharaf Faridi (PLD 1994 SC 105, Islamabad).

Pastor, L. and Veronesi, P., 2012. Uncertainty about government policy and stock prices. The journal of Finance, 67(4), pp.1219-1264.

Baker, S. and Mezzetti, C., 2012. A theory of rational jurisprudence. Journal of Political Economy, 120(3), pp.513-551.

Priest, G.L. and Klein, B., 1984. The selection of disputes for litigation. The Journal of Legal Studies, 13(1), pp.1-55.

Pound, Roscoe. 1963. The spirit of common law. Transaction Publishers.

Reuters (2023) 'What is Israel's new judicial law and why is it causing upheaval?', Reuters, 25 July. Available at: https://www.reuters.com/world/middle-east/what-is-israels-new-judicial-law-why-is-it-causing-upheaval-2023-07-25/ (Accessed: 3 March 2024).

Riaño, J.F., 2021. Bureaucratic nepotism. Available at SSRN 3995589.

Spaeth, H., Epstein, L., Ruger, T., Whittington, K., Segal, J. and Martin, A.D., 2014. Supreme Court database code book. URL: http://scdb.wustl.edu.

Rodrik, D., 2021. Why does globalization fuel populism? Economics, culture, and the rise of right-wing populism. Annual Review of Economics, 13, pp.133-170.

Roodman, D., Nielsen, M.Ø., MacKinnon, J.G. and Webb, M.D., 2019. Fast and wild: Bootstrap inference in Stata using boottest. The Stata Journal, 19(1), pp.4-60.

Rose, E.K., 2021. Who gets a second chance? Effectiveness and equity in supervision of criminal offenders. The Quarterly Journal of Economics, 136(2), pp.1199-1253.

Siddique, O. 2013. An Alien Justice: Pakistan's Experience with Common Law.

Smith, A. 1776. The Wealth of Nations [1776]. Vol.1. 1937 edition.

Sunstein, C.R., Schkade, D., Ellman, L.M. and Sawicki, A., 2007. Are judges political?: an empirical analysis of the federal judiciary. Brookings Institution Press.

Supreme Court of Pakistan. (2019). "ARTICLE 175(A) APPOINTMENT OF JUDGES." Retrieved from: https://www.supremecourt.gov.pk/downloads_judgements/all_downloads/judicial_commission/ARTICLE_175(A)_APPOINTMENT_OF_JUDGES.pdf

TechJuice. 2020. "These Are the 12 Richest and Biggest Business Owners in Pakistan.": https://www.techjuice.pk/12-richest-and-biggest-business-owners-in-pakistan.

Tidmarsh, Jay. 2009. "Resolving cases on the merits." Denv. UL Rev (87): 407.

Ujhelyi, G. (2014)."Civil service reform, " Journal of Public Economics,118,15–25.

Vannutelli, S., 2022. From Lapdogs to Watchdogs: Random Auditor Assignment and Municipal Fiscal Performance (No. w30644). National Bureau of Economic Research.

Voigt, S., Gutmann, J. and Feld, L.P., 2015. Economic growth and judicial independence, a dozen years on: Cross-country evidence using an updated set of indicators. European Journal of Political Economy, 38, pp.197-211.

Viktorov, I., 2019. Russia's network state and reiderstvo practices: The roots to weak property rights protection after the post-communist transition.

Wang, Y., 2018. Relative capture: quasi-experimental evidence from the Chinese judiciary. Comparative Political Studies, 51(8), pp.1012-1041.

World Justice Project., 2023 Home. Available at: https://worldjusticeproject.org/ (Accessed: 15 April 2023).

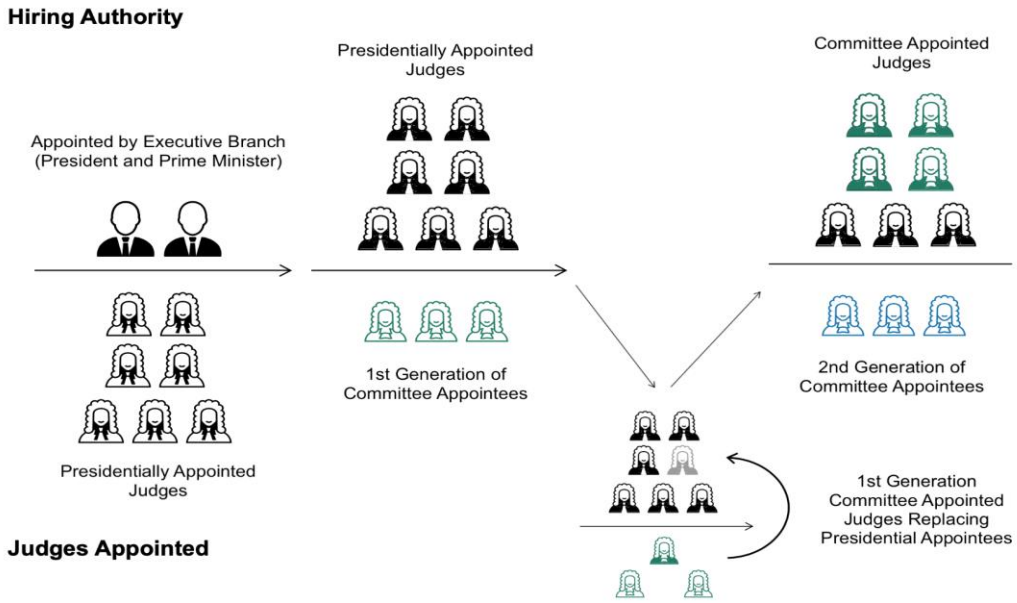
Xu, G., 2018. The costs of patronage: Evidence from the British empire. American Economic Review, 108(11), pp.3170-3198.

Yang, C.S., 2016. Resource constraints and the criminal justice system: Evidence from judicial vacancies. American Economic Journal: Economic Policy, 8(4), pp.289-332.

Zafar, S M (2012), "Inaugural Address on 18th constitutional", IPRI Policy Brief.

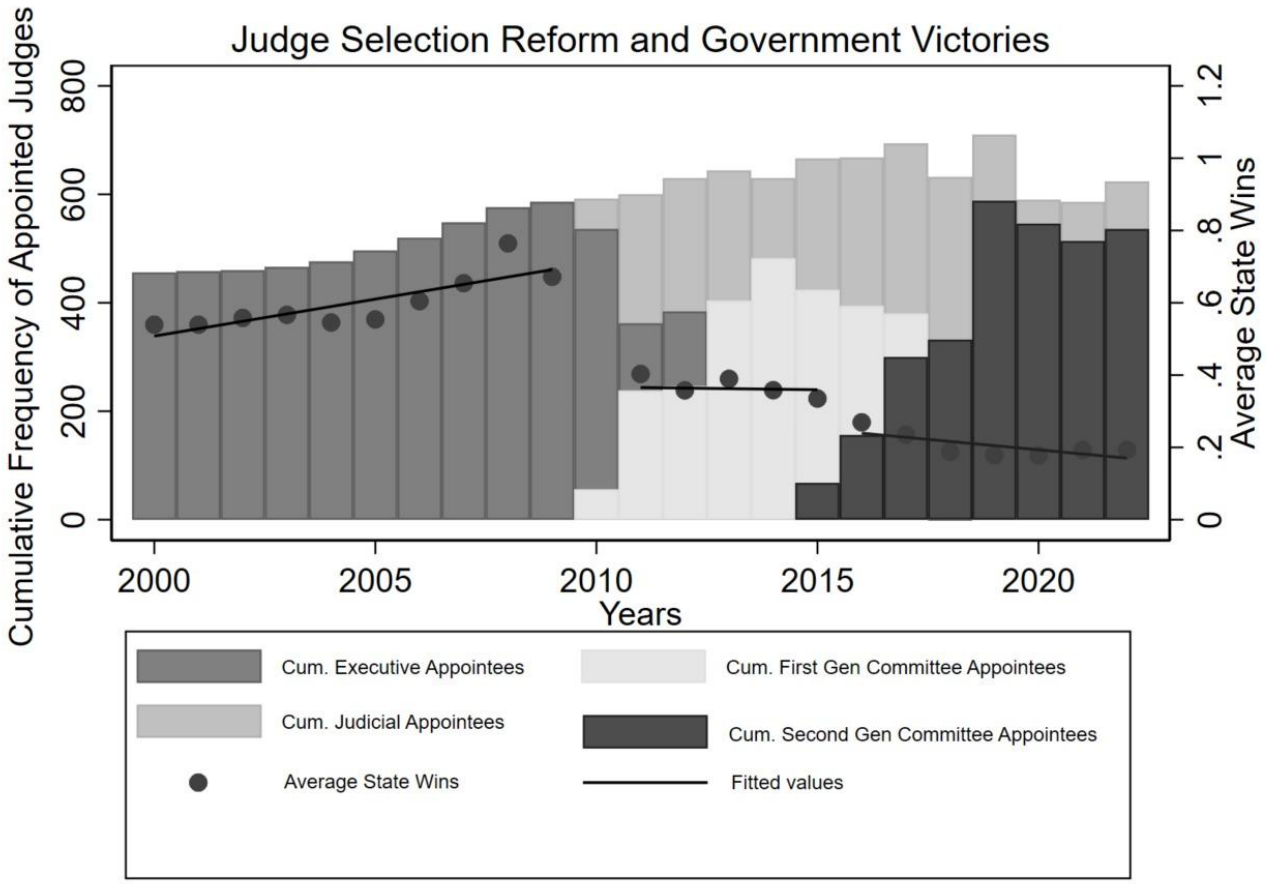
Figures and Tables

Figure 1: The Reform Mutation



Note: The upper half of the figure highlights the appointing authority, with the lower half detailing the appointed judicial positions. In the first stage (pre-2010, situated on the leftmost side of the figure), the hiring authority rests with the President and Prime Minister, responsible for choosing executive-appointed judges. Transitioning to the second stage (at the figure's midpoint), a committee of 7 judges assumes the role of the judicial selection, appointing the first-generation committee-appointed judges—a level removed from direct executive appointments. In the third stage (positioned on the rightmost side of the figure), the hiring authority undergoes a significant transformation, now predominantly composed of committee-appointed judges. Notably, the first-generation committee-appointed judges obtain a 4:3 majority over their executive-appointed counterparts within the judge hiring panel. In this scenario, committee-appointed judges essentially make the committee appointments, resulting in judicial appointments being two degrees removed from the executive branch.

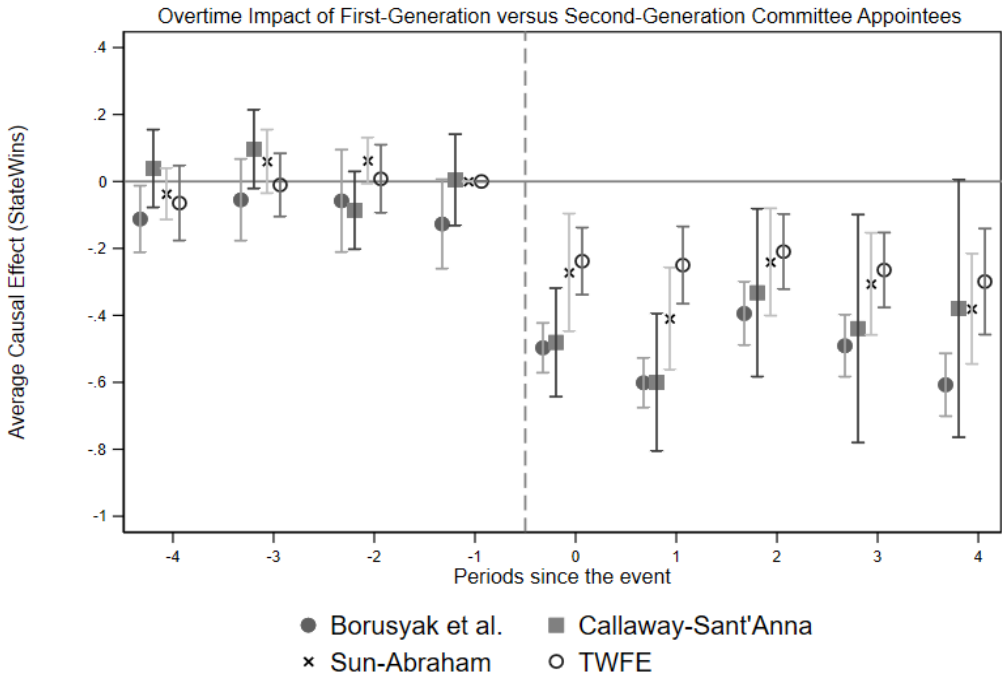
Figure 2: Reform Evolution and Government Victories



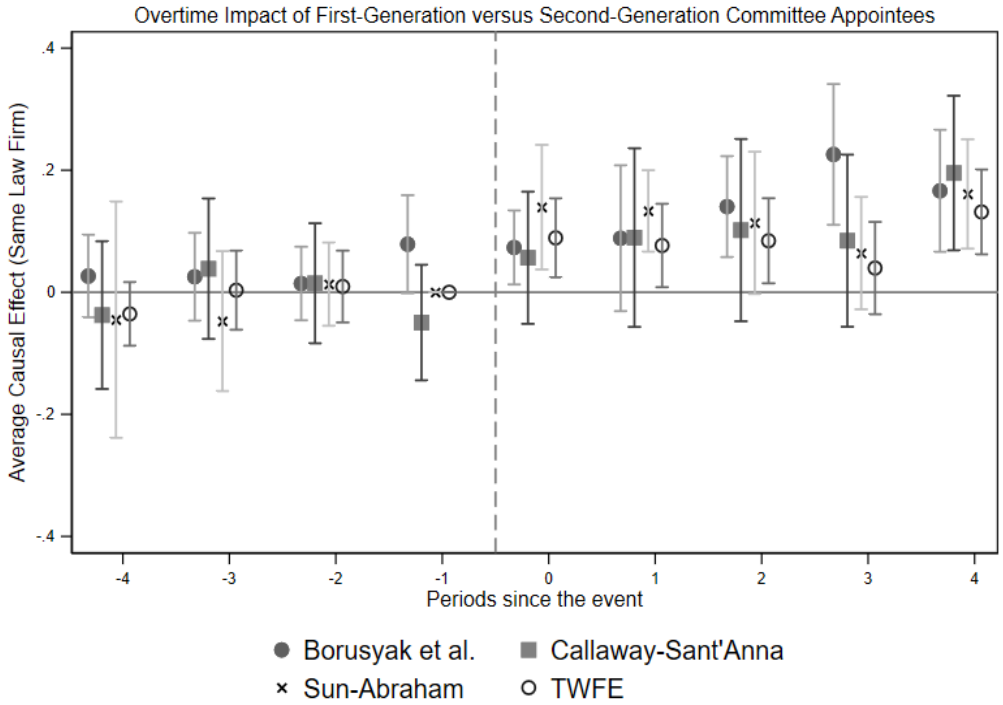
Note: The chart above illustrates the cumulative appointments of judges across three different selection regimes. The dark bars on the left represent the cumulative number of judges appointed by the executive branch (0th-degree of separation from the executive). The lighter bars in the middle represent the first generation of committee-appointed judges (1st-degree of separation from the executive). In contrast, the darkest bars on the right represent the second generation of committee-appointed judges (2nd-degree of separation from the executive). The light grey bars in the background represent the total cumulative judicial appointments across all selection regimes. The chart further illustrates the line of best fit for an average rate of government victories during three critical periods of reform evolution.

Figure 3: Impact of 1st Generation vs 2nd Generation Committee-Appointed Judges

Panel A: On State Wins



Panel B: On Same Law Firms Hires



Notes: This figure presents event-study plots using four different estimators: a dynamic TWFE model with OLS (empty circle markers), Sun and Abraham (2021) (cross markers), Callaway and Sant'Anna (2021) (square markers), and Borusyak, Jaravel, and Spiess (2021) (filled circle markers). In Panel A, the outcome variable is 'State Wins,' an indicator variable for government victory that assumes a value of 1 for government wins and 0 for defeats. In Panel B, the dependent variable is the Same Law Firm judicial hire, a dummy variable that switches on if the newly appointed judge had previously worked at the same law firm as at least one of the judges from the hiring panel's former law firm (where they worked as attorneys before they became judges). The treatment variable activates with the arrival of the first second-generation committee-appointed judge, i.e., when the judicial hiring panel obtains a majority of committee-appointed judges at the judicial bench. The comparison is made with decisions of the 1st generation committee-appointed judges, excluding rulings by executive-appointed judges. The omitted category represents the year before the arrival of the first 2nd-Generation committee-appointed judge at the judicial bench. The figure displays four post periods because estimating additional post periods would require using already treated units as controls for newly treated units, which could introduce bias due to heterogeneous treatment effects. Hence, newly introduced robust estimators limit the maximum number of robustly estimable post-periods to four. We also use for preperiods, as the Borusyak, Jaravel, and Spiess (2021) estimator experiences significantly increased standard errors after four preperiods. The bars depict 95 percent confidence intervals with standard errors clustered at the court-by-bench level. Figure B5 in Appendix B displays a comparable event study estimating all pre-periods by OLS.

Table 1: Impact of First and Second-Generation Committee Appointees on State Wins

	State Wins			
	(1)	(2)	(3)	(4)
Executive Appointments Vs. 1st Degree	-0.248***	-0.233***	-0.0553	-0.0890*
	[0.0220]	[0.0215]	[0.0509]	[0.0485]
Executive Appointments Vs. 2nd Degree	-0.425***	-0.409***	-0.297***	-0.338***
	[0.0191]	[0.0213]	[0.0770]	[0.0766]
Controls	No	Yes	Yes	Yes
Court-by-Bench and Year FE	No	No	Yes	Yes
Court-by-Bench Linear Trends	No	No	No	Yes
The mean of the dependent variable	0.443	0.443	0.443	0.443
R-squared	0.099	0.104	0.152	0.166
Observations	18,578	18,578	18,578	18,578
F-Statistics (1st Degree = 2nd Degree)	99.650	84.130	15.340	17.620
P-Value (1st Degree = 2nd Degree)	p-value	p-value	p-value	p-value
	<0.01***	<0.01***	<0.01***	<0.01***

Note: The table displays the estimated reform multiplier effect on the likelihood of government victories. Coefficients α and β are estimated from equation (1), with the state victory as the dependent dummy variable, coded as one for cases adjudicated in the state's favor and zero otherwise. Executive Appointment Vs. 1st Degree is a variable that activates with the arrival of the first committee-appointed judge at the bench (one degree of separation from the Executive Appointments). The Executive Appointment Vs. 2nd Degree is a variable representing the arrival of the second-generation committee-appointed judge at the district bench (two degrees of separation from the Executive Appointments). Specifically, the variable switches on when the first second-degree committee-appointed judge arrives at the district bench. The comparison of judicial decisions by committee appointees is made relative to cases decided by the judges appointed by the executive branch in both cases. Controls include case type fixed effects (e.g., rights dispute or property dispute), number of pages of judgment order, presence of chief justice on the bench, number of lawyers on the case, number of judges on the case, and total judges on the bench. The estimates from the baseline equation (1) with court-by-bench or district-by-bench fixed effects, year-fixed effects, and bench-specific linear trends are reported in Column 4. Standard errors are reported in brackets and are clustered at the court-by-bench level. ***p<0.01, ** p<0.05, * p<0.1.

Table 2: Impact of Minority versus Majority of Committee Appointees in the Judge Hiring Panel

	(1)	(2)	(3)	(4)
	State Wins			
Committee Appointees Minority 2-5	0.00933			0.000316
	[0.108]			[0.100]
Committee Appointees Minority 3-4		0.0196		0.0206
		[0.0947]		[0.0821]
Committee Appointees **Majority** 4-3			-0.243***	-0.243***
			[0.0524]	[0.0523]
Court-by-Bench and Year FE	Yes	Yes	Yes	Yes
Controls & Court-by-Bench Trends	Yes	Yes	Yes	Yes
Mean of dependent variable	0.283	0.283	0.283	0.283
Observations	8,230	8,230	8,230	8,230
R- squared	0.135	0.135	0.145	0.145

Note: This table reports coefficient estimates on the impact of approaching and achieving a majority of committee-appointed judges in the judge hiring panel. To focus specifically on the reform multiplier effect, the sample is restricted to the committee-appointed judges, excluding cases adjudicated by judges appointed by the executive. The independent variables turn on when committee appointees join the judge hiring panel. Comparisons are made with decisions involving committee-appointed judges, where the judicial selection committee is predominantly composed of executive- versus committee-appointed judges. In the first two columns, the judge hiring panel still has a minority of committee-appointed judges. In contrast, in Columns 3 and 4, the judge hiring panel just attains a majority of committee-appointed judges. Controls include case type fixed effects (e.g., rights dispute or property dispute), number of pages of judgment order, presence of chief justice on the bench, number of lawyers on the case, number of judges on the case, and total judges on the bench. The court-by-bench or district-by-bench fixed effects, year-fixed effects, and bench-specific linear trends are always included. Standard errors are reported in brackets and are clustered at the court-by-bench level. ***p<0.01, ** p<0.05, * p<0.1.

Table 3: Reform's Impact on Concentration within the Judicial Elite

Panel A: Impact on Same Law Firm Hiring (using all data)

	Same Law Firm Hire			
	(1)	(2)	(3)	(4)
Executive Appointments Vs 1st Degree	0.167***	0.120***	0.0405	0.0530*
	[0.0186]	[0.0188]	[0.0284]	[0.0284]
Executive Appointments Vs 2nd Degree	0.658***	0.629***	0.106**	0.156***
	[0.0305]	[0.0372]	[0.0406]	[0.0390]
Court-by-Bench and Year FE	No	No	Yes	Yes
Controls	No	Yes	Yes	Yes
Court-by-Bench Linear Trends	No	No	No	Yes
Mean of dep variable	0.334	0.334	0.334	0.334
R-squared	0.198	0.258	0.336	0.344
Observations	18,578	18,578	18,578	18,578
F-Statistics (1st Degree = 2nd Degree)	404.110	268.540	7.140	19.180
P -Value (1st Degree = 2nd Degree)	p-value <0.01***	p-value <0.01***	p-value <0.01***	p-value <0.01***

Panel B: Impact on Same Law Firm Hiring (using 1st and 2nd Generation judges data)

	Same Law Firm Hire			
	(1)	(2)	(3)	(4)
1st Degree vs 2nd Degree	0.490***	0.510***	0.106***	0.0875***
	[0.0277]	[0.0345]	[0.0259]	[0.0286]
Court-by-Bench and Year FE	No	No	Yes	Yes
Controls	No	Yes	Yes	Yes
Court-by-Bench Linear Trends	No	No	No	Yes
Mean of dep variable	0.522	0.522	0.522	0.522
R-squared	0.202	0.310	0.442	0.450
Observations	7,470	7,470	7,470	7,470

Note: The table displays the estimated reform multiplier effect on judicial hiring. Specifically, in Panel A, the coefficients α and β are estimated from equation (1), focusing on "Same Law Firm Hire" as the dependent variable. This binary variable is assigned a value of one when the adjudicating judge was previously employed by the same law firm as at least one judge on the hiring panel. Executive Appointment Vs. 1st Degree is a variable that activates with the arrival of the first committee-appointed judge at the bench (one degree of separation from the Executive Appointments). The Executive Appointment Vs. 2nd Degree is a variable representing the arrival of the second-generation committee-appointed judge at the district bench (two degrees of separation from the Executive Appointments). Specifically, the variable switches on when the first second-degree committee-appointed judge arrives at the district bench. The comparison of judicial decisions by committee appointees is made relative to cases decided by the judges appointed by the executive branch in both cases. Controls include case type fixed effects (e.g., rights dispute or property dispute), number of pages of judgment order, presence of chief justice on the bench, number of lawyers on the case, number of judges on the case, and total judges on the bench. The estimates from the baseline equation (1) with court-by-bench or district-by-bench fixed effects, year-fixed effects, and bench-specific linear trends are reported in Column 4. Standard errors are reported in brackets and are clustered at the court-by-bench level.***p<0.01, ** p<0.05, * p<0.1.

Table 4: Impact of First and Second-Generation Committee Appointees on State Wins, Case Merits and due Process Followed Controlling for Same Law Firm Judge Hires

	State Wins	*Case Merits*	*Process Followed*
	(1)	(2)	(3)
Same Law Firm Hire	-0.0124	-0.00952	0.0223
	[0.0163]	[0.0161]	[0.0498]
Executive Appointments Vs 1st Degree	-0.0890*	0.0863*	0.317**
	[0.0500]	[0.0437]	[0.132]
Executive Appointments Vs 2nd Degree	-0.333***	0.306***	0.839***
	[0.0657]	[0.0657]	[0.243]
Executive Appointments Vs 1st Degree X	0.00210	0.0265	0.0547
Same Law Firm Hire	[0.0265]	[0.0241]	[0.0730]
Executive Appointments Vs 2nd Degree	-0.00328	0.0269	-0.0107
X			
Same Law Firm Hire	[0.0368]	[0.0483]	[0.194]
Court-by-Bench and Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Court-by-Bench Linear Trends	Yes	Yes	Yes
Mean of dependent variable	0.443	0.679	3.400
R-squared	0.166	0.125	0.085
Observations	18,578	18,578	18,578
F-Statistics (1st Degree = 2nd Degree)	24.450	18.100	6.330
P -Value (1st Degree = 2nd Degree)	p-value < 0.01***	p-value < 0.01***	0.011**

Note: This table reports the estimated reform multiplier effect on pro-government rulings, case merits and rating on due process followed with the same law firm hired judges controlled for in the regression. Same law firm hire is a dummy variable that takes a value of one when the case is decided by a judge hired from the same law firm as at least one of the judges on the judge hiring panel. Specifically, it presents estimates of coefficients α and β from equation (1) with the state victory variable in Column 1, the case merit variable in Column 2, and the process followed variable, which assesses the adherence to due process in the case in Column 3. Executive Appointments Vs. 1st Degree is a variable that activates with the arrival of the first-generation appointed judge at the bench (one degree of separation from the Executive Appointments). The Executive Appointments Vs. 2nd Degree is a variable representing the arrival of the second-generation committee-appointed judge at the bench (two degrees of separation from the Executive Appointments). It switches on when the first second-degree committee-appointed judge arrives at the district bench and remains one. The comparison of judicial decisions by committee appointees is made relative to cases decided by the judges appointed by the executive in both cases. Controls include case type fixed effects (e.g., rights dispute or property dispute), number of pages of judgment order, presence of chief justice on the bench, number of lawyers on the case, number of judges on the case, and total judges on the bench. The court-by-bench or district-by-bench fixed effects, year-fixed effects, and bench-specific linear trends with controls are included in all columns. Standard errors are reported in brackets and are clustered at the court-by-bench level. ***p<0.01, ** p<0.05, * p<0.1.

Table 5: Impact of First and Second-Generation Committee Appointees on Case Merits and Process Followed

	Case Merits		*Process Followed*	
	(1)	(2)	(3)	(4)
Executive Appointments Vs. 1st Degree	0.113***	0.0946**	0.361***	0.336**
	[0.0412]	[0.0441]	[0.127]	[0.131]
Executive Appointments Vs. 2nd Degree	0.330***	0.317***	0.847***	0.820***
	[0.0581]	[0.0628]	[0.245]	[0.248]
Court-by-Bench and Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Court-by-Bench Linear Trends	No	Yes	No	Yes
Mean of dependent variable	0.679	0.679	3.400	3.400
R-squared	0.117	0.124	0.076	0.085
Observations	18,578	18,578	18,578	18,578
F-Stats. (1st Degree = 2nd Degree)	24.120	21.620	7.420	7.860
P -Value (1st Degree = 2nd Degree)	p-value <0.01	p-value <0.01	p-value <0.01	p-value<0.01

Note: This table explores the reform multiplier effect on government victories. Specifically, it presents estimates of coefficients α and β from equation (1) with the case merits and ratings on the process followed as the dependent variables. Executive Appointment Vs. 1st Degree is a variable that activates with the arrival of the first committee-appointed judge at the bench (one degree of separation from the Executive Appointments). The Executive Appointment Vs. 2nd Degree is a variable representing the arrival of the second-generation committee-appointed judge at the district bench (two degrees of separation from the Executive Appointments). Specifically, the variable switches on when the first second-degree committee-appointed judge arrives at the district bench. The comparison of judicial decisions by committee appointees is made relative to cases decided by the judges appointed by the executive branch in both cases. Controls include case type fixed effects (e.g., rights dispute or property dispute), number of pages of judgment order, presence of chief justice on the bench, number of lawyers on the case, number of judges on the case, and total judges on the bench. The baseline equation with court-by-bench or district-by-bench fixed effects, year-fixed effects, and bench-specific linear trends and controls is reported in Columns (2) and (4). Standard errors are reported in brackets and are clustered at the court-by-bench level. ***p<0.01, ** p<0.05, * p<0.1.

Table 6: Comparing Judge Characteristics of First-Generation versus Second-Generation Committee Appointees

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Elite Law School	Age at Appointment	Judge Tenure	Male	Former Office Holder Bar Association	Prior Political Office	Elevated To Supreme Court
1st Degree Vs. 2nd Degree	0.0910***	0.744***	0.311	-0.0110	0.0145	0.0272	-0.0128
	[0.0279]	[0.274]	[0.198]	[0.0174]	[0.0446]	[0.0268]	[0.0166]
Observations	7,456	7,456	7,456	7,456	7,456	7,456	7,456
R-squared	0.456	0.394	0.286	0.090	0.141	0.096	0.218
District by bench and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bench Specific Linear Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Variable	0.541	51.21	3.676	0.948	0.448	0.163	0.0681
				Mediating Variables			
	Elite Law School	Age at Appointment	Judge Tenure	Male	Former Office Holder Bar Association	Prior Political Office	Elevated To Supreme Court
1st Degree Vs. 2nd Degree							
Proportion of Effect Explained	0.396***	0.052	0.071	-0.064	-0.009	0.010	-0.022
by the Mediating Variable	[0.102]	[0.069]	[0.057]	[0.040]	[0.018]	[0.011]	[0.108]
Observations	6,232	6,232	6,232	6,232	6,232	6,232	6,232

Note: Panel A presents the observed characteristics of first and second-generation committee appointees. The *1st Degree vs 2nd Degree* Judge is a dummy variable that switches to one for cases involving second-generation judges and is zero for first-generation judges. The comparison is made between 1st Generation judges, predominantly selected by an executive-appointed judge hiring panel, and 2nd Generation Judges, selected by committee-appointed judges' hiring panel in the majority. Executive-appointed judges are excluded. Each column corresponds to a different dependent variable. For Panel A, in Column (1), the dependent variable is *Elite Law Schools*, a dummy variable equal to one if the judge graduated from one of Pakistan's top 10 law schools and zero otherwise (according to a ranking of best law schools, https://bestlawschools.net/top-law-schools-in-pakistan/). In Column (2), the dependent variable is *Age at Appointment*, representing the age in years of the judge at the time of appointment. In Column (3), the dependent variable is *Judge Tenure*, which represents the time in years of the judge on the bench (judge experience). In Column (4), the dependent variable is *Male*, a dummy variable that switches to one for male judges and is zero for female judges. In Column (5), the *Former Officeholder Bar Association* is a binary indicator for judges who held office at a local or regional bar association prior to their judgeship. In Column (6), the dependent variable is *Prior Political Office*. This is a binary variable where the judge had ever held a political office before elevation to the High Court bench. In Column 7, the dependent variable is *Elevated To Supreme Court*, which switches to one when a High Court judge is promoted to a judgeship at the Supreme Court of Pakistan and zero otherwise. Panel B provides a causal mediation analysis of the variables listed in the corresponding columns. The mediation analysis estimates the extent to which the observed characteristics mediate the relationship between arrival of the second-generation judge and pro-government rulings. The corresponding fraction explained by the mediating variable is reported. Following the methodologies of Heckman and Pinto (2015) and Nguyen et al. (2021), we use the Stata 18 command 'mediate' for these calculations. The results showcase the percentage reduction in pro-government rulings attributable to each variable in the columns. For instance, Column (1) indicates that 39.6% of the reduction in pro-government rulings can be attributed to the variable representing elite law school attendance by second-degree judges. Controls include case type fixed effects (e.g., rights dispute or property dispute), number of pages of judgment order, presence of chief justice on the bench, number of lawyers on the case, number of judges on the case, and total judges on the bench. The baseline equation with court-by-bench or district-by-bench fixed effects, year-fixed effects, and bench-specific linear trends and controls are included in all columns. Standard errors are reported in brackets and are clustered at the court-by-bench level. ***p<0.01, ** p<0.05, * p<0.1.

For Online Publication

Reform Multiplier

By Sultan Mehmood and Bakhtawar Ali

Contents

A. Variable Definitions, data information and Robustness Checks

A1. Variable Definitions

A2. Data Construction and Sources

A3. Pakistan's Court System

A4. Alternative Explanations

A5. Robustness and Discussion

B. Appendix Figures and Tables

A. Variable Definitions, data information and sources

Appendix A1. *Variable Definitions and Sources*

State Wins = The government victory variable is a binary indicator, with one representing a State victory and zero indicating a State defeat. Two independent teams of 5 attorneys each code this variable by examining judgment texts. 97% of the judgments received identical coding from both teams of attorneys. Legal professionals can partially access the text data via an online portal (https://www.pakistanlawsite.com/). We combine this with confidential text data retrieved from High Court case archives from the Registrar's Office. This allowed us to access and randomly sample the universe of High Court judgments. Additional information on the data construction can be found in Appendix A2.

Merit = This variable indicates whether a case was decided based on case merits or evidence rather than technical or procedural grounds. The two teams of attorneys also coded it at the case level by the two teams of attorneys. Specifically, it is a dummy variable that takes the value of one, indicating a case decided on case merits and zero otherwise. This case-level variable was constructed similarly to the government victory variable based on the attorneys' reading of the text of judgment orders. Legal professionals can access the judgment text data through an online portal (https://www.pakistanlawsite.com/). This is supplemented by confidential text data retrieved from High Court case archives at the Registrar's Office, enabling us to access and randomly sample the universe of High Court judgments. Further details on data construction are available in Appendix A2.

Process Followed = This discrete variable rates the extent to which "all relevant jurisdictional, procedural, and evidential requirements were followed in reaching the judicial decision" on a scale from 1 to 5. A higher rating indicates that higher jurisdictional, procedural, and evidential standards were adhered to while making the judicial decision. The rating is also based on assessments made by attorneys, derived from their reading of the text of the judgment order.

Same Law Firm Hire = This is a binary variable that takes the value of 1 if the new judicial hire had previously worked at the same law firm as one of the judges from the judge hiring panel's (previous) law firm. In other words, it indicates whether the newly appointed judge had prior employment at the same law firm as at least one of the judges on the committee responsible for appointing new judges. If the new judicial hire had such previous employment, the variable is coded as 1; otherwise, it takes the value of 0. This data is sourced from a combination of administrative records at the High Court registries and the Law and Justice Commission of Pakistan (LJCP). The independent variables turn on when committee appointees join the judge hiring panel.

Executive Appointed Vs. 1st Degree = This dummy variable transitions to a value of one when the first committee-appointed judge joins the judicial bench. The data for this variable is sourced from the Registrar Offices of High Courts and the Law and Justice Commission of Pakistan.

Executive Appointed Vs. 2nd Degree = This dummy variable switches to one on the arrival of the first judge at the judicial bench shifts the composition of the judge hiring panel to a majority of committee-appointed judges. This data is sourced from the Registrar's Office of High Courts and the Law and Justice Commission of Pakistan.

Committee Appointees Majority 4-3 = This binary variable activates when a judge's arrival shifts the judge-hiring panel to a majority of committee-appointed judges compared to executive-appointed judges. It turns to one when the arrival of the new judge shifts the judge hiring panel to a 4:3 majority within the 7-panel judge hiring committee. Figure 1 visually depicts the judge hiring authority's evolution in selecting first and second-generation committee appointees. The data is compiled from High Court and Supreme Court registries in Pakistan.

Elite Law School = This binary variable takes the value of 1 if the judge was educated at one of Pakistan's top 10 law universities in Pakistan and zero otherwise. The ranking is taken from a global ranking of top law schools worldwide, and we particularly retrieve the "Best Law Schools in Pakistan". For a complete list and methodology behind the rankings, see https://bestlawschools.net/top-law-schools-in-pakistan/

Left-Party or Progressive Group Lawyers = This is a dummy variable that switches on if the judge had previously run bar elections on the Progressive Group party ticket and zero otherwise. The progressive group acts as a union of lawyers and prioritizes principles of social justice, human rights principles, and legal reforms, such as greater representation of women and minorities at the bar and the bench. Data for this is primarily sourced from bar association records and judicial gazettes of the High Courts.

Right-Party or Professional Group Lawyers = This dummy variable depends on whether the judge had previously run bar elections on the Professionals Group party ticket. The professional group runs for bar unions on the principles of greater focus on the interests of lawyers and businesses and calls for a reduction in regulation and more business-friendly reforms. Data is also sourced from bar association records and judicial gazettes.

Number of Lawyers = A count variable for the number of lawyers arguing the particular case. This case-level variable is also retrieved from information in judgment texts. Two teams of 5 attorneys independently code this variable by examining judgment texts, and there is 100% agreement in coding for the number of attorneys between both teams.

Number of Judges = A count variable for the number of judges adjudicating a particular case. This case-level variable is also retrieved from judgment texts. Two teams of five attorneys each code this

variable by analyzing judgment texts. Both teams provided identical coding for the number of judges in 100% of the judgments.

Bench Chief Justice = A dummy variable for the High Court Chief Justice adjudicating in a particular case. This case-level variable is also retrieved from judgment texts. Two teams of 5 attorneys independently code this variable by examining judgment texts, with 100% agreement on whether the case involved the Chief Justice.

Number of Pages of Judgment Orders = A count variable for the number of pages of the judgment order for a particular case. This case-level variable is also retrieved from judgment texts. Two teams of 5 attorneys also independently code this variable with 100% agreement.

Criminal Case Type = A dummy for criminal cases, as indicated in the text of the judgment order. This case-level variable is also retrieved from judgment texts. Two teams of 5 attorneys also code this variable in judgment texts independently, achieving 100% agreement.

Constitutional Case = A dummy for constitutional cases, as indicated in the text of the judgment order. This case-level variable is also retrieved from judgment texts. The same two teams of 5 attorneys also independently code this variable from judgment texts with perfect agreement.

Gender = This variable is a dummy for male judges and was coded in two ways. Firstly, the authors manually checked every judge's name. Secondly, we used Stata to read the strings from administrative data on judge full names starting with "Justice Miss" and "Justice Mrs." as zero (indicating female judges) and the string starting with "Justice Mr." as one (indicating male judges). Both methods yielded an identical number of male and female justices.

Elevated to SC = This dummy variable denotes whether a judge ascended to the Supreme Court, turning on if the High Court judge was promoted and remaining zero otherwise. The information is retrieved from administrative data maintained at the Supreme Court of Pakistan and accessible in the Judicial Gazettes of Pakistan.

Former Lawyer = This dummy variable indicates whether the High Court justice was an attorney, typically true unless they were a former lower court judge. Data sources to construct this variable include annual reports, Bar Council records, and judicial administrative data at High Courts' Registrar offices.

Former Office Holder Bar Association = A dummy variable indicates whether the judge held the position of secretary-general or president in the lawyers' bar association before becoming a High Court judge. Data is sourced from biographical information on bar association websites, offices, and judicial administrative records at High Courts' Registrar Offices.

Former Judge = A dummy variable indicates whether the judge had prior service as a lower court judge (civil, session, or additional session judge). Data sources to construct this variable include

biographical information from annual reports submitted to High Courts, Law and Justice Commission of Pakistan Reports, and judicial administrative records at High Courts' Registrar Offices.

Appendix A2. *Data Construction and Sources*

We randomly sample 20,000 High Court cases in Pakistan from the years 1986 to 2022, constituting 0.3% of the total available cases involving the State. Figure B1 in Appendix B illustrates the evolving nature of the random sample relative to the total universe of cases every year. With the exception of a decline in case filings during the COVID-19 lockdown, there is an observable gradual increase in case filings each year. To be included in our sample, a prerequisite was that the government was one of the parties in the case. The data recorded in the judgment order texts was coded by two teams, each comprising four lawyers under the supervision of a senior attorney (two teams of 5's). Notably, 97% of the judgments share identical coding between both attorney teams, ensuring consistent results regardless of the specific team's coding used. The data was acquired through a contractual agreement with a reputable law firm. The data acquisition involved accessing information from the Pakistan law site portal and utilizing the provided data from the High Court registry. The task assigned to the firm was to code the variables outlined in the preceding section based on their analysis of judgment texts in a team of two with 5 legal experts each.. Part of this data is available partially online at https://www.pakistanlawsite.com/. The website serves as a central repository of cases and is commonly used by lawyers in Pakistan for case preparation for those "Approved for Reporting" by judges. It is a password-protected site accessed through a law firm and is considered a "Central Library" for lawyers practicing within Pakistan's (Common) Law system, where case precedent is highly valued. To ensure a more representative sample, we supplemented the online data with archival data from the High Court's Registrar Offices and record rooms of the High Courts. These archives provide access to cases that may not have been "approved for reporting" by the judges. This confidential judgment text data ensures we have a representative sample. Specifically, the task of recording key information from the judgment texts was carried out by two independent teams of four lawyers each, supervised by a senior lawyer (team of 5). A formal contract was in place to ensure the independence of the teams, and random field visits were conducted to verify compliance with the contract. The two senior lawyers supervising the teams were experts in constitutional law, with over 10 and 12 years of experience as practicing attorneys at the High Court. Their expertise made them well-suited to oversee the coding process, especially for cases involving the state. The remaining eight individuals involved in the coding exercise were legal experts who had all passed the "L.L.M" (called L.L.B in Pakistan) examination, which is equivalent to a two-year master's degree in law. This qualification is a prerequisite for practicing law as an attorney in Pakistan. These coders were essentially junior attorneys at the law firm, and they were not provided with any information about the

research question. Implementing this data collection approach, with coding conducted by two independent teams, initially introduced by Djankov et al. (2003) and influencing a body of literature (see, e.g., Epstein and Segal, 2000; Choi et al., 2013; Aney et al., 2021), ensures the reliability of the coded information for our research study. The dataset on judicial appointments, retirements, affiliations with law firms, and promotions data is curated from diverse sources, encompassing High Court registries, bar association records, Judicial Gazettes of Pakistan, and confidential but mandatory reports submitted to the Law and Justice Commission of Pakistan (LJCP) by each regional High Court. These data combine to delineate judges' past, present, and anticipated career paths, enabling the construction of a network of judicial elites and tracing of judges' entire trajectories, encompassing their journeys from law schools they attended, affiliations with law firms when they were attorneys at to their elevation to judgeships.

Appendix A3. *Pakistan's Court System*

This section provides an overview of the structure and background of the courts in Pakistan. The foundation for the current High Courts in India and Pakistan can be traced back to the Indian High Courts Act of 1861, which granted the British Crown the authority to establish High Courts in the Indian colony. After gaining independence from British colonial rule in 1947, both India and Pakistan introduced various modifications to their legal institutions while maintaining the overall institutional structure, including Common Law jurisprudence, civil and criminal procedures, and the three-tier system of Trial, High, and Supreme Courts. One notable change was the increase in the mandatory retirement age for High Court judges from 60 to 62. India raised the retirement age in 1963, followed by Pakistan in 1969 through constitutional amendments, and it has remained at 62 ever since.

While our study focuses on the Pakistani High Court, it is essential to note that Pakistan's judiciary operates under a three-tier hierarchical system. At the lowest level, there are civil and session courts that handle civil and criminal cases. The High Courts have the authority to review decisions made by these lower courts. In cases involving government, for instance, land expropriation or violations of fundamental rights, the High Court is often the primary and sometimes the only recourse for individuals and firms seeking redress. Cases are assigned to judges within the High Court through a case management system. Although there are only four provincial High Courts and one federal High Court in Islamabad, the High Court system is dispersed across the provinces through 16 regional High Courts or "divisional" benches. Each region has about four judicial benches specializing in property, writ or human rights, tax, and criminal cases. Hence, there are 64 judicial benches across Pakistan according to these four specializations. It is worth noting that Pakistani laws and capacity constraints impose limitations on replacing judicial vacancies at the bench level. In most cases, replacements are determined by mandatory retirements (for more details, see Mehmood, 2022).

High Courts hold significant importance in Pakistan as they provide a platform for individuals to file cases against the government, including constitutional petitions (writ petitions) or criminal appeals, with the State being a party. Constitutional cases involving the State can be filed against the federal, provincial, or local governments or any organ of the state that holds executive authority, such as the office of the Prime Minister. Finally, the Supreme Court of Pakistan, located in the federal capital of Islamabad, serves as the final appellate court. It hears appeals on "technical" grounds for criminal and constitutional cases from the High Courts. The Supreme Court has a minimal number of judges (maximum 17), which restricts the number and scope of cases it can hear. As a result, only a tiny proportion of cases are heard by the Supreme Court.

Appendix A4. Alternative Explanations

In the supplementary text of this appendix, informed by contemporary scholarship, we present evidence that challenges a range of potential alternative factors that could account for the observed reform multiplier effect. We scrutinize various mechanisms including the influence of judicial ideology, career incentives for judges, peer influences or spillover effects, adherence to case law precedents, strategic actions by litigants, and the possibility of corporate or legal profession influence. For each of these channels, we provide a thorough discussion and evidence that counters their role in explaining our findings, detailed below.

Reform Compounding and Judge Ideology. Does the compounding effect of the reform correlate with judge ideology? In order words, are the second-degree committee appointees, with a history of right-wing bar political involvement, more likely to rule against the government compared to their first-degree counterparts, perhaps due to their conservative leanings? A body of literature from the United States court system has argued that judges' decisions are influenced by their ideological leanings on the political spectrum (e.g., Epstein et al., 2007 or Bonica and Sen, 2021 for a review of this literature). Bar politics within Pakistan has been rich and can be traced back to the 1950s, evolving as a platform for lawyers to advocate for their rights and pursue a left or right political agenda. Two prominent factions of lawyers dominate bar politics in Pakistan, members of which are often elevated to judicial positions. These are the *Professionals Group* (right-wing) and the *Progressives Group* (left-wing). These groups represent contrasting ideologies within the legal community, with the Professional Group championing policies favoring deregulation, reduced taxation, and a lower level of government involvement. In contrast, the Progressives Group (also colloquially known as Asma Jehangir’s Group, named after the legendary civil rights activist who headed this group for many decades) focuses on issues of social justice and more progressive ideas, particularly the safeguarding

of minorities, worker rights, and women rights.[27] We build a unique dataset to link judges in our sample who ran for elections as attorneys for these right versus left-wing lawyers' groups, leveraging bar association records. While this data is somewhat incomplete, which does not allow us to match all 720 judges in our sample, we are able to match more than 50% of judges in our dataset (about 400 judges) with one of these two political factions of lawyers. We deem a judge to be aligned with either of these factions if they have contested in bar elections representing one of these groups as an attorney. Appendix B, Table B7 presents the reform multiplier effect obtained by estimating the main Equation (1) on a sub-sample of judges classified as right or left-wing according to their revealed ideology in bar politics. The reform multiplier effect is evident in both left- and right-leaning judges, indicating that this effect transcends judicial ideology.[28] Nonetheless, marginally significant point estimates (p-value = 0.109) suggest that judges with early career ties to the more ring-leaning "professional" group show a greater reform multiplier impact compared to those linked to the left-leaning bar faction, implying that right-affiliated second-generation judges tend to rule against the government more often than their left-affiliated counterparts.

Reform Compounding and Career Concerns. Building on the recent work of Bertrand, Burgess, Chawla, and Xu (2020), which underscores the relevance of career considerations within bureaucracies, we explore this potential channel for the case of Pakistani judges, who, like many bureaucrats across the world, have job security and legislative safeguards against political interference. Prior research, as highlighted by Finan et al. (2017), contends that the promise of promotions may be a critical factor in shaping the choices of public officials. Consequently, it is plausible to consider whether the differential behavior observed between first- and second-generation judges may stem from the varying degrees to which they are exposed to career incentives. Therefore, we examine the prospects of elevation to the Supreme Court for first- and second-generation judges, investigating if the reduced frequency of pro-government decisions by second-generation judges correlates with their increased chances of promotion. Column 7 of Table 6 reports these results, which suggest that first- and second-generation judges are similarly likely to be promoted to the Supreme Court, indicating that career incentives for promotions are not likely the primary explanation for the reform multiplier effect.

Reform Compounding and Peer Effects. Beyond career advancement incentives, there exist alternative pathways through which judges' incentives may explain the reform multiplier effect. For instance, the induction of second-generation judges appointed by committees may create spillover or

[27] There is also an "Independent Group" of lawyers that sometimes also field candidates at lawyer's bar elections, however, they seldom pose a challenge to the two mainstream parties at the bar elections.

[28] This aligns with our previous findings in Table 6, indicating no strong correlation between the reform multiplier effect and the overall political engagement of judges before their judgeships. This includes their decision to run for any bar elections or engage in national politics, as shown in Table 6 (Column 5), and neither factor mediates the reform multiplier effect.

peer effects that sway the rulings of first-generation judges. Table B8 of Appendix B evaluates whether this is the case. The arrival of the second-generation committee appointees does not appear to affect the first-generation judges' pro-government rulings (Columns 1 and 2), decisions on case merits (Columns 3 and 4), or ratings on adherence to due process of law (Columns 5 and 6). A higher fraction of second-generation judges on the judicial bench has no discernible impact on the judicial decisions of first-generation judges. This has two important implications. Firstly, the evidence suggests that any peer or spillover effects from the rulings of second-generation judges on first-generation judges are minimal and possibly even non-existent. This aligns with the previously observed abrupt change in judicial decisions that coincides with a shift in bench composition from a majority of executive appointees to committee appointees in the judicial hiring committee. Secondly, the results of Table B8 also provide support *against* violation of the *Stable Unit Treatment Value Assumption* (SUTVA). This is because the behavior of the first-generation judges (control group) is unaffected by the arrival of the second-generation judges (treated group), which could have violated the SUTVA assumption.

Reform Compounding and Precedent Compliance. Moreover, the results from Table B8 also imply that our findings are unlikely to be ascribed to greater precedent adherence by first-generation judges. If following precedents laid out by second-generation judges by first-generation judges were a key channel behind our results, the arrival of the second generation of judges would have prompted the first generation of committee-appointed judges to align their decisions to be more anti-government. However, we did not observe this to be the case. Several additional patterns in the data also do not support this idea. If more judges followed case precedents over time, we would expect to see (i) gradual changes reflected in our analysis of annual frequency data in the estimation of event studies (ii) a sharp change in coefficient estimates with the inclusion of court-specific linear trends. However, the pronounced declines in pro-government rulings observed in our event-study estimates (Figure 3 Panel A), along with the stability of point estimates after accounting for district-by-year fixed effects and including court-by-bench linear trends (Table 1), imply that greater adherence to case precedents is unlikely to be the primary factor driving our findings.

Strategic Efforts and Case Composition. Another avenue we explore involves delving into whether the reform compounding is explained through changes in strategic efforts exerted by lawyers on the case or decisions of litigants to file a case. It may be that the reform multiplier effect is mediated through lawyers exerting differential effort in cases in front of first- versus second-generation committee appointees or litigants filing more cases when they know the judge is a second-generation versus first-generation appointee. This inquiry builds on the seminal work of Klein and Priest (1984), who contend that the decision to bring cases to court and the ex-post effort exerted by lawyers is a strategic choice, so judicial reforms may influence rulings by changing the effort by lawyers or changing the type of cases that appear in court. We, therefore, explore this potential strategic behavior

exhibited by both the prosecution and defense lawyers. First, building on our discussions with legal professionals in Pakistan, we investigate whether the number of lawyers for both the prosecution and defense changes with the appointment of first- or second-generation judges. The attorney count reflects the State's increased effort in high-stakes cases, supported by qualitative evidence of hiring more lawyers in high-stakes cases (Arshad, 2023). Second, the litigants may strategically opt to file cases, providing a potential mechanism for our observed reform compounding effect, especially over the long term (since, in the short term, the docket of cases that appear before the judge is unlikely to change much considering the severe court congestion in Pakistan).[29] Table B9 in Appendix B presents the results of estimating the baseline Equation (1) with the number of lawyers involved in the case (Columns 1 and 2) and the total number of case filings (Columns 3 and 4) as dependent variables. The number of attorneys representing the defense or the prosecution remains unaffected by the arrival of first and second-generation judges to the bench, nor is there any evidence of a reform multiplier effect on lawyers' efforts (Table B9's Columns 1 and 2). In line with these findings, the total number of cases initiated by both defense and prosecution also does not suggest a reform multiplier effect (Columns 3 and 4). However, an increase in cases filed by the defense against the State was observed following the arrival of first-generation judges relative to executive appointees (Table B9's Columns 3 and 4). Taken together, the rise in the number of lawyers hired or cases filed does not statistically vary for first-generation versus second-generation committee-appointed judges, making this mechanism alone unlikely to explain the reform multiplier effect. This aligns with recent research in Pakistan, where Acemoglu et al. (2020) document that litigants in Pakistani courts maintain similar case-filing behavior despite large changes in court delays. The authors attribute this to severe capacity and informational constraints in courts, hindering litigants from promptly adapting to changes in the court system. Nevertheless, in Columns (5) and (6), we report the stability of our regression analysis at the judge level, as recommended by Lambais and Sigstad (2023), in the context of strategic case filings. They argue that the random assignment of judges within constitutional courts mitigates the conventional concerns associated with Klein-Priest strategic case filings if the results are robust at the judge-level regressions. Overall, the results from Table B9 suggest that the reform compounding is unlikely to be explained by changes in attorney efforts or strategic case filings.

Corporate Capture. Scholars have identified corporate capture of the judiciary, with case studies in China (Wang, 2018) and Russia (Viktorov, 2019) offering suggestive evidence. Specifically, there is an underlying threat that judicial independence from the government may trade off against elevated

[29] Likewise, over time, the first- and second-generation judges are also likely to preside over similar cases because of the high costs associated with forum shopping compounded by legal deterrents, such as the clear precedents set in cases such as Sharaf Faridi v. Federation of Islamic Republic of Pakistan (PLD 1989) and Government of Sindh v. Sharaf Faridi (PLD 1994). This is consistent with the evidence presented earlier in Table B3.

dependence on large corporations. To investigate this channel, we narrow our focus to a subset of cases involving the state versus firms, differentiating between instances where the government opposes small and medium-sized firms (SOEs) and cases involving the government opposing large corporations. In the latter scenario, we consult a list of the 12 largest business conglomerates in Pakistan (Malik, 2022) and scrutinize whether firms belonging to these conglomerates appear as litigants in our sample of cases. The results of this analysis can be found in Table B10 of Appendix B. Intriguingly, many cases against the state and these firms affiliated with 12 conglomerates appear in our dataset. However, mirroring our main results, a similar reform multiplier effect surfaces in cases involving small or medium-sized enterprises, with lower pro-government rulings for second-generation versus first-generation committee appointees. However, upon closer examination of pro-government rulings against large corporations, there is a clear reform multiplier effect. In this scenario, the probability of the government prevailing against large corporations is, if anything, higher for committee appointees versus executive appointees, as shown in Table B10, Column 4 (although the coefficients lack statistical significance). Overall, evidence from Table B10 reveals that the judiciary's increased autonomy from the government does not necessarily trade off against the capture of the judiciary by large corporations.

Lawyers' Capture. Recent qualitative accounts have raised concerns about the potential role of lawyers in capturing the judiciary in Pakistan, suggesting that instead of fostering judicial independence, they may impede the rule of law (as discussed in a case study by Khan, 2019). Our findings, indicating increased hires from the same law firms, raise the possibility of a potential trade-off from greater independence from the government to greater dependence on private law firms. This suggests a possible risk of a few dominant law firms disproportionately swaying judicial decisions. To examine this possibility, our analysis concentrates on cases involving judges and attorneys with connections to the same law firm, specifically where both parties have shared affiliations. We identified approximately 1000 cases involving judges who previously worked at law firms and attorneys practicing from those firms. The findings are detailed in Table B11, Appendix B. Despite a reduced sample, our baseline results continue to hold a significant reform multiplier effect on pro-government outcomes, and decision-quality measures persist even when judges preside over cases involving attorneys from their former law firms.

Appendix A5. Robustness and Discussion

In this section, we illustrate the robustness of our findings by conducting a series of sensitivity and identification checks, subsequently concluding with a discussion on the external validity of our research.

Case Composition and Case Assignment We initiate the robustness checks of the results by examining whether the reform compounding is explained by the change in the types of cases that appear before the court. To explore this aspect, we reevaluate whether the reform compounding is driven by changes in the composition of cases brought to court. Upon reestimating Equation (1) with case characteristics as the dependent variables, the evidence, as presented in Table B3, Columns 1 to 9, suggests that the treatment variables have no discernible effect on the observed case characteristics. We also implement a more rigorous test for joint dependence, as recommended by Bruhn and McKenzie (2009), that jointly evaluates the interplay between all observed case characteristics and treatment variables, as shown in Table B3, Columns 10 and 11. The results from both tests lend empirical support to the idea that changes in case composition are unlikely to explain the reform multiplier effect. The stability in the types of cases observed is likely due to the substantial backlog in the judiciary, which limits the potential for drastic alterations in judges' dockets in the short run. This situation is compounded by the High Court's approach to quasi-randomly assigning cases, a process that is further supported by legal deterrents, including precedents set like those set in Sharaf Faridi v. Federation of Islamic Republic of Pakistan (PLD 1989). Figures B6 to B8 in Appendix B also illustrate the similarity in the distribution of case types among judges, providing visual evidence that supports the random assignment of cases. This consistency in case allocation per judge aligns with the expected balance that a random assignment of cases to judges would produce.

Time Interactions and Fixed Effects. Subsequent to conducting the balance tests, we estimate more demanding specifications that allow for tighter treatment-control comparisons. While the balance of treatment over observed case characteristics is reassuring, and the inclusion of bench-by-district fixed effects helps address level differences across judicial benches, time-varying court-by-bench omitted factors across treated and untreated benches may still influence pro-government rulings. In this subsection, we further delve deeper into this identification concern by extending our difference-in-differences approach to incorporate a broader range of fixed effects and time-interacted baseline characteristics, such as time interactions with all available baseline case and judge characteristics and introduction of district × year fixed effects to focus the comparison on benches within the same district court. These 600 fixed effects with time interactions of baseline case and judge characteristics and district-by-bench fixed effects can account for variations such as changes in government that might alter the composition of the judicial selection committee or fluctuations in case congestion within districts over time, ensuring more precise treatment-control comparisons. Results from this exercise reported in Table B12 maintain statistical and qualitative similarity with the baseline estimates in Table 1. The stability of coefficient estimates suggests that any potential omitted variable bias is likely to be small relative to the treatment effect.

Age and State Victories. It might be presumed that the reform compounding effect is an experience differential effect driven by the newly appointed judges having lower pro-government rulings at the start of their career, only to have higher pro-government rulings in their later years. However, this is unlikely for two major reasons. Firstly, the reform multiplier effect represents the difference in rulings between first- and second-generation judges relative to presidential appointees, and to be age-driven, it would necessitate additional age-specific effect for the first or second-generation committee appointees, which by definition would be the reform multiplier effect. Secondly, our analysis of a bare majority versus an actual majority demonstrated that new appointments—when the judge hiring committee approaches but does not reach a majority of committee-appointed judges, which also coincides with the arrival of a new judge—have no discernible impact on pro-government rulings. Specifically, there is no evidence of a reform compounding effect when new committee-appointed judges are elevated in the judge hiring panel, bringing the judicial committee versus executive appointees in the judge hiring panel close to a majority with a 3:4 ratio. However, a substantial reform compounding effect is observed when the committee-appointed judges reach a 4:3 majority within the judge hiring panel. The statistically significant and abrupt decrease in pro-government rulings is only observed for new judges hired by the judge-hiring panel with a majority of committee-appointed judges, not any new judicial arrivals per se. Finally, we can also observe a minimal correlation between judges' age and their pro-government rulings: judges of different ages are similarly likely to issue pro-government rulings (refer to Figure B13 in Appendix B). These findings align with our regression-form results, which indicated that the judge's tenure (duration in office at the time of decision) is uncorrelated with the judge's appointing authority (Table 6, Column 3). Taken together, our identification strategy and empirical evidence suggest that the judge's age does not drive the reform multiplier effect.

Clustering. We further illustrate the robustness of our results with an alternative clustering choice. Following the recent suggestions by Abadie et al. (2023), we employed court-by-bench level clustering as the baseline since this corresponds to the variation in our treatment assignment. This implies that we have 64 judicial bench-level clusters (Figure B12 in a supplement to Appendix B illustrates how our treatment variables vary across clusters, see figure note to Figure B11). Notably, our cluster count exceeds the recommended rule of thumb of 42 clusters for reliable inference using asymptotic theory, as suggested by Angrist and Pischke (2008, p. 219). Nonetheless, to further validate our findings, we employed bootstrap techniques to construct standard errors. Specifically, we implement Wild Bootstrap clustering, as proposed by Cameron, Gelbach, and Miller (2008), which is particularly suitable for situations with a limited number of clusters. The Wild Bootstrapping yields a confidence

curve representing bootstrapped p-values, establishing confidence set limits.[30] Figure B14 of Appendix B (Panel A and Panel B) shows that the Bootstrapped distributions generated for the first- and second-generation judges do not overlap much. The distribution of second-generation committee appointees (Panel B) is situated firmly to the left of the first-generation committee appointees (Panel A), suggesting a more substantial negative impact—the reform multiplier effect. This underscores the robustness of the reform multiplier effect beyond relying solely on asymptotic theory for statistical inference.

Additional Sensitivity Checks. In this subsection, we report a few additional sensitivity analyses that examine the stability of our main results. We first examine whether specific judicial benches drive our results. We explore the extent to which our results are driven by the operations of the specialized benches dealing with criminal, property, tax, or political rights issues. Table B13 presents these results, where we systematically exclude each specialization bench, one at a time with replacement: regardless of the specialized judicial bench, the reform compounding effect is substantial and statistically significant. Similarly, we explore the potential impact of specific regions (referred to as 'divisions' in Pakistan's official terminology) on our findings. To assess this geographic heterogeneity of our results, we systematically exclude each of the 16 regional high courts one at a time and report these results in Table B14 of Appendix B. The results indicate that the reform multiplier is effective throughout Pakistan, affecting all regions or specific regional High Courts. Last, as suggested by Angrist and Pischke (2008), we used linear models since they best approximate the conditional expectation function, even when working with dummy dependent variables. To further validate our findings, we utilize Logit and Probit models for the dummy dependent variables, State Wins, and Case Merits. The corresponding marginal effects from estimating these non-linear models are reported in Table B15 of Appendix B. Notably, both (Logit and Probit) non-linear models produce results similar to the baseline estimation using the linear model.

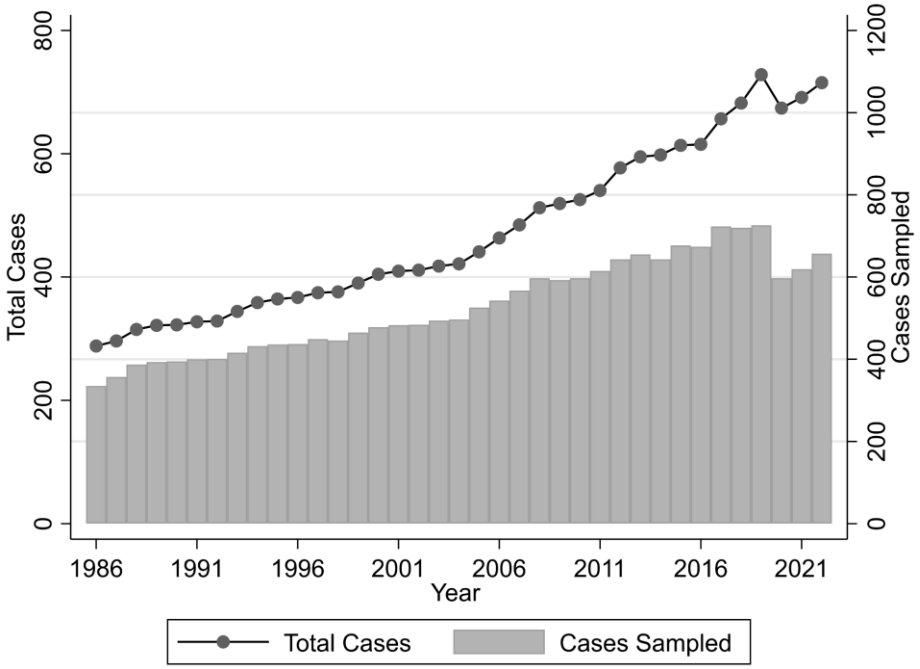
Discussion of External Validity. In the context of the judiciary in Pakistan, its selection procedures, civil and criminal procedural laws, and protocols echo similarities with numerous other developing nations, most notably India and Bangladesh. All three nations inherited the judicial set-up from British Colonial rule in the Indian subcontinent. Collectively, Pakistan, India, and Bangladesh represent more than a quarter of the world's population, and their common inheritance of colonial judicial structures hints at potential external validity for the study. Acknowledging the diversity of settings, List (2020) notes that "all results are externally valid to some setting, and no result will be externally valid to all settings." In our discussion on external validity, we structure our approach based on List (2020)'s

[30] Given the non-symmetric distribution inherent with a small number of clusters, the concept of standard error is not strictly applicable when implementing the wild bootstrap technique. Consequently, we calculate p-values and confidence intervals, adhering to the methodology recommended by Roodman et al. (2019).

SANS (Selection-Attrition-Naturalness-Scaling) criteria. Firstly, in terms of selection, our sample encompasses the universe of 720 High Court judges in Pakistan, serving as the primary, and often sole, avenue for citizens seeking redress for governmental overreach and abuse of authority. Secondly, regarding the naturalness of the setting, time frame, and choice task, we rely on high-stakes and natural measures in the context of the judiciary. Specifically, these judges engage in their 'in the wild' (natural) tasks in the field as we make inferences from observational data on judge hiring decisions (e.g., same law firm hires) and judicial decisions. Lastly, when scaling our paper's results to other contexts, it is essential to note that these judges, including their retirement at age 62, share similarities with counterparts, e.g., in India (where High Court judges also retire at 62). However, we categorize these findings as "wave 1 insights," employing List's (2020) terminology, and advocate for replications to confirm the potential application of effect sizes to other judiciaries in different settings. It is essential to note that while these results hold for Pakistan, a country with a population of 241 million (5th most populous), external validity to other contexts should be verified through further research.

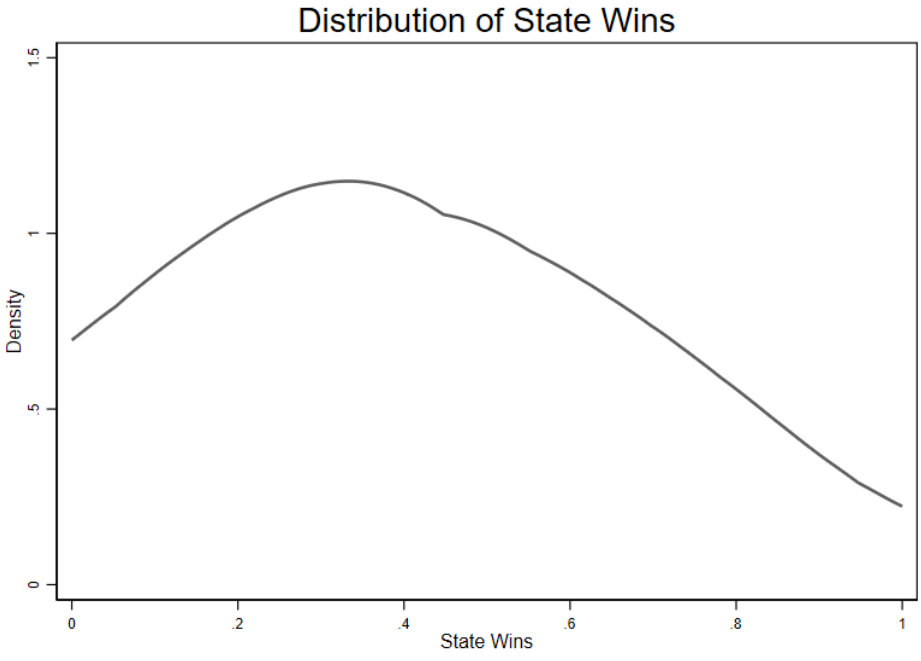
B. Additional Figures and Tables

Figure B1: Total Cases Versus Sampled Cases



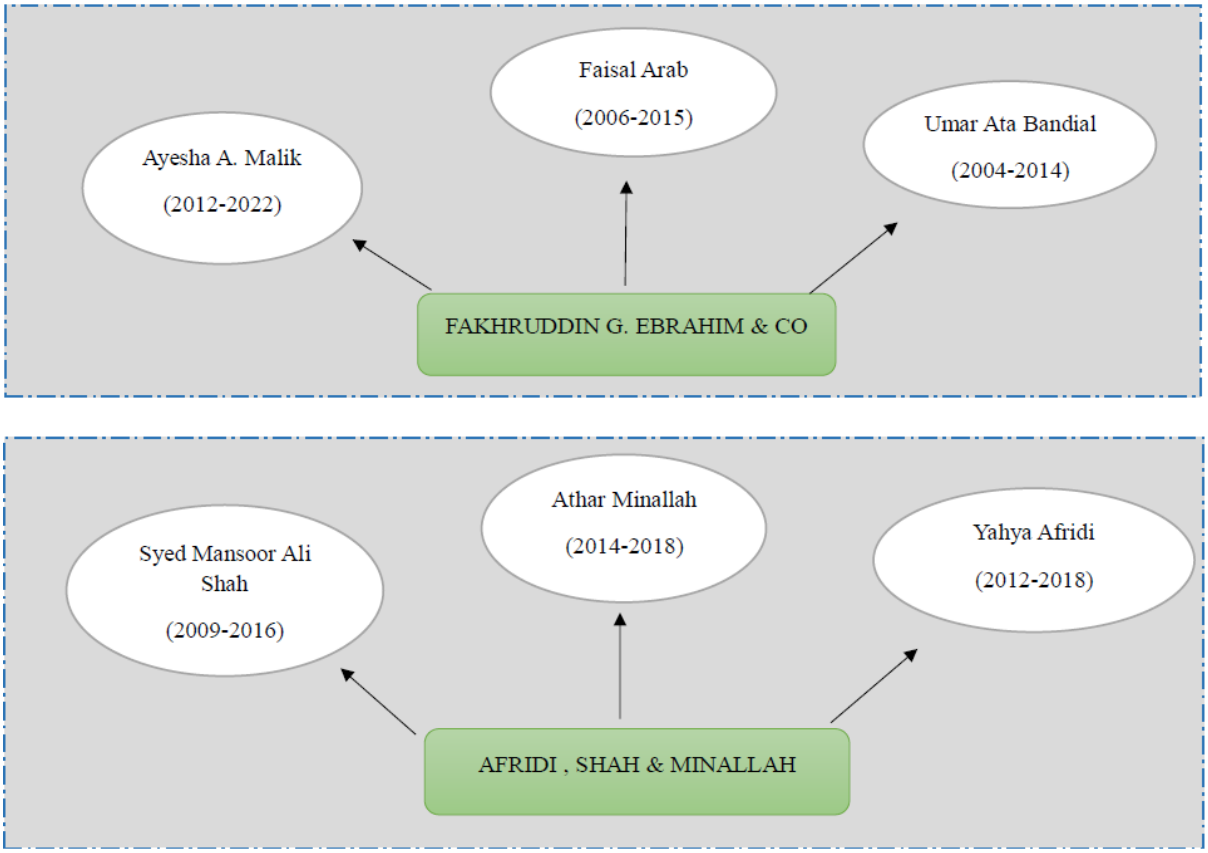
Note: The figure shows 20,000 cases randomly sampled between 1986 and 2022 from the High Courts of Pakistan. The sampled cases represent about 0.3% of the cases decided during this period. To be included in the sample, the case must involve the state as one of the parties, with the State typically acting as a respondent to a citizen's claim. It is worth noting that the State petitions in fewer than 3% of the cases, so the results correspond to citizens petitioning against the government.

Figure B2: Distribution of State Wins



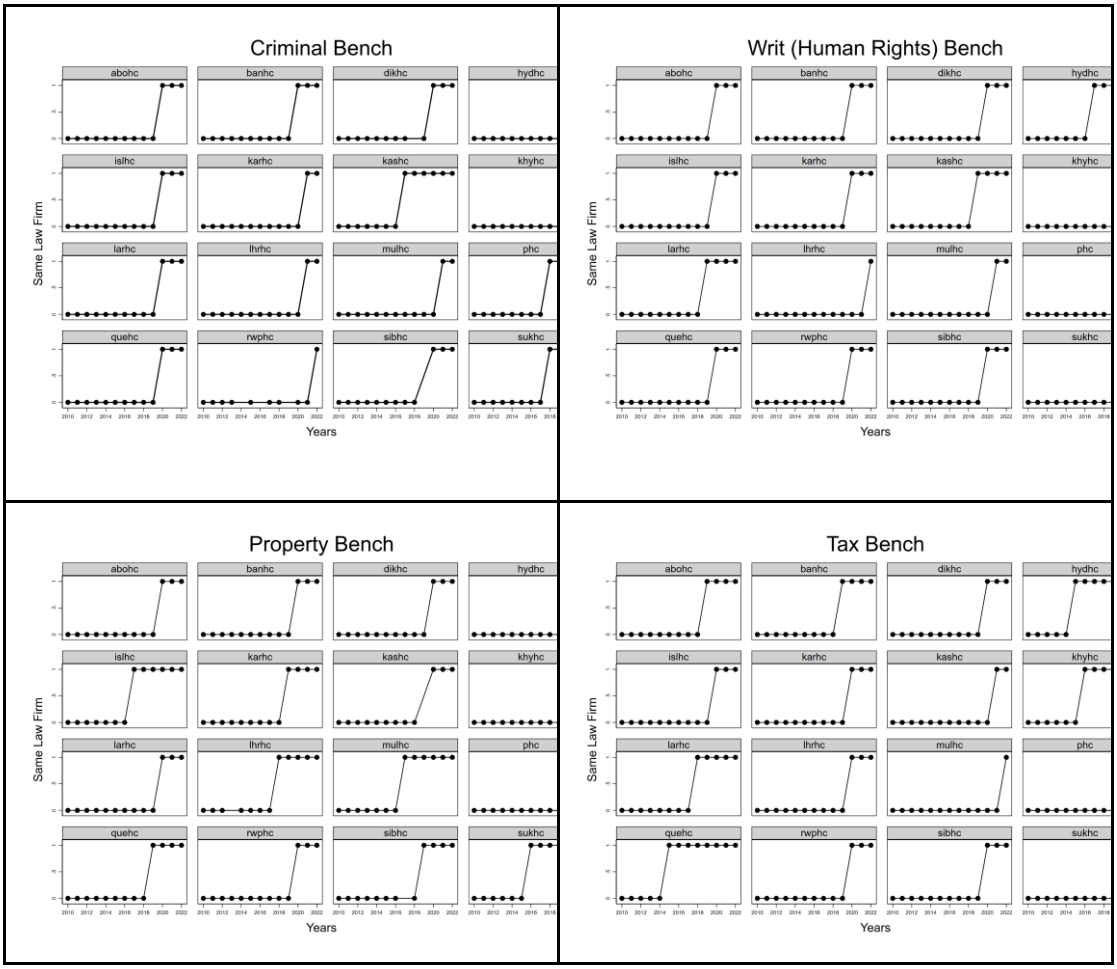
Note: The figure illustrates the estimated probability density of the state victory, which is our principal dependent variable. The distribution's mean is reported at 0.43, with a standard deviation of 0.50.

Figure B3: Illustration of Same Law Firm Judges



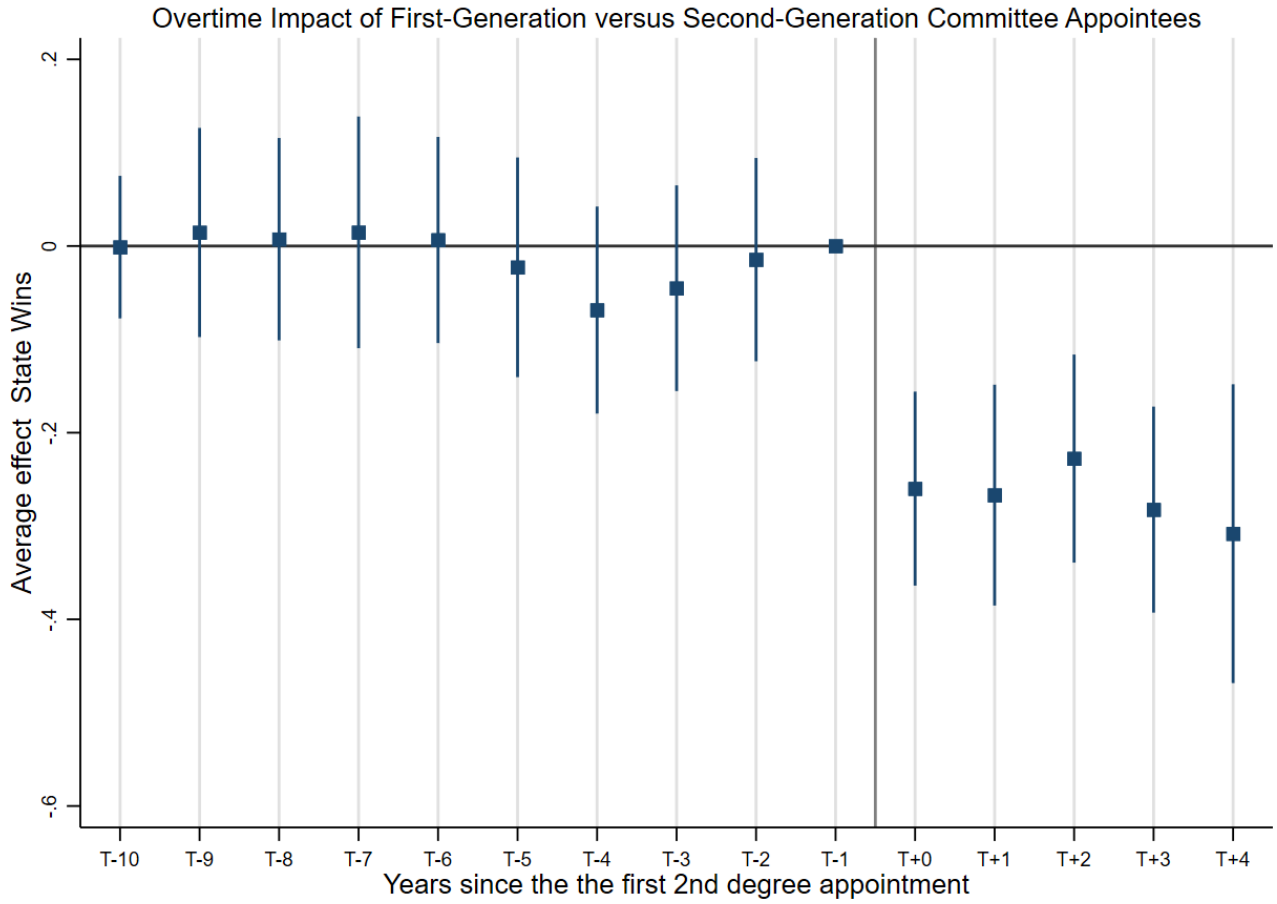
Note: The figure gives examples of judges in Pakistani courts from the same law firms. For instance, Justice Ayesha Malik and Justice Umar Ata Bandial were both associated with Fakhruddin G. Ebrahim & Co. Law firm before their judgeships. Justice Ayesha A. Malik was appointed after the committee selection reform in 2010, while Justice Umar Ata Bandial was in the judge hiring panel making judge selection decisions. The authors emphasize that while the example in this figure is based on factual information, it is used for data visualization purposes and to illustrate a general point. No judgments are made against individual judges in this research.

Figure B4: Same Law Firm Variable



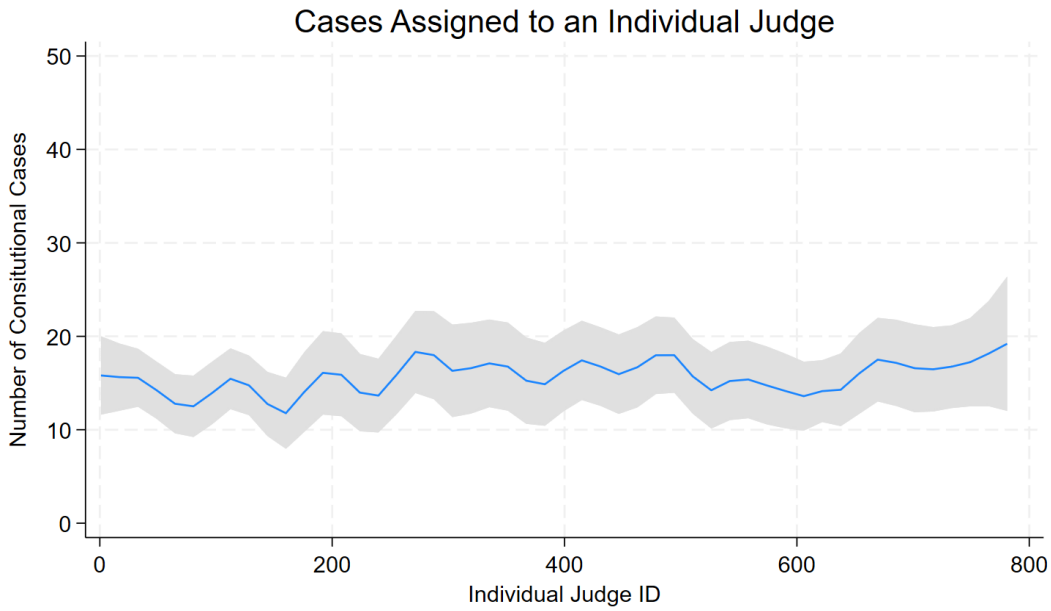
Note: The figure tracks the dynamics of the "Same Law Firm" indicator over time. This binary variable reflects the presence of judges on the bench who have a prior employment history with the same law firms as any member of the Judicial Commission of Pakistan, which is responsible for the appointment of judges post-2010. Initially set to zero, the variable transitions to one upon the appointment of a judge who shares a law firm background with at least one member of the Commission.

Figure B5: Comparative Impact of 1st and 2nd Generation Judicial Committee Appointees on State Wins - All Pretreatment years



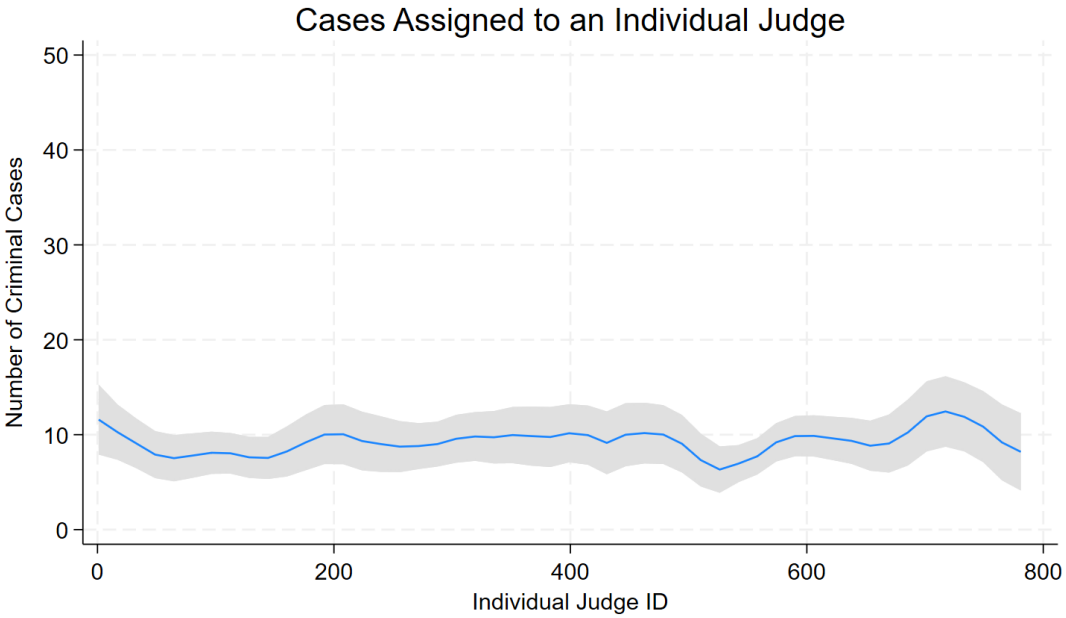
Notes: The figure presents an event-study plot depicting the year-by-year OLS estimation via Equation (2) for all estimable pre-periods (10 years). Specifically, it illustrates OLS estimates for the dynamic version of the TWFE model. The variable of interest is State Wins, a dummy variable indicating whether a case is ruled in favor of the state (1) or not (0). The treatment variable activates with the arrival of second-generation committee appointees at the judicial bench when the judicial hiring panel obtains a majority of committee-appointed judges. The comparison is made with decisions of the 1st generation committee-appointed judges, since rulings by executive-appointed judges are excluded. The omitted category represents the year before the first 2nd-Generation committee-appointed judge arrived. The bars depict 95 percent confidence intervals with standard errors clustered at the court-by-bench level.

Figure B6: Number of Constitutional Cases by Judge



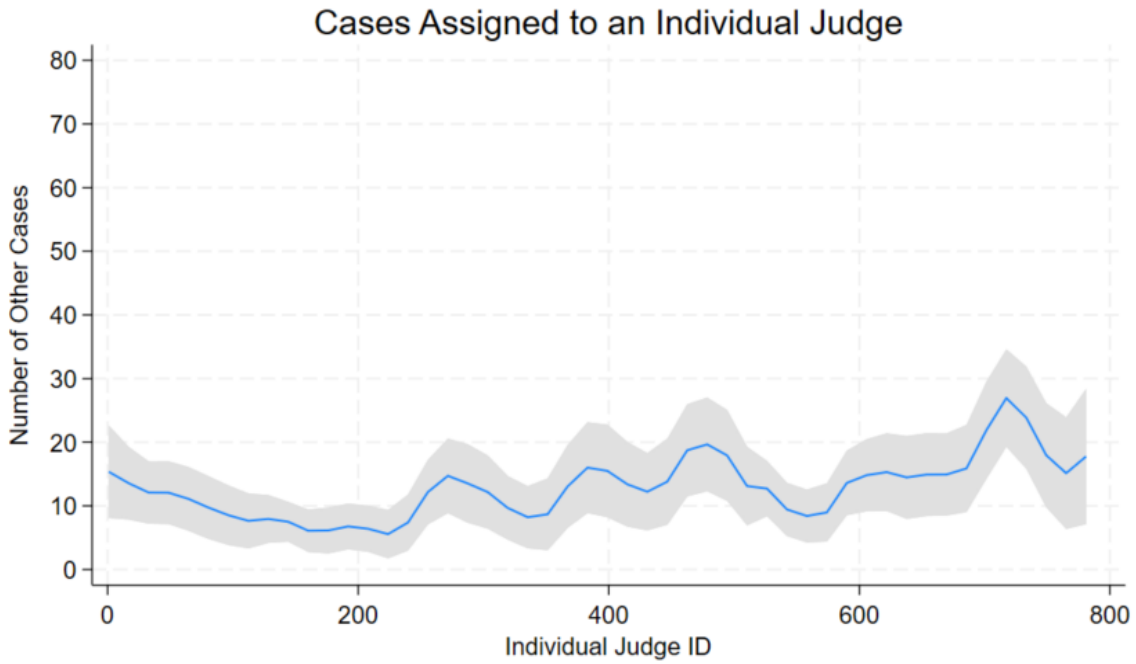
Note: The figure presents a breakdown of caseload per individual judge, sorted by the number of cases each judge handles. Specifically, the figure illustrates the allocation of constitutional cases to each judge within our sample of 720 judges. The judges' names are anonymized and represented by unique numbers along the X-axis, while the Y-axis indicates the number of constitutional cases handled by each judge. 95% confidence intervals are also reported.

Figure B7: Number of Criminal Cases by Judge



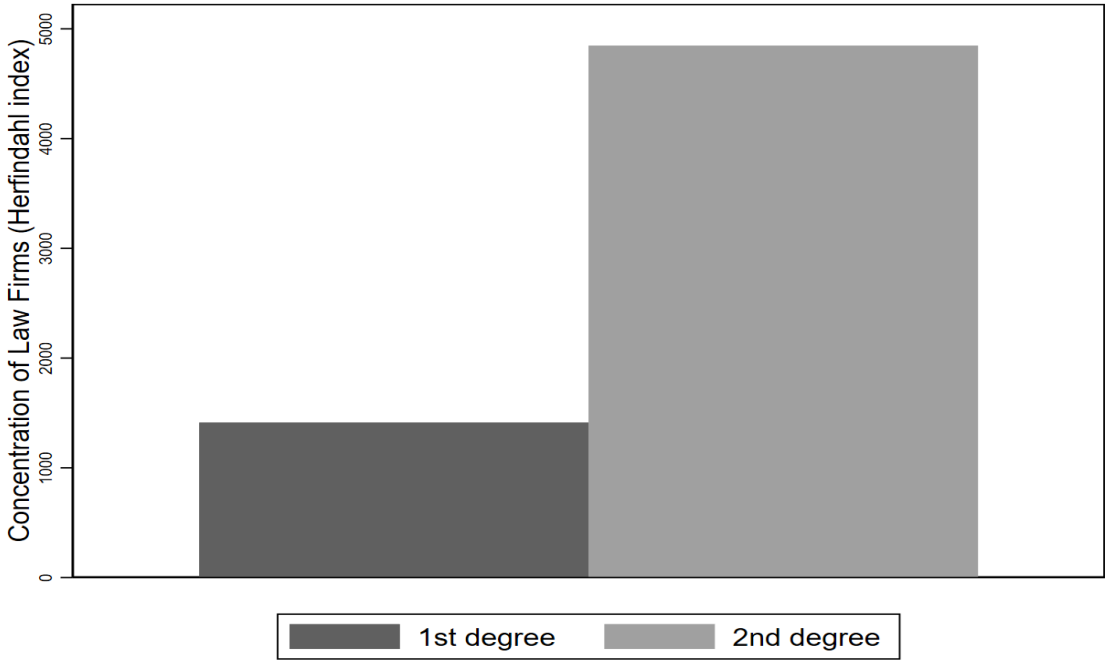
Note: The figure illustrates the allocation of criminal cases to each judge within our sample of 720 judges.The judge names are anonymized and identified as a unique number along the X axis, with the Y axis indicating the number of cases each judge handled. 95% confidence intervals are also reported.

Figure B8: Number of Other Case Types by Judge



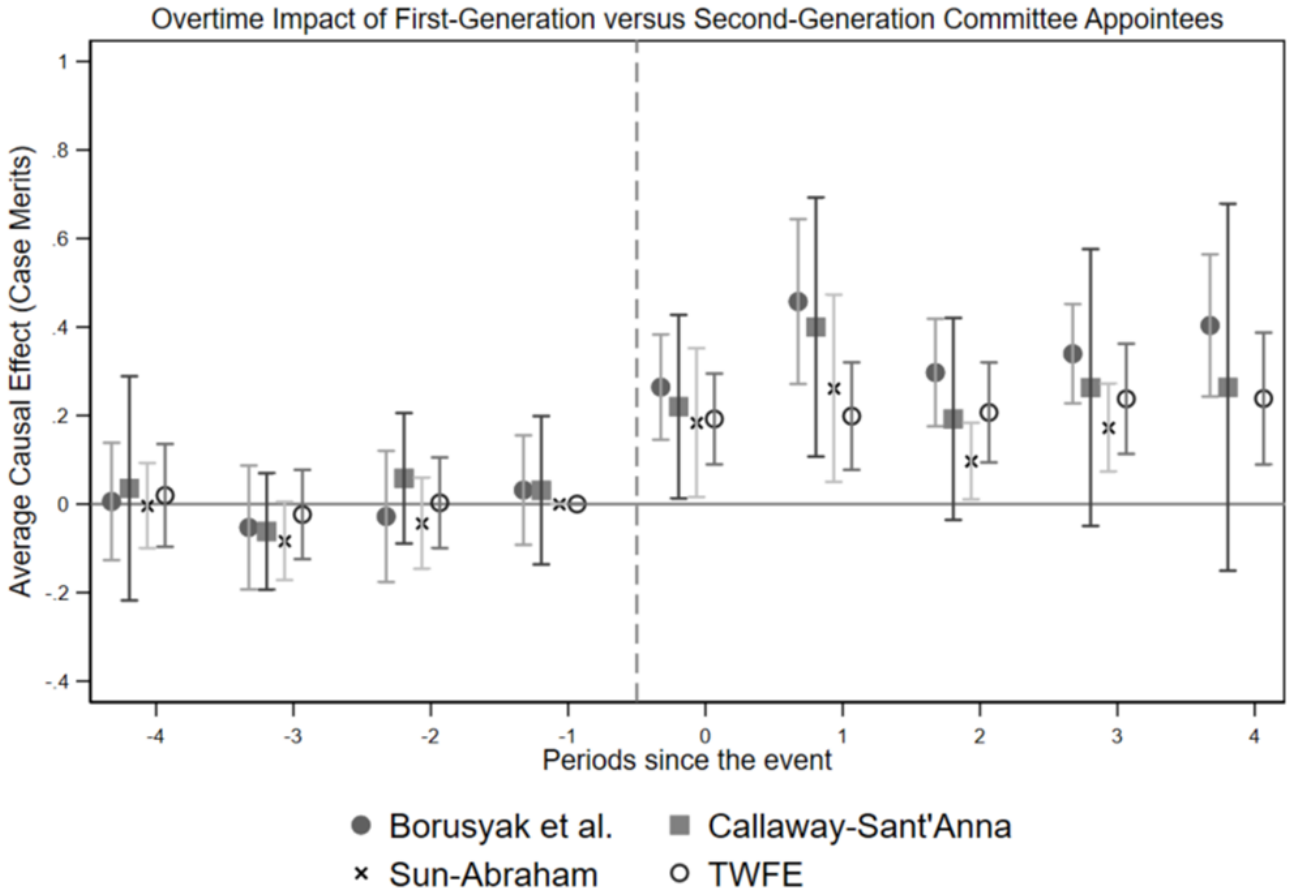
Note: The figure presents a breakdown of caseload per individual judge, sorted by the number of cases each judge handles. Specifically, the figure illustrates the allocation of other type of cases (non-constitutional and non-criminal) to each judge within our sample of 720 judges. The judges' names are anonymized and represented by unique numbers along the X-axis, while the Y-axis indicates the number of constitutional cases handled by each judge. 95% confidence intervals are also reported.

Figure B9: Herfindahl Index of Same Law Firms' Concentration



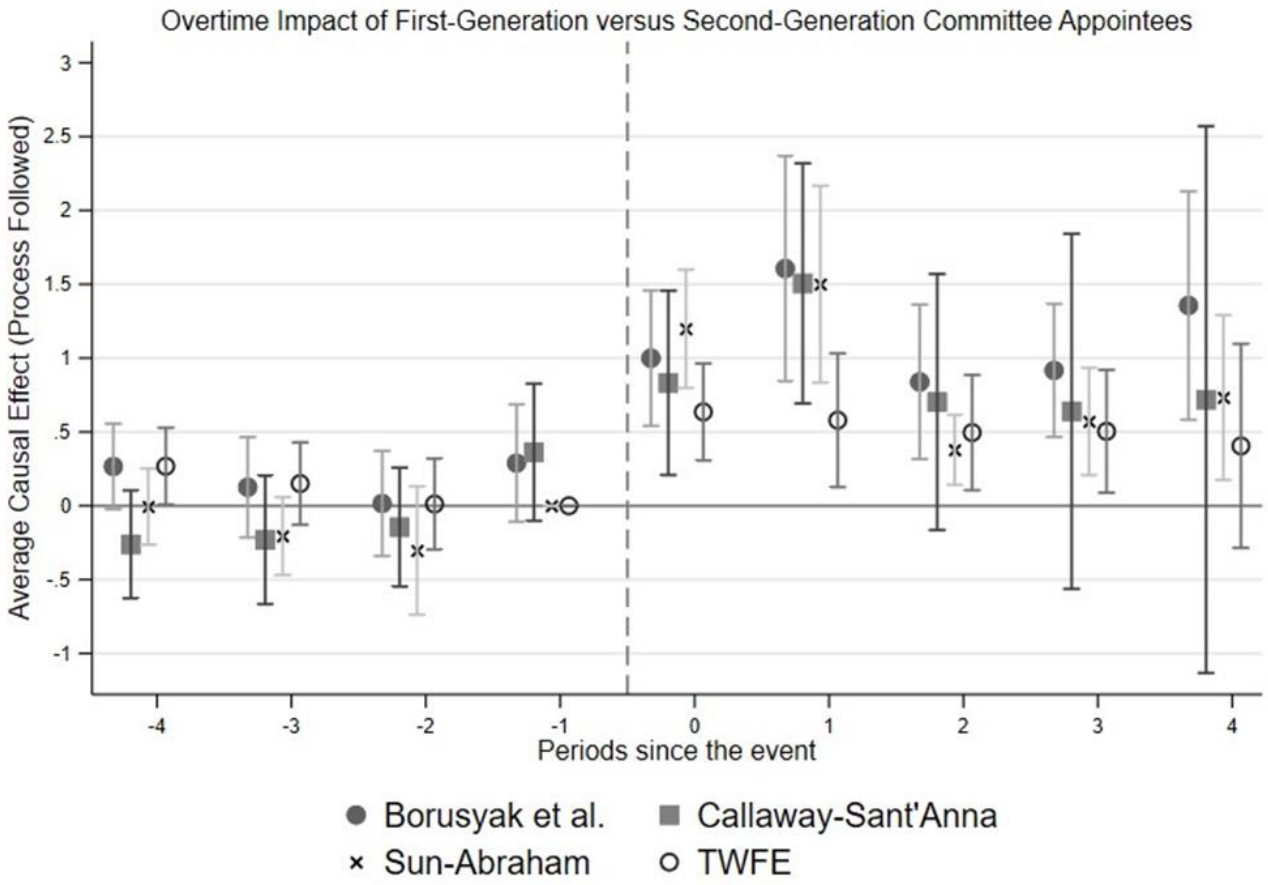
Note: The bars depict the "Herfindahl–Hirschman Index," measuring the market concentration of judges from the same law firm as at least one of the members of the judge hiring panel. The index considers the proportion of judges from the same law firms where they were previously attorneys, reflecting lower values for greater diversity and reaching a maximum value of 10,000. An index below 1000 points is considered an unsaturated market; between 1,500 and 2,500 points are considered moderately concentrated, and markets that exceed 2,500 points are considered highly concentrated. Retrieved from U.S. Department of Justice & FTC Guidelines § 5.3 (2010).

Figure B10: Impact of 1st Degree vs 2nd Degree Judges on Case Merits Over Time



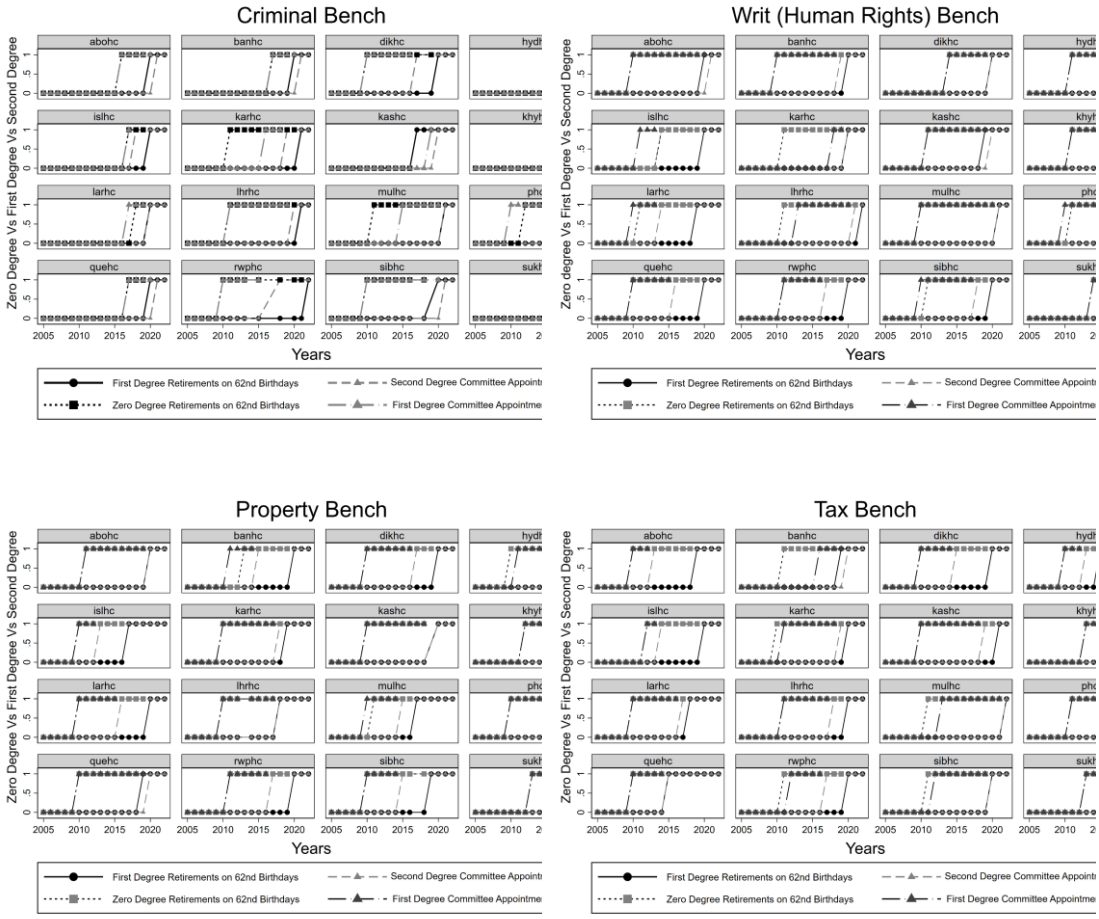
Notes: This figure presents event-study plots using four different estimators: a dynamic TWFE model with OLS (empty circle markers), Sun and Abraham (2021) (cross markers), Callaway and Sant'Anna (2021) (square markers), and Borusyak, Jaravel, and Spiess (2021) (filled circle markers). The dependent variable is Case Merits, a dummy that switches on for the case ruled on the legal merits of the case and zero otherwise. The treatment variable activates with the arrival of the first second-generation committee-appointed judge, i.e., when the judicial hiring panel obtains a majority of committee-appointed judges at the judicial bench. The comparison is made with decisions of the 1st generation committee-appointed judges, excluding rulings by executive-appointed judges. The omitted category represents the year before the arrival of the first 2nd-Generation committee-appointed judge at the judicial bench. The figure displays four post periods because estimating additional post periods would require using already treated units as controls for newly treated units, which could introduce bias due to heterogeneous treatment effects. Hence, newly introduced robust estimators limit the maximum number of robustly estimable post-periods to four. We also use four preperiods, as the Borusyak, Jaravel, and Spiess (2021) estimator experiences significantly increased standard errors after four preperiods. The bars depict 95 percent confidence intervals with standard errors clustered at the court-by-bench level.

Figure B11: Impact of 1st Degree vs 2nd Degree Judges on Process Followed



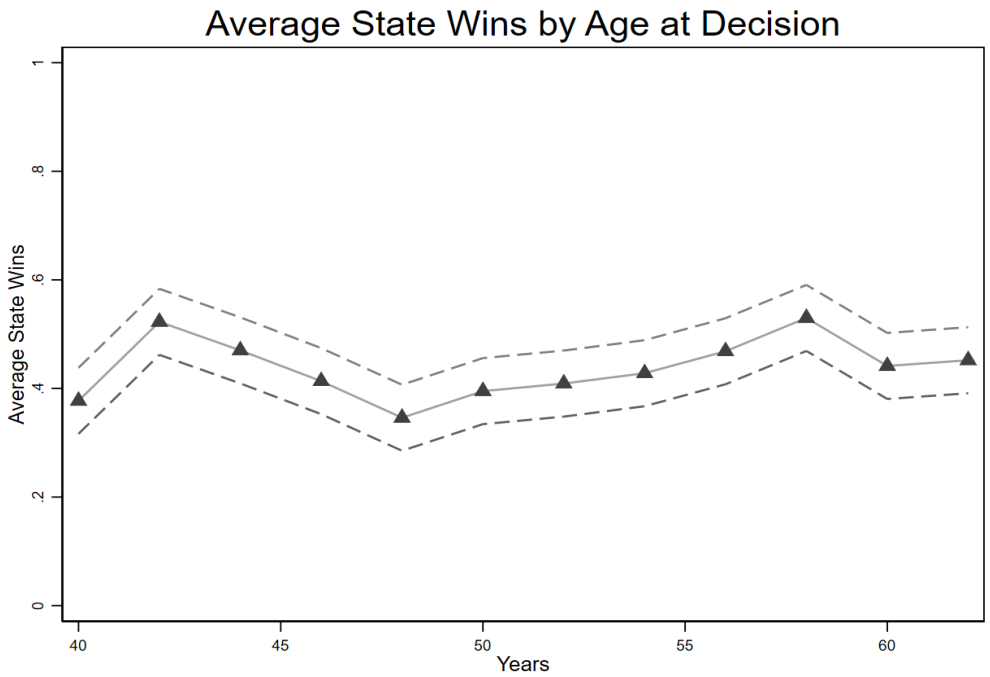
Notes: This figure presents event-study plots using four different estimators: a dynamic TWFE model with OLS (empty circle markers), Sun and Abraham (2021) (cross markers), Callaway and Sant'Anna (2021) (square markers), and Borusyak, Jaravel, and Spiess (2021) (filled circle markers). The outcome variable is Process Followed, a 1-5 rating on due process followed based on a panel of attorneys' evaluation of the judgment text. The treatment variable activates with the arrival of the first second-generation committee-appointed judge, i.e., when the judicial hiring panel obtains a majority of committee-appointed judges at the judicial bench. The comparison is made with decisions of the 1st generation committee-appointed judges, excluding rulings by executive-appointed judges. The omitted category represents the year before the arrival of the first 2nd-Generation committee-appointed judge at the judicial bench. The figure displays four post periods because estimating additional post periods would require using already treated units as controls for newly treated units, which could introduce bias due to heterogeneous treatment effects. Hence, newly introduced robust estimators limit the maximum number of robustly estimable post-periods to four. We also use four preperiods, as the Borusyak, Jaravel, and Spiess (2021) estimator experiences significantly increased standard errors after four preperiods. The bars depict 95 percent confidence intervals with standard errors clustered at the court-by-bench level.

Figure B12: Evolution of Appointments and Mandatory Retirements



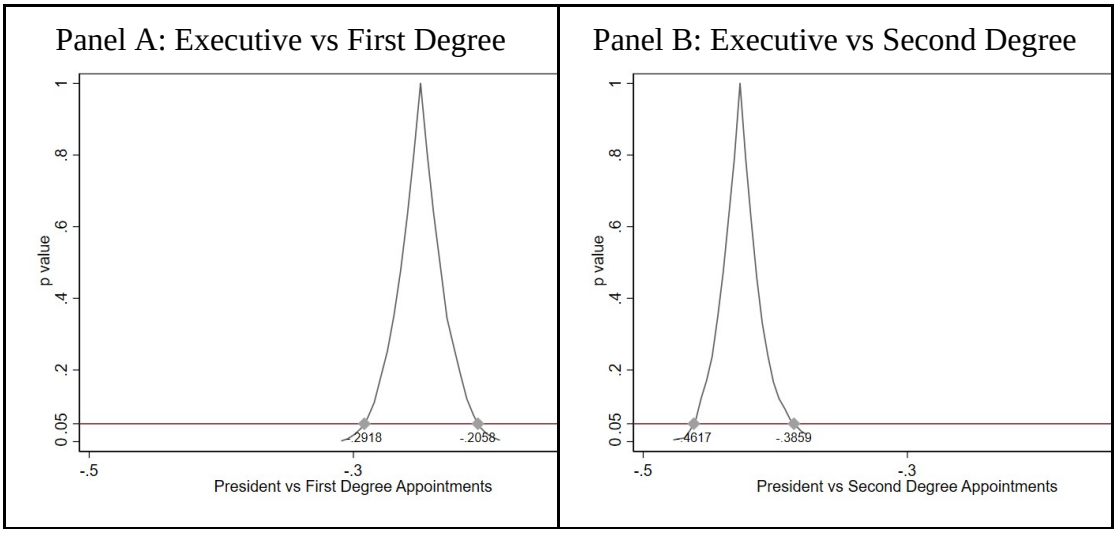
Note: The figure above illustrates the arrival of the first and second generations of committee-appointed judges, and mandatory retirements that often preceded these appointments. It showcases the evolution of appointments and retirements before and after the 2010 judge selection reform across the 16 divisional High Courts or 64 judicial benches.

Figure B13: Age Decision and State Wins



Note: The figure presents a plot showcasing the decisions made by judges at different age brackets throughout their tenure and their corresponding pro-government rulings. The rules for selection and retirement mandate that high court judges must be at least 40 years old upon appointment and must retire before their 62nd birthday. 95% Confidence Intervals are also reported.

Figure B14: Confidence Interval by Wild Bootstrap



Note: The figure displays estimated distributions around our coefficients of interest using wild bootstrap clustering, as per Cameron et al. (2008), which imposes a small cluster correction. Roodman et al. (2019)'s *bootest* in Stata 17 is used to construct this confidence interval with 999 repetitions. Cameron et al.'s (2008) wild bootstrapping procedure utilizes the Rademacher distribution with the null hypothesis imposed. It involves converting the t-statistic from the Wald test and inverting the test to compute the bounds of the confidence set under the imposed number of clusters (in our case, 64 judicial benches). The resulting "confidence curve" is plotted, providing the bootstrapped p-values that represent the limits of the confidence set. The estimated distributions for first- and second-generation judges do not appear to intersect, suggesting a reform multiplier effect, even without strict reliance on asymptotic theory.

Table B1: Descriptive Statistics

Variables	Observations	Mean	Std. Dev.	Min	Max
Panel A: Case Characteristics (varies by case)					
State Wins	18,578	0.443	0.497	0	1
Merit	18,578	0.679	0.467	0	1
Process Followed	18,578	3.400	1.523	1	5
Executive vs. 1st degree	18,578	0.282	0.450	0	1
Executive vs. 2nd degree	18,578	0.120	0.326	0	1
Constitutional	18,578	0.622	0.485	0	1
Criminal	18,578	0.368	0.482	0	1
Land case	18,578	0.368	0.482	0	1
Pages of judgment order	18,578	8.776	7.635	1	81
Chief Justice on Bench	18,578	0.063	0.244	0	1
Number of lawyers	18,578	4.156	1.845	2	32
Number of judges	18,578	1.805	0.833	1	5
Panel B: Appointment and Retirement in the transition year (by court-bench)					
Executive Vs. 1st degree	64	0.341	0.174	0	0.656
Executive Vs. 2nd degree	64	0.196	0.127	0.017	0.598
1st degree Vs. 2nd degree	64	0.330	0.202	0.055	1
Panel C: Judge Characteristics (by judges)					
Tenure at Decision	720	3.354	2.841	0	22
Gender	720	0.961	0.193	0	1
Former LowerCourt Judge	720	0.110	0.313	0	1
Office Holder Bar Asso.	720	0.549	0.498	0	1
Former Lawyer	720	0.893	0.309	0	1
Majority ethnicity	720	0.183	0.387	0	1
Minority ethnicity	720	0.061	0.240	0	1
Promoted to Supreme Court	720	0.289	0.454	0	1
Same Law Firm	720	0.489	0.500	0	1
Executive Appointed judges	720	0.454	0.498	0	1
1st degree judges	720	0.197	0.398	0	1
2nd degree judges	720	0.349	0.477	0	1

Note: This table presents summary statistics for the baseline sample, consisting of 18,578 cases involving 720 judges across 64 judicial benches within the 16 divisional benches of the High Courts of Pakistan. The period covered spans from 1986 to 2022.

Table B2: Impact of Transition vs Non-Transition Mandatory Retirements on State Wins

	State Wins			
	(1)	(2)	(3)	(4)
	Transition Year = 0	Transition Year = -1	Transition Year = -2	Transition Year = 0 to -2
Retirement of Judges (1st Degree)	-0.0687**			-0.0724**
	[0.0312]			[0.0342]
Retirement of Judges (2nd Degree)	-0.135***			-0.133***
	[0.0403]			[0.0397]
Retirement of Judges 1 year Before 1st degree		-0.0396		-0.0247
		[0.0673]		[0.0599]
Retirement of Judges 1 year Before 2nd degree		0.0329		0.0509
		[0.0536]		[0.0497]
Retirement of Judges 2 years Before 1st degree			-0.143	-0.126
			[0.128]	[0.147]
Retirement of Judges 2 years Before 2nd degree			-0.0433	-0.0157
			[0.0574]	[0.0615]
Court-by-Bench and Year FE	Yes	Yes	Yes	Yes
Controls and Court-by-Bench Trends	Yes	Yes	Yes	Yes
Mean of dependent variable	0.443	0.443	0.443	0.443
R-squared	0.163	0.161	0.161	0.163
Observations	18,578	18,578	18,578	18,578

Note: The table displays the estimated reform multiplier effect on the likelihood of government victories by analyzing the impact of mandatory retirement in reform transition versus non-transition years. The dependent variable, "State Wins," is a binary variable that is activated when the government wins a case (1) and remains off when it loses (0). We introduce several dummy variables to capture different scenarios: "Retirement of Judges 1st degree" is a binary variable that turns on when judges reach the mandatory retirement age and are replaced by the 1st generation of committee appointees in the same year. The hiring panel for these judges consisted of a majority of Executive appointees. "Retirement of Judges 2nd degree" is another binary variable that activates when judges are replaced due to mandatory retirements of the 2nd generation committee-appointed judges and their hiring panel had a majority of of committee appointees. "Retirement of Judges 1 year before 1st degree" becomes active when judges exit and are replaced because of mandatory retirements, but in non-transitional years when there is no generational shift or reform mutation (one year before the selection reform). Similarly, "Retirement of Judges 1 year before 2nd degree" activates when judges reach mandatory retirement age and are replaced *one year* before the arrival of the first 2nd-degree judge, whose hiring panel had a majority of committee appointees. In this scenario, the hiring panel of judges still had a minority of committee-appointed judges. And so on. Control variables are as in baseline and include case type fixed effects, the number of pages in the judgment order, the presence of the chief justice on the bench, the number of lawyers involved in the case, the number of judges on the case, and the total number of judges in a district bench. Standard errors are reported in brackets and are clustered at the court-by-bench level. ***p<0.01, ** p<0.05, * p<0.1.

Table B3: Balance over Case and Bench Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Constitutional	Criminal	Land Case	Pages Judgment	Bench Chief Justice	No. of Lawyer	No. of Judge	Total Judges Bench	Avg Case on s Bench	Executive Vs 1st Deg	Executive Vs 2nd Deg
Executive Vs 1st degree	-0.00794	0.00185	0.0202	-0.0851	-0.00145	0.176*	0.0604	0.00195	0.0617		
	[0.0129]	[0.00259]	[0.0226]	[0.457]	[0.0170]	[0.0976]	[0.0524]	[0.212]	[0.283]		
Executive Vs 2nd degree	-0.00818	-0.00271	0.0584*	-0.261	0.000690	0.170	0.152	0.218	0.346		
	[0.0234]	[0.00409]	[0.0334]	[0.694]	[0.0261]	[0.158]	[0.0958]	[0.214]	[0.484]		
Constitutional										-0.00339	-0.00802
										[0.0423]	[0.0345]
Criminal										-0.00175	-0.00682
										[0.0421]	[0.0339]
Land Case										-0.00205	0.00490
										[0.00594]	[0.00445]
Pages Judgment										0.0000175	-0.0000822
										[0.000364]	[0.000270]
Bench Chief Justice										-0.00185	0.00158
										[0.00743]	[0.00575]
No of Lawyer										0.00113	0.000005
										[0.00103]	[0.000826]
No of judge										-0.00110	0.00310
										[0.00329]	[0.00273]
Total Judges Bench										-0.0206	0.0209
										[0.0309]	[0.0131]
Avg cases Bench										-0.00201	0.00242
										[0.00418]	[0.00344]
District by Bench & Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bench Trends & Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	0.622	0.368	0.368	8.776	0.0634	4.156	1.805	7.925	4.111	0.282	0.120
R-squared	0.746	0.772	0.300	0.225	0.038	0.019	0.076	0.981	0.612	0.789	0.800
Observations	18,578	18,578	18,578	18,578	18,578	18,578	18,578	18,578	18,578	18,578	18,578
F Statistic	0.191	1.321	1.531	0.0709	0.0184	1.633	1.291	0.830	0.264	0.503	0.503
p-values	0.827	0.274	0.224	0.932	0.982	0.203	0.282	0.441	0.769	0.867	0.619

Note: The table presents an analysis testing whether the reform's multiplier effect also resulted in changes to the composition of cases. Specifically, the table provides coefficient estimates of equation (1), treating case characteristics as dependent variables (excluding one control variable that becomes the dependent variable). The dependent variables include the number of constitutional cases filed (Column 1), the number of criminal cases filed (Column 2), the number of defense-filed land cases (Column 3), the number of pages in the judgment text (Column 4), a dummy variable indicating the presence of the chief justice among the adjudicating judges (Column 5), the total number of lawyers (defense and prosecution combined) (Column 6), the number of judges adjudicating the case (Column 7), the total number of judges on the bench (Column 8), and the average number of cases on a bench (Column 9). Column 10 presents the joint effect of all case characteristics on Executive Vs. 1st Degree (1st generation judges), while Column 11 presents the joint effect on Executive vs. 2nd Degree (2nd generation judges) as suggested in Bruhn and McKenzie (2009). Standard errors are reported in brackets and are clustered at the court-by-bench level. ***p<0.01, ** p<0.05, * p<0.1.

Table B4: Effect of Judicial Committee Composition on Hiring from Same Law Firms

	Same Law Firm Hire			
	(1)	(2)	(3)	(4)
Committee Appointees Minority 2-5	0.0292			0.0364
	[0.102]			[0.116]
Committee Appointees Minority 3-4		0.00393		0.0124
		[0.0723]		[0.0861]
Committee Appointees **Majority 4-3**			0.0909***	0.0911***
			[0.0284]	[0.0284]
Observations	8,230	8,230	8,230	8,230
R-squared	0.436	0.436	0.438	0.438
District bench and Year FE	Yes	Yes	Yes	Yes
Controls and Bench-Specific Trends	Yes	Yes	Yes	Yes
Mean Dependent Variable	0.501	0.501	0.501	0.501

Note: This table displays the estimated coefficients for the effect of nearing and securing a majority of committee-appointed judges within the judicial hiring panel. The dependent variable, "Same Law Firm Hire," is a binary indicator that assumes a value of one when a judge, who was recruited from the same law firm as one or more members of the appointing panel, adjudicates a case. To ascertain the reform multiplier effect, the comparison specifically homes in on committee-appointed judges, excluding cases adjudicated by judges appointed by the executive. The independent variables turn on when committee appointees join the judge hiring panel. Comparisons are made with decisions involving committee-appointed judges, where the judge selection panel is predominantly composed of executive- versus committee-appointed judges. In the first two columns, the judge hiring panel still has a minority of committee appointees. In contrast, in Columns 3 and 4, the hiring panel just attains a majority of committee-appointed judges. Controls include case type fixed effects (e.g., rights dispute or property dispute), number of pages of judgment order, presence of chief justice on the bench, number of lawyers on the case, number of judges on the case, and total judges on the bench. The court-by-bench or district-by-bench fixed effects, year-fixed effects, and bench-specific linear trends are always included. Standard errors are reported in brackets and are clustered at the court-by-bench level. ***p<0.01, ** p<0.05, * p<0.1.

Table B5: Impact of Minority versus Majority of Committee Appointees on Decision Quality

	Case Merits		*Process Followed*	
	(1)	(2)	(3)	(4)
Committee Appointees **Majority 4-3**	0.227***	0.203***	0.623***	0.575***
	[0.0510]	[0.0450]	[0.210]	[0.170]
Committee Appointees Minority 3-4	-0.00952	0.0131	-0.135	-0.0710
	[0.0587]	[0.0433]	[0.204]	[0.196]
Committee Appointees Minority 2-5	0.0476	0.0633	0.285	0.326
	[0.0485]	[0.0904]	[0.244]	[0.281]
Observations	7,470	7,470	7,470	7,470
R-squared	0.140	0.176	0.096	0.124
District by Bench and Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Specialization bench linear trends	No	Yes	No	Yes
Mean	0.773	0.773	3.684	3.684

Note: This table displays the estimated coefficients for the effect of nearing and securing a majority of committee-appointed judges within the judicial hiring panel. In Columns (1) and (2), we present estimates using the Case Merit variable as the dependent variable, which is a dummy variable taking the value of one for cases ruled on merit and zero otherwise. Columns (3) and (4) provide estimates with the dependent variable 'process followed,' measured by a rating from one to five, reflecting the degree of due process observed in the case according to assessments by attorneys. Executive Appointments Vs. 1st Degree is a variable that activates with the arrival of the first-generation appointed judge at the bench (one degree of separation from the Executive Appointments). The Executive Appointments Vs. 2nd Degree is a variable representing the arrival of the second-generation committee-appointed judge at the bench (two degrees of separation from the Executive Appointments). It switches on when the first second-degree committee-appointed judge arrives at the bench and remains one. The comparison of judicial decisions by committee appointees is made relative to cases decided by the judges appointed by the executive in both cases. Standard errors are reported in brackets and are clustered at the court-by-bench level. ***p<0.01, ** p<0.05, * p<0.1. Tables not included but referenced in the manuscript, notably those associated with the discussion of robustness and sensitivity analyses in Online Appendix A5, are accessible HERE.