Technology and the Production of Justice: System-Wide Effects of Indonesia's E-Court Reform

Sultan Mehmood ⓡ*
New Economic School

Resuf Ahmed ⓡ
Stanford U.

Alexandra Filimonova ⓡ
Bocconi University

August 23, 2026

Abstract

Delays in judicial proceedings are a pervasive institutional challenge. We study how technology changes judicial production using Indonesia's e-court reform, which introduced e-filing, e-payments, and virtual hearings for civil but not criminal cases. Using case-level data linking trial decisions to appeals, we compare civil and criminal cases handled by the same judge in the same court before and after reform. E-courts reduced case duration by about 7 days, or 15 percent, while preserving judicial quality. If anything, appeals and reversals fell, judgments became more concise, and readability and statutory citations were unchanged. The gains extended through the judicial system: case duration in untreated appellate courts fell by about 9 days, or 20 percent. The pattern is consistent with the reform freeing judicial capacity for more complex disputes, highlighting the potential for technology to expand capacity by reorganizing judicial production and relieving system-wide congestion.

JEL Codes: K40, H11, O33

Keywords: technology, organizations, development.

*Author order is randomized. Randomization was conducted using the AEA Author Randomization Tool (AEA Randomization ID: KtE1CtFRZtyt). We thank Margarita Kudaeva and Mariya Salikhova for outstanding research assistance. We thank Thomas Fujiwara, Imran Rasul and Mushfiq Mobarak for their helpful comments and suggestions. We also thank seminar participants, conferences, and workshops in Tokyo and Chicago. Contact: Sultan Mehmood, New Economic School (NES), smehmood@nes.ru; Resuf Ahmed, Stanford University, resuf@stanford.edu; Alexandra Filimonova, Bocconi University, alek2000sandra@gmail.com.

"Justice too long delayed is justice denied."

Martin Luther King Jr.

1 Introduction

Strong state capacity is central to economic development (Besley et al., 2022). A central challenge is how states can expand their effective capacity when increases in personnel, infrastructure, and other conventional inputs are relatively costly and slow to scale. One possibility is to use existing organizational resources more effectively. Technology can facilitate such changes by reducing the costs of performing existing activities and changing how work is organized. By substituting for routine activities and complementing more complex work, technology can change the allocation of worker effort (Autor et al., 2003). Recent evidence from the private sector shows that new technologies can raise productivity by changing how existing workers perform their jobs (Brynjolfsson et al., 2025). In many large organizations, especially in the public sector, however, production is organized across interdependent activities, so improvements at one stage may be constrained by capacity elsewhere in the organization. Technology can therefore do more than raise productivity in the activity it directly affects. By freeing time at one stage, it can shift capacity toward other activities and affect performance throughout the production process.

Courts provide an especially informative setting in which to study these organizational effects. Judges allocate limited time across many cases of varying complexity, while cases themselves move through linked stages of adjudication and appeal. A technology that reduces the time required for one class of cases can therefore affect not only those cases, but also the other disputes competing for the same judicial capacity and, potentially, subsequent stages of the judicial system. These organizational interactions matter because legal capacity, the state's ability to enforce rights and resolve disputes in a timely and reliable manner, is a key dimension of state capacity and economic development (Besley and Persson, 2009). Frictions at any stage of judicial production can constrain this capacity and propagate delays through the system. Limited information technology, inadequate support for judges, and weak court administration have been identified as important sources of such dysfunction (Papaioannou and Karatza, 2018). When cases remain unresolved, creditors cannot recover loans, property rights become less secure, and the economic value of formal legal rights declines (Djankov et al., 2003). Repeated across large caseloads and multiple levels of adjudication, these frictions can become a system-wide constraint on the state's ability to provide justice.

The resulting burden on judicial systems can be enormous. India alone has more than 51 million pending cases, and at current disposal rates, clearing the existing stock could take centuries (New York Times, 2024). Severe congestion is not confined to developing countries: U.S. immigration courts have accumulated more than two million pending cases (U.S. Government Accountability Office, 2024), while the Crown Court of England and Wales had more than 75,000 outstanding cases at the end of 2024 (Institute for Fiscal Studies, 2025). At this scale, judicial delay is not simply a matter of individual cases taking too long, but a constraint on the capacity of the judicial system as a whole.

These constraints have made judicial administration an important target for technological reform. Over the past two decades, governments across the world have introduced e-court systems that move filing, payments, summons, and hearings onto electronic platforms (World Bank, 2023a,b, 2024). By 2019, roughly one-third of countries worldwide had adopted at least one such feature. The effects of these reforms may extend well beyond the court processes they directly target. E-courts can reduce procedural burdens, change the demands placed on judicial attention across cases, and alter how cases move through different levels of the court system. Whether these changes improve performance, however, is ambiguous. Faster resolution at the point of reform may come at the cost of decision quality, while freed capacity may induce additional filings by litigants, limiting reductions in congestion, or shift congestion further down the judicial hierarchy. Understanding the effects of e-courts, then, requires tracing not only what happens within the courts directly targeted by reform, but also how those effects transmit, or fail to transmit, through the judicial hierarchy.

Indonesia provides an especially useful setting in which to do so. Beginning in 2018, the judiciary introduced an e-court system with e-filing, e-payments, and virtual hearings for civil cases, while leaving criminal cases outside the reform. We combine this differential exposure with case-level data constructed from millions of records in the Indonesian Supreme Court's online repository, covering civil and criminal cases across the judicial hierarchy from 2010 to 2020. We link trial decisions to subsequent appeals and observe the full text of judgments at both levels. The resulting data allow us to measure case duration, appeals, reversals, and judicial writing, and to trace whether changes in trial-court production propagate through the judicial hierarchy.

A central identification concern, however, is that reform placement may itself be endogenous. Comparing reformed and unreformed courts may conflate the effects of e-courts with the factors that determine where reform occurs. This issue is especially salient in the judiciary, where reform adoption responds directly to pre-existing differences in court performance and capacity (Chemin, 2021, p. 9). We address it by restricting the baseline analysis to reform-exposed courts and comparing treated civil cases with untreated criminal cases handled by the same judges before and after implementation. The specification includes judge, case-type, and court-by-month-year fixed effects. Identification therefore comes from the differential post-reform change in non-criminal relative to criminal cases within the same court and month, rather than from comparisons between courts selected into and out of reform. We find little evidence that this within-court comparison is accompanied by changes in the underlying mix of cases: detailed case-type shares remain broadly stable around adoption. As a complementary check on the restriction to reform-adopting courts, we re-estimate the main outcomes using a broader sample of nearly 800 courts and more than three million cases. With the corresponding fixed-effect structure, the estimates are similar to those in the baseline sample.

Our baseline within-court comparison yields a sizable reduction in case duration. Following adoption, case duration for reform-eligible non-criminal cases falls by about 7 days relative to criminal cases handled by the same judges in the same courts, equivalent to roughly 15 percent of mean case duration. Aggregated over the sample period, these reductions amount to more than half a million judicial days saved. To put this magnitude in perspective, achieving a comparable reduction in case-processing time through additional hiring would have required at least 320 judges, at an estimated salary cost of approximately USD 447.4 million over the same period (ERI Economic Research Institute, 2025). In comparison, government documents indicate that implementing and operating the e-court system cost approximately USD 71.6 million (Supreme Court of the Republic of Indonesia, 2020). These figures imply an equivalent personnel-cost benchmark of about USD 6.25 for each dollar spent on E-Courts, or USD 3.12 under the more conservative assumptions. These comparisons are necessarily approximate and serve only as rough benchmarks for the magnitude of the estimated reduction in delay.

These gains could be attenuated if shorter case durations led to additional disputes entering the courts and reintroducing congestion. We find little evidence of such a response: case filings remain broadly unchanged following adoption. Thus, despite the lower cost of obtaining a faster decision, litigants do not appear to substantially increase their use of the courts.[1]

Equally important, faster case resolution does not appear to come at the expense of decision quality. Reform-affected non-criminal cases are substantially less likely to be appealed, and the share of first-instance decisions that are ultimately reversed also falls, with both appeals and reversals declining by roughly one half relative to their sample means. Text

[1]This pattern is consistent with Acemoglu et al. (2020), who find that substantial reductions in case delay do not necessarily translate into higher filing rates, and emphasize the role of information and other barriers in shaping access to formal courts.

analysis of the available full-text judgments further suggests that opinions become more concise, without any corresponding deterioration in readability, statutory citations per page, or reasoning indicators. We also find no evidence of greater bias in appellate review, suggesting that improvements in speed are not offset by changes in fairness.

The reduction in trial-court appeals also appears to have relaxed capacity constraints higher in the judicial hierarchy. High-court decision time fell by about 9 days, or 20 percent relative to the mean, consistent with lower appellate inflows easing upstream congestion. This suggests that changes in trial-court resolutions can propagate through the judicial hierarchy, consistent with the reform's effects extending beyond the courts it directly targeted.

We further examine whether the upstream effect reflects changes in the composition of appellate cases. Using pre-reform appellate linkages, we measure for each higher court the share of its non-criminal caseload originating in trial courts that later adopt e-courts. If the reform eases congestion by reducing appeals, case duration should fall more in higher courts historically more exposed to these trial courts. This is exactly what we find: case duration falls more in these higher courts. Because exposure is defined using pre-reform case flows, the result is not mechanically driven by post-reform changes in the composition of appeals.

The effects also vary systematically with case complexity. Reductions in duration are concentrated among relatively routine matters, such as birth-certificate corrections, inheritance filings, and other documentation-change cases. Medium-complexity disputes, including child-custody cases, also become faster, alongside improvements in measured quality. In contrast, case duration changes little for more complex and contentious marital and divorce proceedings. This pattern is consistent with e-courts economizing on routine procedural tasks and reallocating scarce judicial capacity toward more demanding disputes, as suggested by organizational theories of task allocation and specialization (Holmstrom and Milgrom, 1991; Garicano, 2000).

Our design restricts attention to courts that adopt e-courts and compares treated civil cases with untreated criminal cases handled by the same judges before and after implementation. The causal interpretation therefore requires that, absent the reform, civil and criminal cases within these courts would have evolved similarly, and that implementation did not itself alter case assignment in ways correlated with outcomes. Several features of the setting support this comparison. Judges concurrently handle both criminal and non-criminal cases, and case assignment is administered through the electronic case-management system subject to docket-balancing and other assignment rules. Our identifying assumption is not that criminal and non-criminal cases are identical, but that, absent the reform, their relative evolution would not have changed discontinuously at E-Court adoption. Consistent with this assumption, the shares of detailed criminal and non-criminal case types show little systematic break around implementation. The estimated reduction in case duration also remains negative and sizable when we condition more flexibly on case type and restrict attention to more comparable judge-by-case-type and judge-by-case-subtype comparisons. Event-study estimates further show no pre-trends, and placebo assignments of treatment produce null effects tightly centered at zero. Bounding exercises following Rambachan and Roth (2023) demonstrate that only deviations several times larger than those observed could overturn the results, and estimates are unchanged when aggregating to the quarter. These exercises support the view that the reform, rather than spurious dynamics, drives our results.

Another concern is that the reform may have changed the composition of cases, so that the observed civil–criminal difference partly reflects a changing mix of disputes rather than faster case resolution. We first examine the shares of narrowly defined case types and find little change around implementation for either within civil or criminal cases. The estimates are also similar when we include detailed case-type fixed effects and, more demanding still, case-type-by-judge-fixed effects that absorb judge-specific differences in how each case type is handled. Last, we restrict the sample to narrowly defined case subtypes that the same judge handles both before and after reform. For example, we compare divorce cases handled by a given judge before reform with divorce cases handled by that same judge after reform. The estimates attenuate somewhat but remain large and negative, making changes in case composition an unlikely explanation for the decline in duration.

Finally, a separate concern is statistical precision. Because the reform was implemented in counties that are institutionally and geographically similar, outcomes across nearby courts may be correlated in ways that conventional clustering does not fully capture. We therefore consider several complementary approaches to spatial dependence and inference. We apply Conley-type spatial corrections at multiple distance thresholds (Conley, 1999), cluster standard errors at broader geographic levels, and implement the Thresholding Multiple Outcomes (TMO) procedure (DellaVigna et al., 2025). We also apply the spatially robust procedure of Müller and Watson (2022) and the wild cluster bootstrap (MacKinnon and Webb, 2018). Across these exercises, the estimated effects remain similar in sign and magnitude, although statistical precision varies somewhat across corrections and some outcomes are no longer significant under the most conservative procedures. Overall, the results are not particularly sensitive to alternative approaches to inference that allow for dependence across courts.

Related Literature

This paper relates to several strands of the literature. First, we contribute to the literature on technology and the organization of production. A large body of work emphasizes that technological change reshapes production by changing the tasks performed by workers and the allocation of effort across them (Autor et al., 2003; Acemoglu et al., 2007; Bresnahan et al., 2002; Acemoglu and Autor, 2011). Organizational theories similarly emphasize how specialization and the allocation of routine and complex tasks determine the use of scarce expertise within organizations (Holmstrom and Milgrom, 1991; Garicano, 2000). Consistent with these theories, recent work shows that organizational performance depends on how workers and managerial attention are allocated within firms (Gumpert et al., 2022; Minni, 2026). Emerging evidence on generative AI further shows that new technologies can raise productivity in knowledge-intensive work and alter how professionals perform their tasks (Noy and Zhang, 2023; Brynjolfsson et al., 2025). We complement this literature by studying a technology-enabled reform within a large public organization in which production is organized across heterogeneous tasks and interdependent stages. The reform changes neither the number of judges nor their formal responsibilities, but shifts filing, payments, and summons onto electronic systems, reducing the procedural time required to move affected cases through the court. We show that the resulting gains are likely to be concentrated in relatively routine disputes and propagate beyond the tasks directly affected by the technology: if anything, decision quality improves and lower appellate inflows reduce congestion higher in the judicial hierarchy. These findings suggest how technology can expand effective organizational capacity by changing the allocation of scarce inputs across tasks and across stages of production.

Second, we relate to the growing literature on digital and technology-enabled reforms in the public sector. This work studies biometric smartcards in welfare programs (Muralidharan et al., 2016), electronic voting machines (Fujiwara, 2015; Duarte, 2025), technology-aided instruction in public education (Muralidharan et al., 2019), e-governance systems in large public programs (Banerjee et al., 2020), and mobile platforms in agriculture (Fabregas et al., 2019).[2] We extend this literature to the judiciary, an important but comparatively understudied dimension of state capacity. Legal capacity is central to the ability of the state to enforce rights and support economic exchange (Besley and Persson, 2009), and existing work shows that judicial effectiveness shapes credit, production, and resource allocation (Ponticelli and Alencar, 2016; Boehm and Oberfield, 2020; Li and Ponticelli, 2022; Aberra and Chemin, 2021). Examining Indonesia's E-Court reform, we provide evidence that technology-enabled reforms to court administration can expand effective legal capacity by reducing case-processing time without sacrificing decision quality. Moreover, because trial and appellate courts are linked through case flows, we show that the effects of technological reform need not remain confined to the service directly targeted: by relaxing capacity constraints at

[2] For a recent review of this literature, see Haug et al. (2024).

one stage of the judicial system, technology can improve performance elsewhere in the state organization.

Last, the paper contributes to the emerging literature on technology in the judiciary. Existing work has studied how court design, judicial selection, and executive constraints shape court performance (Ash and MacLeod, 2024; Mehmood, 2022; Liu et al., 2022; Chemin, 2021). A smaller but growing literature examines how new technologies interact with judicial decision-making: machine-learning tools can improve predictions relevant to bail decisions and reveal systematic features of judicial behavior (Kleinberg et al., 2018; Ludwig and Mullainathan, 2024), while technology-enabled transparency can alter judicial conduct (Chen et al., 2026). Our setting instead studies a technology-enabled reform that changes the procedural interactions by which cases move through the court system, allowing us to shed light on how such changes can affect courts. This distinction is increasingly relevant as courts adopt AI-based tools that can be incorporated into electronic judicial workflows.

The remainder of the paper is organized as follows. Section 2 provides institutional background on the E-Court reform and its implementation. Section 3 describes the data and key variables. Section 4 outlines the empirical strategy. Section 5 presents the main results on decision time. Section 6 examines the implications of the reform for decision quality. Section 7 investigates whether digitization altered patterns of judicial bias. Section 8 explores heterogeneity, mechanisms, and discusses robustness. A final section concludes. Online Appendix A defines all variables, documents data construction, and provides additional institutional background on the Indonesian court system. Online Appendix B presents supplementary figures and tables.

2 Institutional Background

Judicial Structure. Indonesia's judiciary is formally organized as a three-tier system. At the base are first-level trial courts located in municipalities, which exercise original jurisdiction in both civil and criminal cases. We refer to these as *municipal courts*, reflecting their local reach and relatively small scale. There are roughly 850 such courts nationwide, with some municipalities hosting multiple courts while others are served by a single court.[3] Above them are second-level courts, the High Courts, seated in provincial capitals, which hear appeals. At the national level, the Supreme Court (*Mahkamah Agung*) functions as the final appellate authority, while the Constitutional Court hears constitutional reviews, elec-

[3]These include all public or secular courts (*Peradilan Umum*) that handle civil and criminal disputes; Islamic Courts or *Peradilan Agama* and administrative courts (*Peradilan Tata Usaha Negara*) oversee disputes with state agencies. These courts, together, constitute the entirety of the Indonesian adjudication at the local level.

tion disputes, and conflicts between state organs. The overwhelming bulk of judicial activity takes place in the municipal courts and, to a lesser extent, the High Courts. Accordingly, our analysis focuses on these local municipal and provincial high courts, which together constitute almost the entirety of judicial case flow in Indonesia. We assemble records for 824 municipal courts, representing nearly 97 percent of all courts, and for all appellate high courts. Figure A1 presents the corresponding case volumes: first-instance Secular Courts (*Pengadilan Negeri*) handled 3,231,294 cases, and Religious Courts handled 656,462 in 2020. The High Courts collectively reviewed tens of thousands of appeals, and the Supreme Court served as the final appellate authority (Supreme Court of the Republic of Indonesia, 2020). In Figure A2, the shaded courts contain treated cases; within each court, only a subset of case types was assigned to e-filing, e-summons, or e-hearings, predominantly civil dockets, whereas criminal cases proceeded as before. Figure A1 in Appendix A illustrates the system before and after the E-Court reform, showing the shift from manual, paper-based filing to e-filing and from in-person hearings to e-hearings.

Technology-enabled court administration has become increasingly common across judicial systems. Figure 1 documents the global prevalence of these reforms using the World Bank's Court Automation Index. By 2019, a substantial number of countries had adopted at least one electronic court feature, including e-filing, electronic service of process, online fee payments, or online access to judgments.

Against this backdrop, the Supreme Court of Indonesia introduced its E-Court system through Supreme Court Regulation No. 3 of 2018 and Regulation No. 1 of 2019. Implementation proceeded in two stages. The first phase, initiated in July 2018, introduced electronic case registration, electronic summons, and online fee payments. The second phase, in August 2019, expanded the system to include electronic submission of evidence and documents, electronic trials, and remote hearings. Because the two phases differed in both timing and scope, we later report estimates separately by reform wave.

Importantly, although the reform was implemented at the court level and adoption was staggered across courts, it also generated rich within-court variation: within any adopting court, only civil cases transitioned to the digital workflow, whereas criminal cases handled by the same judges, in the same court, and in the same year continued under the prior system. Specifically, the reform applied to civil cases in the secular courts (*Pengadilan Negeri*), family and other civil matters in the religious courts (*Peradilan Agama*), and administrative disputes in the state administrative courts (*Peradilan Tata Usaha Negara*), but not to criminal cases (Mahkamah Agung Republik Indonesia, 2018; Peraturan Mahkamah Agung Republik Indonesia, 2019). The cases in these reform-adopting courts form the basis of our analysis. In practice, the system was first introduced in 32 courts in 2018 and then extended in 2019, for a total of 69 unique first-instance courts designated to adopt the E-Court system (Putra, 2020). We observe case-level electronic records for 54 of these treated courts.[4] As a baseline, we restrict attention to within-court comparisons in these courts. We later compare treated and untreated courts and discuss what this implies for the external validity of our baseline estimates. Appendix Table B1 reports the distribution of treated and untreated courts by province and reform wave. Only about 10 percent of courts were treated before COVID-19 disrupted further expansion.

Case Assignment. The assignment of cases to judges in Indonesia follows procedures that are intended to promote impartiality and roughly balance workload across judges, and in our analysis, we treat this process as approximately random. Specifically, within each municipal court, cases are allocated to judges through an electronic case management system (Sistem Informasi Penelusuran Perkara), which assigns cases across judges within a given month according to an algorithm that, conditional on court and time, yields as-if random assignment (Mahkamah Agung Republik Indonesia, 2023). While this electronic system is designed to limit forum shopping and reduce discretion by court administrators, in practice, assignments are subject to some constraints: judges must be balanced across civil and criminal dockets, and senior judges may be more frequently assigned to complex or high-profile cases. These features imply that the assignment is not perfectly random. Nevertheless, conditional on court, judge, and time, it plausibly retains elements of as-if random variation. Our interpretation, based on the relevant legislation and administrative practice, is analogous in spirit to the argument in Chandra et al. (2023), who note that although politically salient cases can be steered, the large majority of cases can reasonably be treated as approximately randomly assigned.

Research Design. Our design exploits differential exposure to the E-Court reform within reform-adopting courts. We compare non-criminal and criminal cases handled by the same judge within the same court before and after implementation. Table 1 examines whether this comparison is affected by changes in the composition of cases around adoption, reporting changes in the shares of the ten most frequent civil and criminal case types before and after the reform. We focus on these high-frequency categories, where the number of observations and thus statistical power are greatest. Once we condition on both judge and time fixed effects, differences in case composition largely disappear, which is consistent with quasi-random case assignment and suggests that our specification is not driven by strategic sorting or systematic changes in the mix of cases handled by treated judges. Our main

[4] The administrative records list 69 courts. Our estimation sample includes 54 courts after imposing data requirements such as judge and year fixed effects, which drop courts with no within-court variation. A further 4 courts never enter our sample because no cases from these courts appear on the Supreme Court website from which we construct the dataset.

analysis focuses on 828 judges in 54 treated courts, drawn from an original universe of 24,100 judges. We retain judges who remain in a single treated court and hear both criminal and non-criminal cases in both the pre- and post-reform periods. Figure A2 documents each step of this sample construction. The design exploits within-judge variation, so that the same judge is observed handling different dockets before and after the reform. The key identifying assumption is that, conditional on judge fixed effects and common time shocks, judges are not systematically assigned a different mix of cases after the reform. The formal case assignment rules, together with an additional balance over observable judge and treated court characteristics reported in Table B2, also support this interpretation.

3 Data

Our analysis draws on a newly consolidated dataset of Indonesian court cases from 2010 to 2020. The data come from the Supreme Court of Indonesia's online repository, one of the largest publicly accessible collections of court records worldwide. We harmonize information from lower and higher courts, combining metadata on judges, courts, case types, filing and verdict dates, and case summaries, and link full-text judgments. The resulting dataset provides a system-wide view of judicial activity across Indonesia, covering both first-instance decisions and appeals. Appendix A provides additional detail on the data, including variable definitions, construction of each measure, and the underlying sources. We next briefly describe our main outcome variables and the treatment variable.

Case Duration. We begin with our primary outcome, case duration, defined as the number of days between case registration and verdict.[5] We measure it using the Indonesian Supreme Court's publicly available metadata, which reports filing and decision dates together with court identifiers, case types, and judge names. From these records, we construct the case-level decision time and classify cases using the Supreme Court's official case-type categories. We do so for both lower (first-instance) courts and higher courts. To study potential spillovers, we link first-instance courts that were directly exposed to the reform to their corresponding unexposed appellate courts in the judicial hierarchy.

Appeals and Reversals. Because appellate courts are tasked with reviewing potential legal and factual errors, appellate outcomes are often interpreted as a proxy for underlying case quality beyond the trial court (Kaplow, 1994; Shavell, 1995). Recent evidence shows that judges systematically adjust future sentencing in the direction of appeals and reversals,

[5]The estimation sample is defined using the date of case registration rather than the date of verdict. The final included registration cohort is February 28, 2020, while the administrative records were collected through June 30, 2021, providing at least 16 months of follow-up for all cases in the estimation window. By comparison, mean case duration in the estimation sample is approximately 50 days.

indicating that these outcomes provide high-content signals about decision quality (Bhuller and Sigstad, 2025). We therefore examine whether a case was appealed and, unconditional on appeal, whether the first-instance verdict was reversed. To construct these measures, we link appellate decisions to their underlying first-instance cases using the Supreme Court's online metadata. Appeal verdict summaries report the first-instance case number and the appellate outcome, which allows us to see from the court metadata whether a case was appealed. We then use GPT-assisted text analysis to code reversals and validate these codings against human readings, following Ludwig et al. (2025) in treating LLM outputs for estimation as requiring researcher-collected validation data to assess and correct misclassification.

Text-Based Measures. As a further dimension of case quality, we analyze the full-text judgments available for a subset of cases from our main dataset, which covers about 16% of the sample. These texts contain the complete court decision, including legal considerations and citations of relevant statutes. From the judgment texts, we construct measures of length (pages), statutory citation density (citations per page), and readability (reading-ease indices based on sentence and word structure). These indicators provide summary information on how detailed, citation-heavy, and accessible trial-level reasoning is, and they complement appellate outcomes when we examine whether the E-Court reform is associated with changes in judgment quality.

Treatment Definition. Our treatment indicator equals one for non-criminal cases filed after E-Court adoption in the relevant court, and zero otherwise. Treatment status is assigned using the case filing or registration month. Post-Reform is a court-by-month indicator that switches to one in the month of implementation—July 2018 for first-wave courts and August 2019 for second-wave courts.[6] Non-criminal identifies administrative and civil cases, with criminal cases serving as the comparison group. Treatment is therefore given by the interaction of Post-Reform and Non-criminal, combining variation in reform timing with differential exposure across case types within reform-exposed courts. For the baseline sample, we retain judges who handle both criminal and non-criminal cases before and after reform. This ensures that the comparison is drawn from judges observed across both case categories on each side of implementation, rather than from changes in the composition of judges over time. Appendix Table B3 reports summary statistics for all variables used in the analysis, Appendix Table A1 reports the main case types, and Online Appendix A1 provides detailed descriptions of these variables.

[6] In robustness exercises, we estimate the effects separately by reform wave.

4 Empirical Methodology

Baseline Specification. Our empirical design exploits variation within reform-adopting courts. The E-Court reform applied to non-criminal cases but not to criminal cases, even when both were heard in the same court and by the same judge. This differential exposure provides a tight comparison between treated and untreated cases. We restrict the sample to reform-adopting courts and compare non-criminal and criminal cases handled by the same judge before and after implementation. The design, therefore, holds fixed the court environment, the local timing of reform adoption, and persistent differences across judges. The estimating equation is

$$Y_c = \alpha \left(\text{Non-Criminal}_{k(c)} \times \text{Post E-Court Reform}_{i(c)t(c)}\right) + \delta_{k(c)} + \gamma_{j(c)} + \theta_{i(c)t(c)} + \mu_{i(c)k(c)} + \varepsilon_c,$$

where c indexes individual cases, and $i(c)$, $j(c)$, $t(c)$, and $k(c)$ denote the corresponding court, judge, case filing (registration) month-year, and non-criminal case category. Y_c is the case-level outcome. The indicator $\text{Non-Criminal}_{k(c)}$ equals one for administrative and civil cases, and zero for criminal cases. $\text{Post E-Court Reform}_{i(c)t(c)}$ equals one if case c was filed in court $i(c)$ on or after that court adopted E-Court, and zero otherwise.[7] Because adoption is staggered in two waves across courts, our dynamic estimates use Sun and Abraham (2021), which avoids the contamination that can arise when treatment effects vary across adoption cohorts or over time. The coefficient of interest, α, is therefore a relative intention-to-treat effect: it captures the differential post-reform change in reform-eligible non-criminal cases relative to criminal cases handled by the same judges within reform-adopting courts.[8]

The specification includes non-criminal case category fixed effect, $\delta_{k(c)}$, judge fixed effects, $\gamma_{j(c)}$, court-by-month-year fixed effects, $\theta_{i(c)t(c)}$, and court-by-non-criminal-case-category fixed effects $\mu_{i(c)k(c)}$.[9] These account, respectively, for persistent differences across case categories and judges, and for shocks common to cases heard in the same court and month, including changes in local congestion, staffing, and administrative conditions. Identification thus comes from the differential post-reform change in non-criminal relative to criminal cases within the same court and month, rather than from comparisons across reforming and non-reforming courts.

[7] Cases filed before adoption are therefore classified as pre-reform even if they remain pending and receive a verdict after implementation.

[8] Because judges handle both case types, any positive spillover of E-Court onto criminal cases through shared judicial capacity enters the comparison group. The baseline estimate therefore captures a net differential effect and is likely conservative relative to the effect on treated cases in the absence of such spillovers.

[9] The court-by-month-year fixed effects, $\theta_{i(c)t(c)}$, are also defined using the same filing month-year, so neither treatment status nor the time fixed effects depend on the verdict date or realized case duration.

Case Composition. Before turning to the effects of E-Court on case outcomes, we examine whether adoption coincided with changes in the mix of cases in the estimation sample. Table 1 reports the pre-reform share and the estimated post-adoption change in the share of each of the ten most frequent non-criminal and criminal case types. Among the non-criminal case types directly exposed to the reform, none of the ten estimated changes is statistically significant. Among criminal cases, eight of the ten shares are similarly stable, while theft and gambling cases increase by 1.29 and 0.62 percentage points, respectively. Overall, these estimates, therefore, provide little evidence of systematic re-sorting across detailed case types around E-Court implementation.

5 Main Results

Figure 2 presents the raw comparison underlying our empirical design. Before E-Court reform, mean case duration is about 52 days for criminal cases and 38 days for non-criminal cases. After adoption, criminal-case duration rises to roughly 55 days, while non-criminal-case duration falls to about 34 days. The relative gap therefore widens from approximately 14 days before reform to about 21 days afterward. Equivalently, the unconditional difference-in-differences is around -8 days, with Figure 2 providing a simple visual summary of the relative change underlying the design.

Building on this descriptive evidence, we next estimate the conditional difference-in-differences specification. Table 2 reports the results for case duration. Column 1 includes a non-criminal-case fixed effect. Column 2 adds judge fixed effects, so that the comparison exploits changes among civil and criminal cases handled by the same judges over time. Column 3 adds county-court-by-month-year fixed effects, absorbing shocks common to all cases in a court at a given point in time, including changes in local congestion. Column 4, our most saturated specification, further includes county-court-by-non-criminal-case fixed effects, absorbing persistent differences between civil and criminal cases within each court. Once judge fixed effects are included, the estimates remain broadly similar across more demanding specifications, ranging from about 7 to 9 days. In the most demanding specification, non-criminal cases become about 7 days faster after the reform relative to criminal cases handled by the same judges, a reduction equal to about 15 percent of mean case duration In aggregate terms, the reform freed nearly 700,000 judicial days, or more than 19 centuries of court time.[10] The scale of this effect highlights how seemingly modest efficiency gains at the case level can yield transformative institutional capacity when multiplied across large caseloads.

[10] This calculation is based on 100,000 civil cases filed after the reform in lower courts (in the full sample), each experiencing an average decision time reduction of about 7 days.

We next ask how realized case duration evolved relative to the path implied by pre-reform relationships. Figure 3 reports the results. We estimate the model using only pre-reform cases and then use the resulting judge and court-by-case-group effects to predict how case duration would have evolved after implementation had those relationships remained unchanged. The predicted series therefore incorporates changes over time in the composition of judges and cases, while holding fixed their pre-reform relationship with duration. We normalize the predicted series to match the pre-reform mean within criminal and non-criminal cases so that the comparison focuses on changes around implementation rather than differences in levels. Before the reform, realized and predicted duration track each other closely in both groups. After implementation, realized duration for non-criminal cases falls increasingly below its predicted path, while criminal cases remain much closer to the counterfactual. The divergence emerges only after implementation and is concentrated among reform-exposed cases.

The reduction in case duration may also have lowered the time costs associated with litigation. To provide a sense of the potential magnitude, Table B4 presents a back-of-the-envelope benchmark that values the time of litigants, lawyers, judges, and court staff using prevailing hourly wages. Under the benchmark assumptions reported in the table, including an eight-hour workday, a seven-day reduction in case duration corresponds to an illustrative wage-equivalent value of approximately USD 4,474 per case. Applied to 100,000 cases, this amounts to roughly USD 447 million. Because reductions in elapsed case duration need not translate one-for-one into active labor hours, we treat these calculations as helpful benchmarks rather than estimates of realized labor-cost savings.

We compare these benchmark values with the estimated expenditure associated with implementing the E-Court system. Based on Supreme Court budget documents and reported allocations, we estimate implementation expenditures of approximately USD 71.6 million over 2018–2020. Relative to the wage-equivalent value of time saved above, this yields an illustrative value-to-expenditure ratio of about 6.25 to 1. Panels D and E relax the benchmark assumptions in two directions. Assuming four rather than eight hours per day lowers the ratio to 3.12 to 1, while assuming a five-day rather than seven-day reduction in duration yields a ratio of about 4.46 to 1. We also report ten-year projections under the assumptions stated in the table, although these longer-run calculations are necessarily more sensitive to assumptions about costs, case volumes, and the persistence of the estimated reduction in delay.

Identification Checks. A key identification concern in our setting is that the observed decline in decision time could reflect pre-existing differences across case types rather than the causal effect of the reform. Absent the reform, decision time for civil and criminal cases would have followed parallel paths both within and across courts. If this assumption is violated, for instance, if treated non-criminal cases were already experiencing faster improvements in case management before the reform, then our estimates could overstate the true effect of the reform. We address this concern through a series of identification checks. First, the raw data suggest that treated and untreated cases follow similar pre-trends before the E-Court reform, reducing concerns that the results are driven by underlying dynamics (Figure B1). Consistent with this, Figure 4 reports event–study coefficients from a standard two–way fixed effects specification and from the Sun–Abraham estimator (Sun and Abraham, 2021), which is robust to treatment effect heterogeneity; in both, pre-treatment coefficients are close to zero, and decision time sharply declines after implementation.

Second, we implement the "honest" difference-in-differences approach of Rambachan and Roth (2023), which partially relaxes the parallel-trends assumption by allowing bounded deviations in pre-treatment trend slopes. This approach provides a quantitative sense of how severe a violation would need to be to overturn the results. This method introduces a bound M on potential deviations in the slope of pre-treatment trends and then rotates the post-treatment estimates accordingly, considering all paths within the bounded set. Figure 5 reports these exercises: Panel A replicates the baseline event-study with linear trends; Panel B rotates the estimates to account for linear deviations from parallel trends; and Panel C extends the analysis to nonlinear departures, showing robustness even when allowing for violations up to multiple times the observed pre-treatment differences between treated and control units. Across a range of values of M, the estimated effect on decision time remains statistically significant or marginally significant at conventional levels. Third, since our baseline specification uses month–year fixed effects, we assess whether the results depend on this level of temporal granularity by re-estimating the event study with quarterly aggregation. As reported in Figure B2, the estimates are nearly identical to the baseline, with no indication of differential pre-trends and a treatment effect of comparable size. This exercise provides additional reassurance that the findings are not an artifact of fine-grained timing.

Third, we examine whether changes in case composition could account for the decline in duration. If the reform shifted non-criminal cases toward disputes that are inherently quicker to resolve, the civil–criminal comparison could change even without faster processing of comparable cases. We first examine this possibility in Figure B6. The figure traces, month by month around implementation, the share of judges' caseloads accounted for by the 6 most common case types for both non-criminal and criminal cases. Among non-criminal cases, these include document modifications, child custody, divorce and marital disputes, and contract cases; among criminal cases, they include narcotics, theft, corruption, and assault. Some categories exhibit gradual changes over the event window, but there is little evidence of a break in the composition of either non-criminal or criminal caseloads at the date of reform.

We next ask whether the result is sensitive to differences in the types of disputes judges handle. Column (2) of Table B5 adds mutually exclusive case-type fixed effects, absorbing persistent differences in duration across categories such as divorce and marital disputes, contract, narcotics, and theft cases. The estimated reduction falls somewhat, from 6.9 to 4.7 days, but remains sizable and statistically significant. The table then allows case-type differences to vary even more flexibly across courts and, in the most demanding specification, includes judge-by-case-type fixed effects. The latter compares changes within the same judge handling the same sub-type case over time. It yields a still sizable reduction of case duration.

Last, Table B6 imposes a still more direct like-for-like restriction. It retains only detailed case-subtype cells observed for the same judge both before and after implementation. Thus, a judge's divorce, contract, or other narrowly defined cases contribute to the comparison only when that same subtype is observed for that judge on both sides of the reform, rather than allowing changes in the mix of cases handled over time to identify the effect. The estimated reduction remains about 5.5 days in the most saturated specification. Overall, the relative stability of case shares around implementation and the persistence of the duration effect under increasingly restrictive within-case comparisons provide little indication that changes in case composition account for the observed decline in duration.

6 Decision Quality and Upstream Spillovers

Decision Quality. A separate concern is whether the reduction in case duration came at the expense of decision quality. Faster resolution need not represent an improvement in court performance if judges achieve it by devoting less attention to each dispute or deciding cases more superficially. This potential speed–quality trade-off is an important concern in recent literature (Kondylis and Stein, 2023; Chemin et al., 2026). Our setting allows us to examine this dimension directly because a common case identifier lets us trace disputes through the judicial hierarchy, recording whether a trial-court decision is appealed and, unconditional on appeal, whether it is reversed. Appeals and reversals are not complete measures of decision quality: the decision to appeal is itself selective, and appellate courts may also disagree with lower courts for reasons other than error. Nevertheless, they provide consequential benchmarks for whether first-instance decisions are contested and subsequently overturned and, as recent work suggests, are informative about broader dimensions of decision quality (Bhuller and Sigstad, 2025).

Panel A of Table 3 reports these outcomes using the case-level metadata. Columns 1–2 show that the E-Court reform reduces the share of non-criminal cases that are appealed. In the more saturated specification, the decline is about 1.15 percentage points, or roughly 43 percent relative to the mean appeal rate. Columns 3–4 show a similar pattern for the share of cases ultimately reversed on appeal: the estimate in Column 4 is a decline of about 0.45 percentage points, approximately half of the corresponding mean.[11] These outcomes are of course not complete measures of decision quality, but they provide little indication that the reduction in duration was accompanied by greater contestation upstream or more frequent reversal of first-instance decisions.

Panel B provides complementary evidence from the subset of cases for which we observe the full text of the judgment. These data allow us to examine features of the written decision that are not available in the broader case metadata, including judgment length, readability, statutory citations, and the density of factual and legal reasoning. Several patterns emerge. First, judgments become more concise: Column 1 indicates a reduction of about 1.5 pages relative to a mean of 14.7 pages. Second, this shortening is not accompanied by a detectable deterioration in readability. The estimate in Column 2 is positive but small relative to its standard error. Third, we find no corresponding decline in the density of legal content. Citations to statutes, codes, and legal provisions increase by about 0.09 per page relative to a mean of 0.55—an increase of roughly 17 percent—although the estimate is imprecise at conventional significance levels. Reasoning per page also has a positive but imprecisely estimated coefficient. The text evidence, therefore, suggests that the reform produced shorter judgments without a detectable deterioration in readability, citation density, or reasoning density. Viewed alongside the decline in appeals and reversals, these patterns are difficult to reconcile with a simple account in which faster case resolution was achieved by sacrificing observable decision quality.

Upstream Spillovers. The decline in appeals also raises a broader question: whether improvements at the trial level propagate through the judicial hierarchy. Because higher courts receive cases originating in first-instance courts, fewer appeals at the source can ease upstream caseload pressure and thereby shorten delays beyond the courts directly affected by the reform. We examine this possibility in Table 4, using cases decided in higher courts serving jurisdictions exposed to the E-Court reform.

The estimates reveal sizable reductions in higher-court case duration. Across specifications, non-criminal cases in reform-exposed jurisdictions are resolved between 6.6 and 9.9 days faster. In the most saturated specification (Column 4), duration falls by about 8.9 days relative to a mean of 47.2 days, a reduction of roughly 20 percent. Importantly, these

[11] Appendix Table B7 further shows no detectable change in reversal rates conditional on appeal, suggesting that the decline in unconditional reversals operates primarily through fewer cases reaching appeal.

higher courts were not themselves directly exposed to the E-Court reform; their exposure arose through their appellate jurisdiction over first-instance courts in treated areas. The results therefore suggest meaningful vertical spillovers: improvements introduced lower in the judicial hierarchy are associated with shorter case duration further upstream, consistent with the decline in appeals easing congestion faced by higher courts.

The aggregate magnitude is also nontrivial. Among the 7,470 post-reform non-criminal higher-court cases in treated jurisdictions, applying the Column 4 estimate corresponds to approximately 66,000 fewer case-days spent pending. This calculation is best viewed as an approximate measure of the aggregate reduction in case duration rather than as a direct measure of judicial labor or capacity saved. Together with the decline in appeals, these patterns suggest that the effects of the reform may have extended beyond the first-instance courts in which it was implemented and reached further up the judicial hierarchy.

Higher-Court Exposure. This interpretation raises a separate measurement issue. The baseline specification classifies a higher-court jurisdiction as exposed if it contains at least one treated first-instance court, even though jurisdictions differ considerably in how closely their caseloads are linked to treated courts. Post-reform changes in which cases reach appellate courts could also affect measured higher-court duration. We therefore complement the binary indicator with continuous measures of higher-court exposure constructed entirely from pre-reform case flows. Table B8 reports three such measures: Panel A uses the higher court's pre-reform non-criminal caseload share; Panel B uses the pre-reform non-criminal caseload share of treated first-instance courts within its appellate jurisdiction; and Panel C uses the pre-reform share of non-criminal appellate inflow originating from treated first-instance courts. These measures capture variation in how closely each higher court was linked to treated non-criminal caseloads before reform.[12]

Because these exposure measures are fixed before reform, they are not mechanically affected by subsequent changes in which cases are appealed. In the richer specifications, all three measures yield the same qualitative pattern: higher courts with greater pre-reform exposure to treated non-criminal caseloads experience larger post-reform reductions in decision time. This pattern is reassuring because it is not specific to the binary jurisdiction-level exposure measure used in the baseline analysis and remains when exposure is measured using predetermined differences in the intensity of links to treated first-instance courts. We interpret this evidence as consistent with upstream spillovers being stronger where higher courts were more exposed to treated lower courts.

[12]The Supreme Court is excluded from this exercise because its nationwide jurisdiction does not admit the same localized exposure measure.

7 Judicial Bias

The evidence above suggests that faster case resolution did not come at the expense of observable decision quality. A related concern is whether the E-Court reform changed how lower-court decisions were treated on appeal across different groups of judges. Recent work has raised the broader concern that technology-based reforms may interact with existing disparities in public decision-making (Beraja and Yuchtman, 2025). In our setting, the reform could, in principle, widen such differences if its benefits accrued unevenly among judges, or narrow them if more standardized procedures reduced the scope for differential treatment. We examine this question using two judge characteristics emphasized in prior work: gender and religion (Shayo and Zussman, 2011; Ash et al., 2025).[13]

We first consider gender. Previous work documents differences in appellate reversal rates for decisions authored by female lower-court judges (Ash et al., 2024). We therefore ask whether the E-Court reform changed the relative likelihood that decisions authored by female judges were reversed on appeal. Panel A of Table 5 reports the results. The estimates become negative as we add judge and court-by-time fixed effects and are quite stable across the more saturated specifications. In Column 4, the relative reversal rate of decisions authored by female judges falls by about 0.50 percentage points, compared with an overall reversal rate of 0.87 percent. This estimate should not be interpreted as a complete measure of gender bias, but it provides no indication that the reform increased adverse differential treatment of decisions authored by female judges. If anything, the differential in appellate review moves in the opposite direction.

Panel B examines the corresponding pattern by religion. We ask whether the reform changed the relative reversal rate of decisions authored by non-Muslim judges. Across specifications, the estimates are small relative to their standard errors and statistically indistinguishable from zero. In the most saturated specification, the estimated change is 0.08 percentage points with a standard error of 0.21. We therefore find little evidence that the reform changed differential appellate review along this dimension.

These results provide no evidence that the gains in speed documented above were accompanied by greater differential treatment in appellate review along the judge characteristics we can observe. The gender differential declines after the reform, while the corresponding differential by religion changes little. These results are necessarily narrower than a general test of judicial bias, but they provide some indication that the reform did not generate an observable trade-off along these dimensions.

[13] The available case metadata do not systematically record litigant gender or religion. We therefore cannot examine differential treatment at the litigant–judge level and instead study whether appellate courts treat decisions authored by different lower-court judges differently after the reform.

8 Mechanisms and Robustness

Heterogeneous Effects. We next examine where the reductions in case duration are concentrated. Figure 6 reports estimates separately across major categories of non-criminal cases, with criminal cases serving as the comparison group in each specification. Panel A shows sizable reductions for family property and inheritance disputes and for child custody cases, while the estimate for divorce and marital disputes is small and statistically indistinguishable from zero. Panel B shows similarly sizable reductions for document modifications, contract disputes, and the residual category of other civil cases, whereas tort and other non-contract disputes change little. Thus, the average effect masks meaningful heterogeneity across types of disputes.

The pattern is suggestive of larger gains in cases for which procedural and documentary steps are relatively important. Document modifications, for example, experience a clear reduction in duration, while more contentious divorce and marital disputes show little change. At the same time, the reductions for family property, inheritance, and contract disputes caution against interpreting the results as a simple distinction between "easy" and 'difficult" cases. The larger reductions in relatively standardized case types, however, raise the possibility that E-Court changed how judges allocate time across tasks. We examine this next.

Task Reallocation. If these differential time savings free judicial capacity, some of that capacity may be redirected toward tasks that are more complex and harder to automate (Autor et al., 2003; Garicano, 2000). We first examine this possibility using predetermined measures of case complexity. Table B9 shows that the reduction in duration attenuates with pre-reform complexity: treatment effects are smaller for case types with greater ex ante complexity. Because these measures are constructed before implementation, they capture persistent differences in the demands of different case types rather than responses to the reform itself. The pattern is therefore consistent with larger time savings in case types requiring less legal work before implementation.

We next ask whether this pattern extends to other cases handled by the same judges. For each judge, we construct a predetermined exposure measure—the pre-reform share of cases that were non-criminal—and examine subsequent changes in criminal-case duration. Table B10 shows that criminal-case duration falls more after reform for judges with greater pre-reform non-criminal exposure, including in specifications with judge-by-month-year fixed effects. Because criminal cases themselves were untreated, this differential improvement is consistent with judges reallocating some capacity toward other parts of their caseload.

The corresponding binned relationship in Figure B7 is also negative, although less precisely estimated ($p = 0.117$). We do not find a comparable increase in the number of criminal or total cases resolved, suggesting that the adjustment appears primarily in processing time rather than case counts. Overall, these patterns are consistent with some reallocation of judicial time across tasks.

Spillovers. The within-judge evidence above asks whether changes in untreated criminal cases are larger for judges with greater exposure to reform-eligible work. A related question is whether criminal cases improve on average after E-Court adoption. Our baseline design cannot identify the level of any such spillover: because it compares non-criminal and criminal cases within the same reform-adopting court and period, changes common to both types of cases are absorbed by the court-by-time fixed effects.

To obtain a rough benchmark for changes in criminal cases themselves, we therefore turn to the broader sample and compare criminal cases in reform-adopting courts with criminal cases in courts that did not adopt E-Court. Table B11 reports these results. In the most saturated specification, criminal-case duration falls by about 2.5 days in reform-adopting courts relative to non-adopting courts, even though criminal cases were not directly covered by the reform. Non-criminal cases experience an additional reduction of about 6.6 days, close to the 6.9-day estimate in our baseline sample and specification. The decline among untreated criminal cases is consistent with an average spillover from the reform, although we interpret this comparison cautiously because it relies on differences between adopting and non-adopting courts.

Broader Sample. Beyond its use in the spillover exercise, the broader sample provides a separate check on whether the baseline estimates are specific to the courts and judges included in our preferred design. Our baseline analysis focused on reform-adopting courts and on judges who remain in the same court and handle both criminal and non-criminal cases before and after implementation. This restriction provides the within-court comparison used for identification, but it also means that the baseline estimates are most directly informative about this subset of judges and courts. Appendix Table B12 shows that reform-adopting regions differ from other regions along several pre-existing characteristics: they are larger, less poor, and have higher expenditure, as well as more businesses, markets, and hospitals. Appendix Table B13, similarly documents differences in case composition between the treated-judge sample and the remainder of the court system. These differences motivate our decision not to use never-treated courts as the primary comparison group.

As a broader-sample check, Table B14 re-estimates the main outcomes using more than three million cases from nearly 800 courts, including courts that never adopted the E-Court reform. The estimates become progressively closer to the baseline as the specification absorbs judge heterogeneity, court-specific time shocks, and persistent differences between criminal and non-criminal cases within courts. In the most saturated specification, the estimates for case duration, appeals, and reversals are very close to those in the treated-courts sample. We view this as evidence that the main patterns are not specific to the more selected baseline sample, while retaining the treated-courts specification as our baseline design.[14]

Filing Responses. A distinct question is whether faster case resolution changes the demand for courts. By lowering the time required to obtain a decision, the E-Court reform could induce additional cases to enter the system. Such a response would itself be economically meaningful, but it could also absorb part of the capacity released by the reform and make sustained reductions in case duration harder to achieve. Table B15 examines this possibility. Across specifications, the estimated changes in filings are small, statistically indistinguishable from zero, and exhibit no systematic pattern in sign. We therefore find little evidence that the reform materially changed the inflow of new cases over our sample period. The reduction in case duration thus does not appear to have been offset by an accompanying increase in court use. This muted filing response is consistent with evidence showing that reductions in case delay need not translate into higher filing rates when information and other barriers continue to shape access to formal courts (Acemoglu et al., 2020).

Reform Waves. The staged rollout also allows us to examine whether the effects vary across successive phases of the E-Court reform. The first wave began on July 13, 2018, covering 30 of 809 municipal courts and introducing electronic registration and payments. The second wave began on August 19, 2019, extending the reform to an additional set of courts and adding electronic submissions and hearings to the functions introduced in the first wave (Mahkamah Agung Republik Indonesia, 2019). First-wave courts also received these additional functions, while second-wave courts entered E-Court with the expanded set of features. The staged rollout therefore allows us to examine whether the estimated effects differ across waves as the reform expanded in scope, although differences across waves cannot be attributed solely to the additional features because the waves also involve different courts and implementation periods. Table B16 reports estimates separately for the two waves. Case duration falls in both waves across specifications. The estimated reduction is larger for the first wave in the most saturated specification, where equality of the two wave-specific coefficients is rejected, although the differences are smaller and not statistically distinguishable in the intermediate specifications. The appellate outcomes show a more

[14] Appendix C. reports analogous broader-sample exercises for the remaining decision-quality measures constructed from judgment texts, using the subsample for which full-text decisions are available.

similar pattern: appeals and reversals decline in both waves, although some wave-specific estimates are imprecisely estimated, and the equality tests do not reject common effects across waves. The results suggest some heterogeneity in the case-duration response as the E-Court system expanded, although the evidence is not sufficiently consistent for us to draw strong conclusions.

Permutation Inference. We conduct placebo exercises that randomly reassign treatment across case types as an identification check. The resulting distributions of placebo estimates, displayed in Figure B3, are tightly centered around zero with no pronounced structure. This pattern is consistent with the view that our main effects are not solely artifacts of random variation and supports the identifying assumption that criminal cases within a court and judge provide a relevant counterfactual for civil cases before and after the reform. Evidence from raw trends, event-studies, bounding exercises, alternative temporal aggregation, and permutation tests is therefore consistent with the interpretation that the estimated decline in decision time reflects the impact of the reform rather than purely spurious dynamics.

Standard Errors. The results also do not appear sensitive to alternative clustering choices. In the baseline, we cluster at the municipal court level, which corresponds to the level of variation in reform implementation. In Table B32, we re-estimate the main specification clustering at the narrower judge level (Panel A) and at the broader municipality level (Panel B). Across these alternatives, the estimated effects remain statistically significant. There are also concerns about moderate cluster counts since our inference relies on asymptotic theory. To speak to this concern, we implement the Wild Cluster Bootstrap method of MacKinnon and Webb (2018), which does not depend on large-sample approximations. Figure B4 presents bootstrap confidence intervals and p-values for the main outcomes, and inference remains significant at conventional levels. We next address potential spatial correlation across districts using Conley standard errors (Conley, 1999) with distance cutoffs based on the geographic location of courts. While Conley's approach is now standard, it can understate uncertainty if the decay of spatial correlation is misspecified or does not decays with distance. As a complementary check, we implement the spatially robust inference procedure of Müller and Watson (2022), which treats spatial dependence as a worst-case scenario and does not rely on distance-based cut-offs. Following Müller and Watson (2022), we impose an upper bound ρ on the average pairwise spatial correlation and construct standard errors and critical values from the corresponding worst-case covariance matrix. Under this procedure, confidence intervals retain nominal coverage for any data-generating process with spatial dependence no stronger than ρ. Table B33 reports these robustness checks. The estimates remain stable across bandwidth choices and worst-case spatial correlation scenarios, suggesting that spatial dependence is not a key driver of our results.

Multiple Hypotheses. We also examine whether inference on our main outcomes is sensitive to multiple-hypothesis adjustments. We treat case duration, appeals, and reversals as a single family of outcomes and report both sharpened q-values, which adjust for the false discovery rate, and family-wise-error-rate adjusted p-values in Table B34. The conventional p-values for the three outcomes range from 0.003 to 0.026. After adjustment, the sharpened q-values range from 0.009 to 0.026, while the family-wise-error-rate adjusted p-values range from 0.009 to 0.027. Thus, the reductions in case duration, appeals, and reversals remain statistically distinguishable from zero under both approaches.[15] These results provide some reassurance that inference on the main outcomes is not particularly sensitive to accounting for multiple testing.

Alternative Fixed Effects. We also examine whether the results are sensitive to more demanding fixed-effects structures. Panel A of Table B35 allows for province-by-year shocks and persistent differences across non-criminal case types within provinces. Panel B provides a particularly stringent version of our within-judge comparison by including judge-by-month-year fixed effects. Because judges in our estimation sample remain in the same court, this specification compares criminal and non-criminal cases handled by the same judge in the same court and month-year, absorbing any contemporaneous shocks specific to that judge. It therefore rules out identification from criminal and non-criminal cases handled by different judges within the same court-month. The estimates remain similar to the baseline. In the most saturated specification, case duration falls by 7.36 days, appeals by 0.80 percentage points, and reversals by 0.40 percentage points. Importantly, this specification sharpens the within-judge comparison, but the coefficient continues to capture the relative change in non-criminal versus criminal cases: any reform-induced improvement common to both case types handled by the same judge in a given month is differenced out. As discussed in the Spillovers subsection above, we separately used the broader sample to examine whether untreated criminal cases themselves also improve following adoption.

Sample Restrictions. We also assess the robustness of our findings to alternative sample restrictions designed to address influential observations and partial exposure to the reform. First, we exclude the largest courts, which may otherwise exert disproportionate influence on the estimates. Second, we drop observations from the implementation quarter to mitigate concerns about partial reform exposure. Table B36 reports the results. The estimated effects across outcomes remain broadly consistent with the main specifications,

[15] Because the judgment-text measures are available only for a smaller subsample and capture a distinct dimension of the analysis, we treat them as a separate family and report the corresponding multiple-testing adjustments in Appendix C.

suggesting that the findings are not driven by specific subsets of the sample.

Manual Validation of Reversals. In addition to these robustness checks, it is important to verify that our outcome measures themselves are accurately coded. To benchmark the accuracy of our automated procedure for identifying case reversals, we manually validated a randomly drawn subsample of 935 appeals to determine if they were indeed reversed. Each of these appeals was hand-coded to determine whether it resulted in a reversal, the original order was upheld, or could not be reliably classified. Panel A of Table B23 reports the results of the GPT-based classification, while Panel B presents the outcomes of the manual validation. Out of the 758 manually coded appeals, 281 (37.07%) resulted in a reversal, 461 (60.82%) were upheld, and 16 (2.11%) could not be reliably classified as reversed or upheld. The GPT-based model produces distributions that are essentially similar with the manual benchmark: 32.6% of appeals were identified as reversals by GPT compared to 37.1% in the hand-coded sample, while the unclassified share is higher in the automated classification (77.8%). This pattern suggests that the model takes a more cautious stance than human coders, leaving cases unclassified when the evidence is ambiguous and thereby reducing the risk of mislabeling outcomes.

Debiasing GPT-Classified Reversals. Building on this validation exercise, we next examine whether any systematic differences between GPT-based and manually coded outcomes translate into bias in our regression estimates. Following the econometric framework suggested in Ludwig et al. (2025), we use this validation sample to estimate the extent of measurement bias and adjust the regression coefficients in Table 3 accordingly. Panel C of Table B23 reports regression estimates on the validation data, where outcomes coded manually are compared against those classified by GPT for the same specifications and sample size. The difference between the two sets of coefficients yields a small misclassification bias of 0.001. We then adjust the main GPT-based estimate in Table 3 (-0.451) by this bias term, yielding a debiased estimate of -0.450, which is virtually identical to the baseline estimate. The validation exercise for the LLM-classified reversals, therefore, suggests that GPT-based classification introduces only a modest bias, which can be explicitly corrected using the manual subsample. Importantly, only the reversal outcome relies on GPT classification; other outcomes, such as decision time and appeals, are measured directly from the court records. This suggests that concerns about misclassification are limited in scope, and the overall pattern of results is unlikely to be materially affected by this source of measurement error.[16]

[16]For brevity, we presented robustness checks for our primary outcomes in this section. Appendix C reports the corresponding exercises for the remaining outcomes, including the text-based measures of decision quality, which are available for a smaller subsample.

9 Concluding Remarks

This paper provides empirical evidence on how technology-enabled changes in administrative procedures affect the functioning of an interdependent public organization. We study Indonesia's phased rollout of the E-Court reform, exploiting the fact that it applied to non-criminal but not criminal cases within the same reform-adopting courts. Our design compares changes across these cases while absorbing judge, case-type, and court-by-month-year fixed effects. We find that case duration falls by about 15 percent following the reform. We find little evidence that this reduction in delay came at the expense of the dimensions of decision quality we can observe: appeals and reversals decline, while judgments become shorter without detectable declines in readability or statutory citation density. Case filings and the composition of cases also change little around implementation.

A central feature of the setting is that courts form an interdependent system: decisions made at one level generate workload at the next. The effects extend beyond the first-instance courts directly exposed to the reform. The decline in appeals provides one potential channel through which changes at the trial level may relax capacity constraints further up the judicial hierarchy. Consistent with this possibility, case duration in higher courts serving reform-exposed jurisdictions falls by about 9 days, or roughly 20 percent relative to the mean. The same pattern emerges when higher-court exposure is measured using only pre-reform case flows, with larger reductions in duration among courts more exposed ex ante to treated caseloads. These results suggest that changes in case processing at one stage of the judicial system can have consequences for workload and delay at subsequent stages. This vertical spillover is important for evaluating technologies introduced into organizations in which output at one stage becomes an input or source of workload at another.

The pattern across cases also provides some indication of where the reductions in delay are concentrated. Reductions in duration are larger for case types in which procedural and documentary steps appear relatively important, including documentation changes and inheritance cases, and extend to medium-complexity disputes such as child custody. By contrast, duration changes little for more complex marital and divorce disputes. Predetermined measures of case complexity show a similar pattern, with smaller reductions for case types that required more legal work before the reform. Our results are consistent with E-Court generating larger time savings where standardized procedural work accounts for a greater share of case processing. The main results are also robust to alternative approaches to case composition, pre-existing trends, treatment timing, and spatially correlated inference.

How far these findings generalize will naturally depend on the institutional setting and on which constraints limit court performance. The reform we study is comparatively incremental: it changes how cases are filed and paid for, rather than changing substantive law, or the formal allocation of authority. The findings may thus be most relevant to settings in which procedural and administrative frictions contribute importantly to case delay. In such settings, our results suggest that relatively targeted changes in case processing can generate time savings across large caseloads and, where stages of production are linked, may have consequences beyond the point at which the technology is introduced.

More broadly, our findings suggest how the effects of administrative technology can depend on the organization in which it is introduced. In organizations built around interconnected tasks, the consequences of a technology need not be confined to the workers or activities that use it directly. Changes in one part of the production process can alter the demands placed on others. The Indonesian E-Court reform provides one example: a targeted change in court procedure was associated with shorter case duration among directly affected cases, little deterioration in the observable dimensions of decision quality, and reductions in delay further up the judicial hierarchy.

References

Aberra, Adam and Matthieu Chemin, "Does legal representation increase investment? Evidence from a field experiment in Kenya," *Journal of Development Economics*, 2021, *150*, 102612.

Acemoglu, Daron, Ali Cheema, Asim I Khwaja, and James A Robinson, "Trust in state and nonstate actors: Evidence from dispute resolution in Pakistan," *Journal of Political Economy*, 2020, *128* (8), 3090–3147.

_ **and David Autor**, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4B, Elsevier, 2011, chapter 12, pp. 1043–1171.

_ **, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti**, "Technology, Information, and the Decentralization of the Firm," *The Quarterly Journal of Economics*, 2007, *122* (4), 1759–1799.

Ash, Elliott and Stephen Hansen, "Text algorithms in economics," *Annual Review of Economics*, 2023, *15*, 659–688.

_ **and W. Bentley MacLeod**, "Mandatory Retirement for Judges Improved the Performance of US State Supreme Courts," *American Economic Journal: Economic Policy*, February 2024, *16* (1), 518–48.

_ **, Daniel L Chen, and Arianna Ornaghi**, "Gender attitudes in the judiciary: Evidence from US circuit courts," *American Economic Journal: Applied Economics*, 2024, *16* (1), 314–350.

_ **, Sam Asher, Aditi Bhowmick, Sandeep Bhupatiraju, Daniel Chen, Tanaya Devi, Christoph Goessmann, Paul Novosad, and Bilal Siddiqi**, "In-group bias in the Indian judiciary: Evidence from 5 million criminal cases," *Review of Economics and Statistics*, 2025, pp. 1–45.

Autor, David H., Frank Levy, and Richard J. Murnane, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 2003, *118* (4), 1279–1333.

Banerjee, Abhijit, Esther Duflo, Clement Imbert, Santhosh Mathew, and Rohini Pande, "E-governance, accountability, and leakage in public programs: Experimental evidence from a financial management reform in India," *American Economic Journal: Applied Economics*, 2020, *12* (4), 39–72.

Beraja, Martin and Noam Yuchtman, "Generalized Disruption: Society, Work, and Property Rights in the Age of AI," *NBER Chapters*, 2025.

Besley, Timothy and Torsten Persson, "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, 2009, *99* (4), 1218–44.

_ , Robin Burgess, Adnan Khan, and Guo Xu, "Bureaucracy and development," *Annual Review of Economics*, 2022, *14* (1), 397–424.

Bhuller, Manudeep and Henrik Sigstad, "Feedback and learning: The causal effects of reversals on judicial decision-making," *Review of Economic Studies*, 2025, *92* (4), 2359–2397.

Boehm, Johannes and Ezra Oberfield, "Misallocation in the Market for Inputs: Enforcement and the Organization of Production," *The Quarterly Journal of Economics*, 2020, *135* (4), 2007–2058.

Bosio, Erica and Philip McGrath, "Reforming Justice: Fostering Engagement Between Statutory Courts and Customary Law Systems," Technical Report, World Bank 2025.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt, "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *The Quarterly Journal of Economics*, 2002, *117* (1), 339–376.

Brynjolfsson, Erik, Danielle Li, and Lindsey R. Raymond, "Generative AI at Work," *The Quarterly Journal of Economics*, 2025, *140* (2), 889–942.

Butt, Simon and Tim Lindsey, *Indonesian law*, Oxford University Press, 2018.

Chandra, Aparna, Sital Kalantry, and William H. J. Hubbard, *Court on Trial: A Data-Driven Account of the Supreme Court of India*, New Delhi: Penguin Random House India, 2023.

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

_ , Daniel L. Chen, Vincenzo Di Maro, Paul Kimalu, Momanyi Mokaya, and Manuel Ramos-Maqueda, "Making Information Actionable: Experimental Evidence from Kenyan Courts," *Journal of Development Economics*, 2026, *183*, 103852.

Chen, Heng, Yuyu Chen, and Qingxu Yang, "Women in the Courtroom: Technology and Justice," *The Review of Economic Studies*, 2026, *93* (3), 1574–1601.

Conley, T.G., "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

DellaVigna, Stefano, Guido Imbens, Woojin Kim, and David M Ritzwoller, "Using Multiple Outcomes to Adjust Standard Errors for Spatial Correlation," Working Paper 33716, National Bureau of Economic Research 2025.

Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Courts," *The Quarterly Journal of Economics*, 2003, *118* (2), 453–517.

Duarte, Raúl, "Voting Technology, Political Competition, and Clientelism," February 2025.

ERI Economic Research Institute, "Occupation-Specific Salary Estimates for Indonesia," 2025. Judge & Magistrate Salary, Lawyer Salary, and Court Clerk Salary. Accessed October 8, 2025. Available at: https://www.salaryexpert.com/salary/job/judge-and-magistrate/indonesia; https://www.salaryexpert.com/salary/job/lawyer/indonesia; https://www.salaryexpert.com/salary/job/court-clerk/indonesia.

Fabregas, Raissa, Michael Kremer, and Frank Schilbach, "Realizing the potential of digital development: The case of agricultural advice," *Science*, 2019, *366* (6471), eaay3038.

Fujiwara, Thomas, "Voting technology, political responsiveness, and infant health: Evidence from Brazil," *Econometrica*, 2015, *83* (2), 423–464.

Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production," *Journal of political economy*, 2000, *108* (5), 874–904.

Gumpert, Anna, Henrike Steimer, and Manfred Antoni, "Firm Organization with Multiple Establishments," *The Quarterly Journal of Economics*, 2022, *137* (2), 1091–1138.

Haug, Nathalie, Sorin Dan, and Ines Mergel, "Digitally-induced change in the public sector: a systematic review and research agenda," *Public Management Review*, 2024, *26* (7), 1963–1987.

Holmstrom, Bengt and Paul Milgrom, "Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design," *The Journal of Law, Economics, and Organization*, 1991, *7* (special_issue), 24–52.

Institute for Fiscal Studies, "Crown Court backlog exacerbated by post-pandemic productivity slump," June 2025. Accessed September 24, 2025.

Kaplow, Louis, "The Value of Accuracy in Adjudication: An Economic Analysis," *Journal of Legal Studies*, 1994, *23* (1), 307–401.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan, "Human Decisions and Machine Predictions," *The Quarterly Journal of Economics*, 2018, *133* (1), 237–293.

Kondylis, Florence and Mattea Stein, "The Speed of Justice," *The Review of Economics and Statistics*, 2023, *105* (3), 596–613.

Li, Bo and Jacopo Ponticelli, "Going bankrupt in China," *Review of Finance*, 2022, *26* (3), 449–486.

Liu, Ernest, Yi Lu, Wenwei Peng, and Shaoda Wang, "Judicial independence, local protectionism, and economic integration: Evidence from China," Technical Report, National Bureau of Economic Research 2022.

Ludwig, Jens and Sendhil Mullainathan, "Machine Learning as a Tool for Hypothesis Generation," *The Quarterly Journal of Economics*, 2024, *139* (2), 751–827.

_ , _ , and Ashesh Rambachan, "Large Language Models: An Applied Econometric Framework," Working Paper 33344, National Bureau of Economic Research January 2025.

MacKinnon, James G and Matthew D Webb, "The wild bootstrap for few (treated) clusters," *The Econometrics Journal*, 2018, *21* (2), 114–135.

Mahkamah Agung Republik Indonesia, "Peraturan Mahkamah Agung Republik Indonesia Nomor 3 Tahun 2018 tentang Administrasi Pengadilan secara Elektronik," 2018. Supreme Court Regulation No. 3 of 2018 on Electronic Court Administration. Accessed via the official repository of the Government of Indonesia.

_ , *Buku Panduan e-Court Mahkamah Agung 2019* Mahkamah Agung Republik Indonesia 2019. Accessed: 2026-08-23.

_ , "Menuju Usia Ke-78, Mahkamah Agung Luncurkan 5 Aplikasi Baru," 2023. [Accessed 24 September 2025].

Mehmood, Sultan, "The impact of Presidential appointment of judges: Montesquieu or the Federalists?," *American Economic Journal: Applied Economics*, 2022, *14* (4), 411–445.

Minni, Virginia, "Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs," *The Quarterly Journal of Economics*, 2026, *141* (3), 1871–1920.

Müller, Ulrich K and Mark W Watson, "Spatial correlation robust inference," *Econometrica*, 2022, *90* (6), 2901–2935.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J Ganimian, "Disrupting education? Experimental evidence on technology-aided instruction in India," *American Economic Review*, 2019, *109* (4), 1426–1460.

_ **, Paul Niehaus, and Sandip Sukhtankar**, "Building state capacity: Evidence from biometric smartcards in India," *American Economic Review*, 2016, *106* (10), 2895–2929.

New York Times, "A Lifelong Nightmare: Seeking Justice in India's Overwhelmed Courts," 2024. [Accessed 22 November 2024].

Noy, Shakked and Whitney Zhang, "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence," *Science*, 2023, *381* (6654), 187–192.

Papaioannou, Elias and Stavroula Karatza, "The Greek Justice System: Collapse and Reform," CEPR Discussion Paper DP12731, Centre for Economic Policy Research February 2018.

Peraturan Mahkamah Agung Republik Indonesia, "Peraturan Mahkamah Agung Republik Indonesia Nomor 1 Tahun 2019 tentang Administrasi Pengadilan secara Elektronik," 2019. Supreme Court Regulation No. 1 of 2019 on Electronic Court Administration. Accessed via the official repository of the Government of Indonesia.

Ponticelli, Jacopo and Leonardo S. Alencar, "Court Enforcement, Bank Loans, and Firm Investment: Evidence from a Bankruptcy Reform in Brazil," *The Quarterly Journal of Economics*, 2016, *131* (3), 1365–1413.

Porta, Rafael La, Florencio Lopez de Silanes, and Andrei Shleifer, "The economic consequences of legal origins," *Journal of economic literature*, 2008, *46* (2), 285–332.

Putra, Dedi, "A modern judicial system in Indonesia: legal breakthrough of e-court and e-legal proceeding," *Jurnal Hukum dan Peradilan*, 2020, *9* (2), 275–297.

Rambachan, Ashesh and Jonathan Roth, "A more credible approach to parallel trends," *Review of Economic Studies*, 2023, *90* (5), 2555–2591.

Shavell, Steven, "The Appeals Process as a Means of Error Correction," *Journal of Legal Studies*, 1995, *24* (2), 379–426.

Shayo, Moses and Asaf Zussman, "Judicial ingroup bias in the shadow of terrorism," *The Quarterly journal of economics*, 2011, *126* (3), 1447–1484.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of econometrics*, 2021, *225* (2), 175–199.

Supreme Court of the Republic of Indonesia, "Peroleh Tambahan Anggaran Lewat BA BUN, MA Prioritaskan Implementasi E-Court dan Perumahan Hakim [Received Additional Budget via BA BUN, Supreme Court Prioritizes Implementation of E-Court and Judges' Housing]," May 2019. News release.

_, "Annual Report 2020: Judicial Reform and E-Court Expansion (Laporan Tahunan 2020)," 2020.

Tim, Lindsey and Helen Pausacker, *Crime and Punishment in Indonesia*, London: Routledge, 2020.

U.S. Government Accountability Office, "U.S. Immigration Courts See Significant and Growing Backlog," 2024. [Accessed 29 November 2024].

World Bank, "Digital Progress and Trends Report 2023," 2023a. [Accessed 29 November 2024].

_, "Digitizing Court Systems: Benefits and Limitations," World Bank Open Knowledge Repository 2023b. [Accessed 27 November 2024].

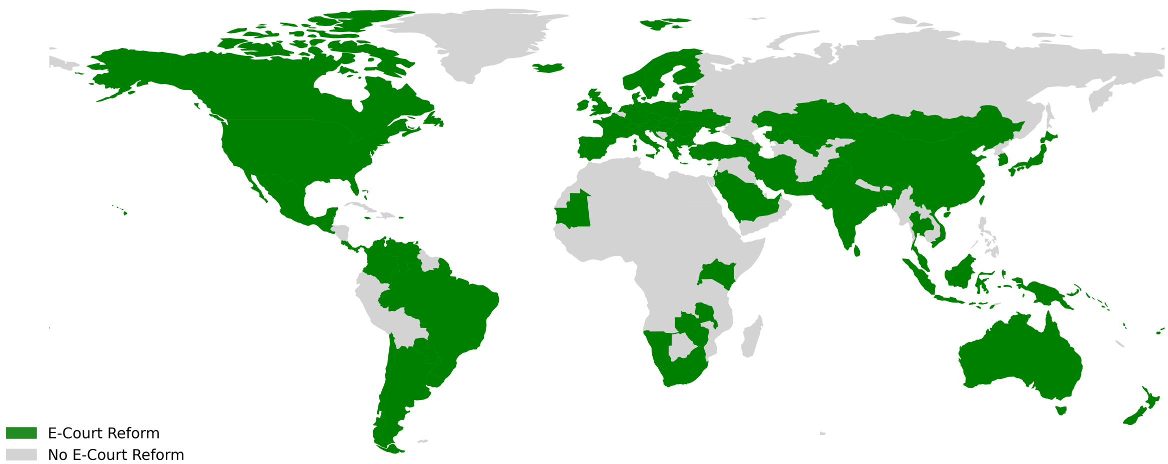
_, "Reforming Justice: Improving Service Delivery through Technology," Technical Report, World Bank 2024.

World Bank Group, "World Bank Doing Business Database," 2019. [Accessed: 23 August 2025].

Yoong, Pui Shen and Frederico Gil Sander, "Structural Transformation and Labor Productivity in Indonesia: Where Are All the Good Jobs?," Technical Report, World Bank, Washington, DC 2020.

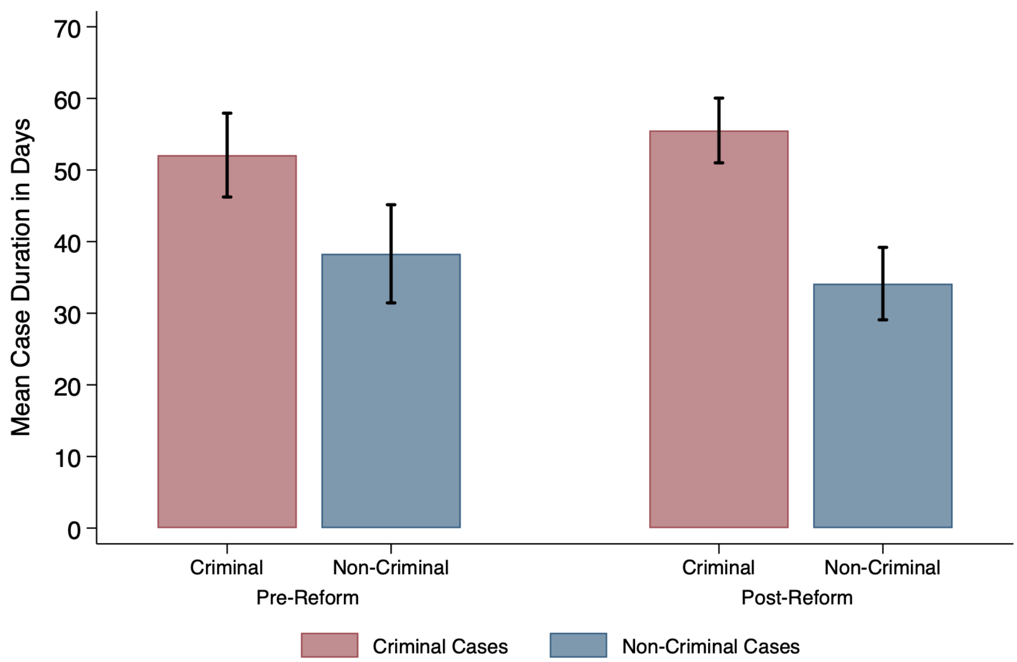
Figures and Tables

Figure 1: E-Court Reforms across the World



Note: The figure presents a map showing the presence of court automation features in 2019 across countries. Economies are shaded in green if the World Bank Court Automation Index is greater than zero, indicating the presence of at least one of the following features: (i) electronic filing of the initial complaint, (ii) electronic service of the initial complaint, (iii) electronic payment of court fees, or (iv) online public availability of commercial judgments. Economies with a value of zero are shaded in gray. Many of these electronic features were introduced in countries such as the United States, Brazil, Argentina, India, China, Malaysia, Kenya, Tanzania, and many others. A discussion of some of these reforms in specific countries can be found in Table B37. The data are from the World Bank's Doing Business report (World Bank Group, 2019).

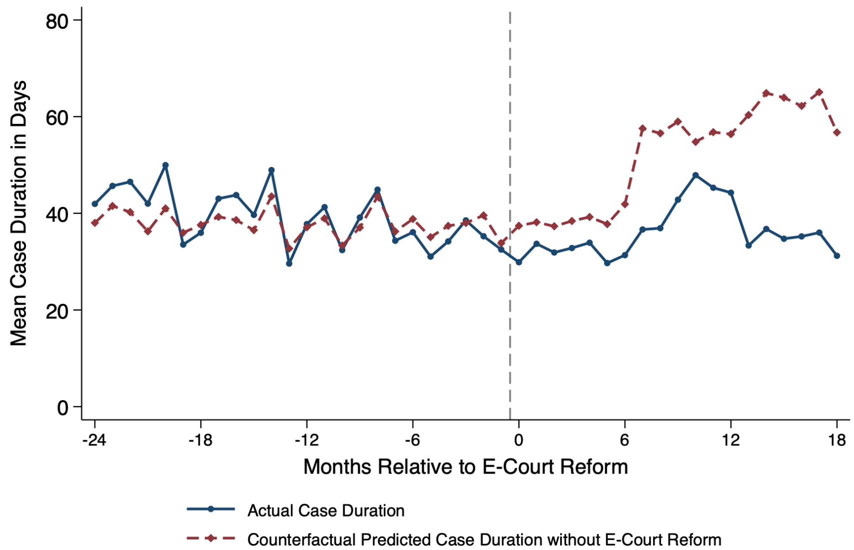
Figure 2: Mean Case Duration Non-criminal vs Criminal Case, Before and After E-Court Reform



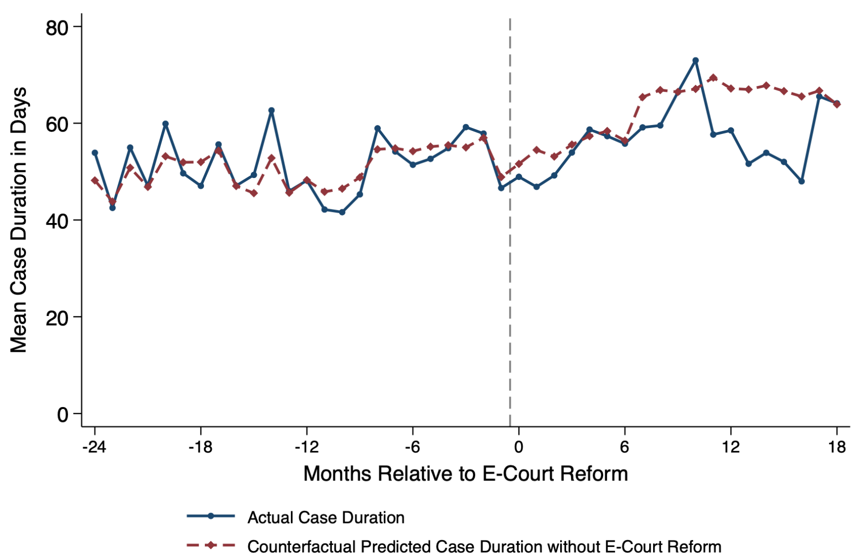
Note: The figure reports realized mean case duration separately for criminal and non-criminal cases before and after E-Court reform. Bars show group-specific means. Black vertical lines report 95 percent confidence intervals based on standard errors clustered at the county-court level. The unit of observation is a court case. The sample comprises cases filed between April 1, 2009, and February 28, 2020, in the treated-courts sample. Judges remain in the same court throughout the sample period and are observed before and after reform implementation with both civil and criminal cases.

Figure 3: Actual and Counterfactual Case Duration

Panel A. Non-criminal Cases

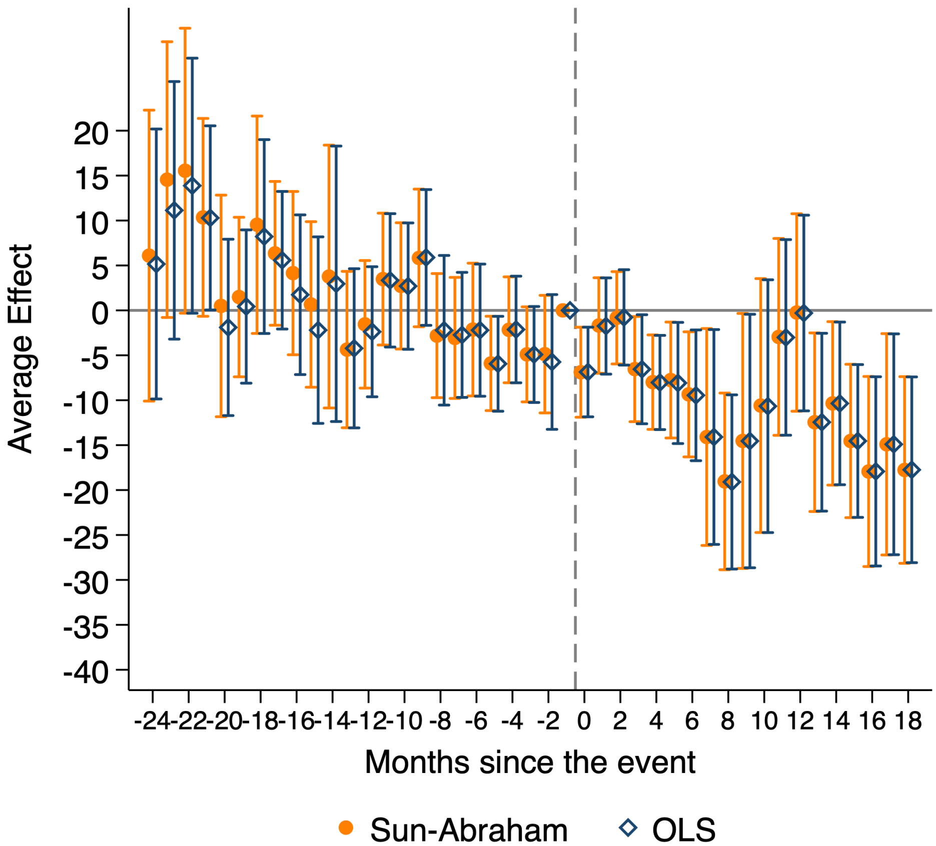


Panel B. Criminal Cases



Note: The figure compares realized mean case duration with a counterfactual predicted duration series after E-Court reform. The counterfactual is constructed in three steps. First, the model is estimated only on pre-reform cases, using case duration in days as the dependent variable and absorbing judge fixed effects, a non-criminal-case fixed effect, and county-court-by-non-criminal-case fixed effects. Second, the pre-reform fixed-effect components are applied to all cases before and after reform implementation. The resulting series therefore answers: what average duration would be expected in each event month if the relationship between duration and the pre-reform composition of judges, courts, and the criminal/non-criminal margin remained unchanged and E-Court had no direct effect. Third, the predicted values are normalized to match the pre-reform mean duration within each case group, separately for non-criminal and criminal cases, so levels are comparable to realized duration. Actual Case Duration is the realized mean duration in the same event month and case group. The vertical dashed line marks E-Court reform implementation. The sample is the main treated-courts case-duration sample: treated courts, cases filed between April 1, 2009, and February 28, 2020, and judges who remained in the same court and adjudicated both criminal and non-criminal cases before and after reform implementation.

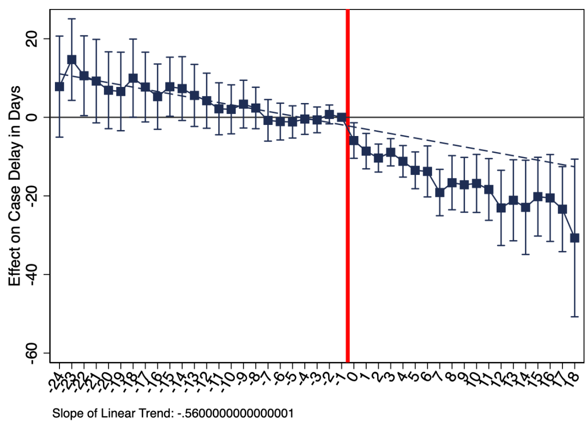
Figure 4: Case Duration in Days Over Time – Difference-in-Differences – Treated Courts Sample



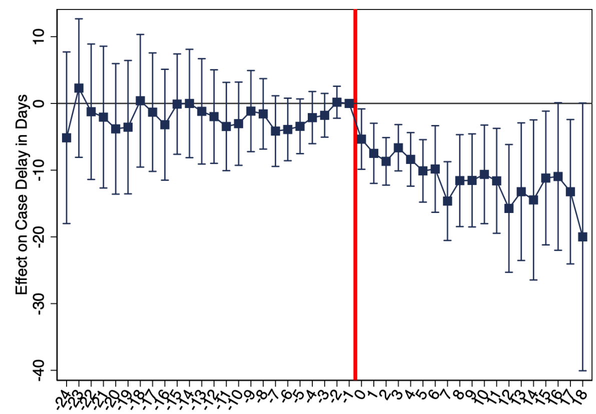
Note: The figure reports event-study estimates of the impact of the E-Court Reform rollout on case duration for non-criminal cases relative to criminal cases in treated public county courts. Data are aggregated to the county court-judge-month-case type level. Event-time indicators are interacted with an indicator for non-criminal cases; criminal cases serve as the control group. Outcomes are regressed on event-time indicators in two specifications: the Sun and Abraham (2021) estimator and a dynamic two-way fixed effects model estimated by OLS. All regressions include judge and county-court × month–year fixed effects, and do not include a separate fixed effect for case type. The omitted category is the month immediately prior to the reform, shown at event month -1. The sample comprises cases filed between April 1, 2009, and February 28, 2020. The Treated Courts sample includes judges observed both before and after the reform and adjudicating both criminal and non-criminal cases. Standard errors are clustered at the county-court level, and 95% confidence intervals are displayed.

Figure 5: E-Court Reform and Case Delay in Days for Treated Courts - Rambachan and Roth (2023)'s Credible Approach to Parallel Trends

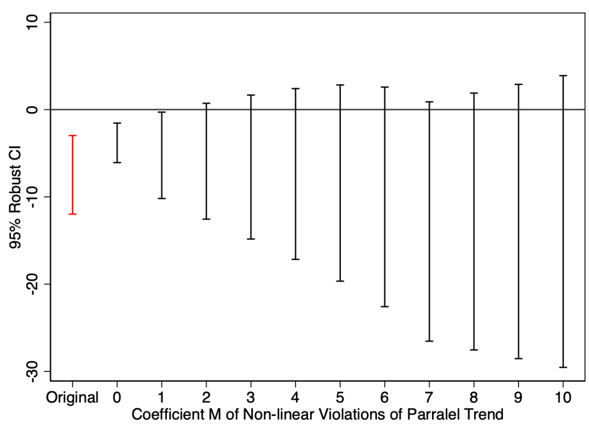
Panel A: Estimated Linear Trend



Panel B: Rotated Event Study



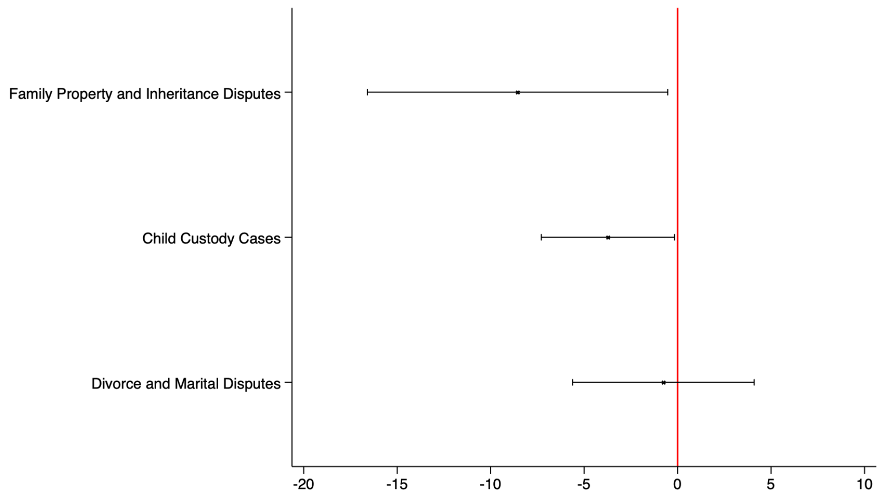
Panel C: Sensitivity to Potential Violations of Parallel Trends



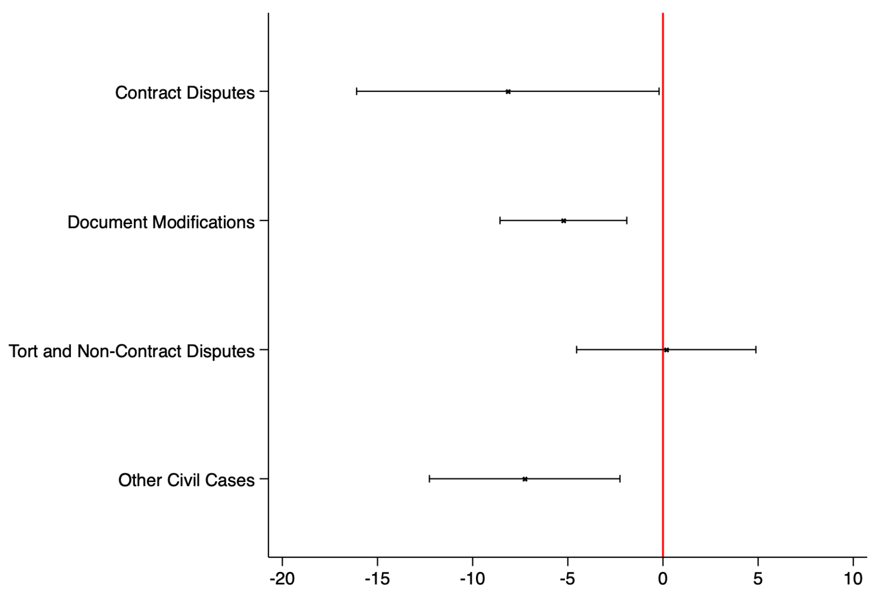
Note: This figure assesses potential violations of the parallel trends assumption in Figure 4, Panel A (treated courts sample). Panel A overlays the event-study coefficients with a linear trend estimated from pre-treatment periods and extrapolated to post-treatment periods. Panel B reports the event-study coefficients corrected for this estimated trend. Panel C shows the sensitivity of these results to linear extrapolation of the pre-treatment coefficients using the honest parallel trends approach of Rambachan and Roth (2023). For Panel C, we report the confidence sets of Rambachan and Roth (2023) for the average post-treatment effect, allowing the slope of the pre-trend coefficients to vary by at most M across consecutive periods.

Figure 6: Heterogeneous Effects on Case Duration – Treated Courts Sample

Panel A: Marriage and Family Cases



Panel B: Document, Contract and Tort Cases



Note: The figure reports heterogeneous effects of the E-Court Reform on case duration in treated courts. Panel A reports estimates for family-related civil cases: family property and inheritance disputes, child custody cases, and divorce or marital disputes. Panel B reports estimates for other civil cases: contract disputes, document modifications, tort or other non-contract disputes, and residual civil cases not included in the family categories or the listed other-civil categories. Each coefficient comes from a separate regression estimated on a subsample containing the listed civil category and all criminal cases, which serve as the comparison group. The sample comprises cases filed between April 1, 2009, and February 28, 2020, in treated first-instance courts. Judges remain in the same court throughout the sample period and are observed before and after the reform with both civil and criminal cases. All specifications include judge fixed effects, case-type fixed effects, county-court-by-month-year fixed effects, and county-court-by-case-type fixed effects. Standard errors are clustered at the county-court level. Horizontal bars denote 95% confidence intervals.

Table 1: Case-Type Composition Before and After E-Court Adoption

	Pre-Reform Share	Post-Reform Change in Share (p.p.)	
10 Most Frequent Non-Criminal Case Types			
% Divorce and Marital Disputes	3.421	0.919	0.892
	(0.080)	(0.841)	(0.721)
% Family Property Cases	0.544	-0.121	-0.078
	(0.032)	(0.123)	(0.107)
% Document Modifications Cases	19.035	-0.798	-0.553
	(0.173)	(1.641)	(1.489)
% Child Custody Cases	11.641	-0.434	-0.227
	(0.141)	(0.884)	(0.779)
% Contract Cases	0.394	0.376	0.307
	(0.028)	(0.252)	(0.191)
% Act of Unlawfulness	2.006	0.445	0.493
	(0.062)	(0.340)	(0.339)
% Land and Housing Cases	0.320	0.116	0.124
	(0.025)	(0.126)	(0.132)
% Breach of Contract	0.559	0.206	0.236
	(0.033)	(0.200)	(0.209)
% Bankruptcy Cases	0.025	0.240	0.173
	(0.007)	(0.251)	(0.185)
% Consumer Rights Cases	0.050	0.067	0.072
	(0.010)	(0.043)	(0.045)
10 Most Frequent Criminal Case Types			
% Narcotics Cases	19.589	0.443	0.393
	(0.175)	(0.862)	(0.790)
% Theft Cases	12.697	1.188*	1.286**
	(0.147)	(0.635)	(0.595)
% Gambling Cases	2.656	0.626**	0.621**
	(0.071)	(0.275)	(0.272)
% Assault Cases	2.596	0.249	0.248
	(0.070)	(0.203)	(0.202)
% Defamation Cases	0.278	0.058	0.072
	(0.023)	(0.089)	(0.086)
% Embezzlement Cases	2.878	0.202	0.275
	(0.074)	(0.239)	(0.230)
% Corruption Cases	1.693	0.032	-0.011
	(0.057)	(0.481)	(0.477)
% Fraud Cases	1.876	0.322	0.307
	(0.060)	(0.229)	(0.228)
% Stolen Property Cases	1.660	0.038	-0.001
	(0.056)	(0.251)	(0.254)
% Forgery Cases	0.369	0.044	0.065
	(0.027)	(0.115)	(0.117)
Non-Criminal FE		Yes	Yes
Judge FE		Yes	Yes
Month-Year FE		Yes	Yes
County Court × Non-Criminal Case FE		No	Yes
Number of Judges		713	713
Number of Courts (Clusters)		52	52
Observations		98810	98810

Note: Each coefficient reports the effect of the post-reform indicator (E-Court Reform) on case-level outcomes, measured as the percentage-point share of cases belonging to the listed type. The first column reports pre-reform shares, with standard errors of the mean in parentheses. Reported Non-Criminal and Criminal case types correspond to the ten most frequent categories in each domain. All columns are estimated on the common case-level sample from the strictest specification. Robust standard errors for the regression columns, clustered at the county-court level, are shown in parentheses. The unit of observation is a case. The sample comprises 98,810 cases filed between April 1, 2009, and February 28, 2020. It is restricted to cases assigned to judges who remained in the same court throughout the sample period and who adjudicated cases both before and after the reform, as well as both criminal and non-criminal cases. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table 2: E-Court Reform and Case Duration in Treated Courts

	Case Duration in Days			
	(1)	(2)	(3)	(4)
E-Court Reform × Non-Criminal Case	-4.161*	-9.099***	-7.506**	-6.947***
	(2.350)	(2.237)	(2.844)	(2.241)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	No	Yes	Yes	Yes
County Court × Month-Year FE	No	No	Yes	Yes
County Court × Non-Criminal Case FE	No	No	No	Yes
Mean Dep Var	47.2	47.2	47.2	47.2
R-squared	0.059	0.254	0.336	0.363
Number of Clusters	52	52	52	52
Observations	98810	98810	98810	98810

Note: The table reports estimates of the effect of the E-Court Reform on case duration, measured as the number of days between filing and first-instance decision. The sample comprises cases filed between April 1, 2009, and February 28, 2020, in the treated-courts sample. Judges remain in the same court throughout the sample period and are observed before and after reform implementation with both civil and criminal cases. The coefficient of interest is the interaction between E-Court reform implementation and the civil-case indicator; criminal cases serve as the within-court comparison group. All columns include the non-criminal case indicator. Column (2) adds judge fixed effects. Column (3) adds county-court-by-month-year fixed effects. Column (4), the preferred saturated specification, additionally includes county-court-by-non-criminal-case fixed effects. Mean Dep Var is calculated over the full estimation sample in each column. Robust standard errors, clustered at the county-court level, are reported in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table 3: E-Court Reform and Decision Quality in Treated Courts

Panel A: Decision Quality and Contestability Measures Based on Available Metadata

	% of Appeals		% of Case Reversals	
	(1)	(2)	(3)	(4)
E-Court Reform × Non-Criminal Case	-1.648***	-1.154**	-0.560***	-0.451**
	(0.393)	(0.451)	(0.149)	(0.197)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	Yes	Yes	Yes	Yes
County Court × Month-Year FE	Yes	Yes	Yes	Yes
County Court × Non-Criminal Case FE	No	Yes	No	Yes
Mean Dep Var	2.676	2.676	0.873	0.873
R-squared	0.117	0.121	0.093	0.094
Number of Clusters	52	52	52	52
Observations	98810	98810	98810	98810

Panel B: Decision Quality Measures Based on Available Full Judgment Texts

	Number of Pages	Reading Ease	Citations per Page	Reasoning per Page
E-Court Reform × Non-Criminal Case	-1.501***	0.348	0.092	0.141
	(0.403)	(0.433)	(0.059)	(0.131)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	Yes	Yes	Yes	Yes
County Court × Month-Year FE	Yes	Yes	Yes	Yes
County Court × Non-Criminal Case FE	Yes	Yes	Yes	Yes
Mean Dep Var	14.65	-14.47	0.55	4.99
R-squared	0.396	0.491	0.342	0.258
Number of Clusters	31	31	31	31
Observations	15599	15599	15599	15599

Note: The table examines whether faster adjudication under the E-Court Reform is accompanied by changes in decision quality in the treated-courts sample. Panel A reports appellate-review outcomes: whether a first-instance decision was appealed and whether it was reversed on appeal. Columns (1) and (3) include non-criminal case fixed effects, judge fixed effects, and county-court-by-month-year fixed effects; columns (2) and (4) additionally include county-court-by-non-criminal-case fixed effects. Panel B uses the subsample of cases with complete written judgments, restricted to judgments whose page length lies within the trimmed 95% range of the page-count distribution. All Panel B columns use the saturated specification with non-criminal case fixed effects, judge fixed effects, county-court-by-month-year fixed effects, and county-court-by-non-criminal-case fixed effects. Reading Ease is the inverted Gunning Fog measure, so higher values indicate more readable text. Citations per Page measures explicit references to statutes, codes, or legal provisions per page. Reasoning per Page measures the combined number of factual-background and legal-reasoning sections per page. The sample comprises cases filed between April 1, 2009, and February 28, 2020. Judges remain in the same court throughout the sample period and are observed before and after reform implementation with both civil and criminal cases. Mean Dep Var is calculated over the full estimation sample in each column. Robust standard errors, clustered at the county-court level, are reported in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table 4: Vertical Spillovers to High Courts

	Case Duration in Days			
E-Court Reform × Non-Criminal Cases	-9.920***	-6.593***	-9.486***	-8.862***
	(0.749)	(0.947)	(1.349)	(1.354)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	No	Yes	Yes	Yes
County Court × Month-Year FE	No	No	Yes	Yes
County Court × Non-Criminal Case FE	No	No	No	Yes
Mean Dep Var	47.17	47.17	47.17	47.17
R-squared	0.259	0.444	0.620	0.632
Number of Clusters	32	32	32	32
Observations	24185	24185	24185	24185

Note: The table uses data from high courts–including the High Court, High Religious Court, High State Administrative Court, and the Supreme Court–and comprises cases filed in treated and untreated high-court jurisdictions. Treated high courts are defined as those where the corresponding high court oversees at least one treated first-instance court. The sample is restricted to cases assigned to judges who remained in the same court throughout the sample period. High courts are included when they are observed both before and after the reform and adjudicate both criminal and non-criminal cases during the sample period. The dependent variable is case duration in days, right-trimmed at the 95% confidence interval. Of 66 high courts, 19 were treated in the first wave of the E-Court Reform and 36 in the second wave, including the initial 19. The Supreme Court is included in all specifications. Robust standard errors, clustered at the county court level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Impact of E-Court Reform on Judicial Bias

Panel A: Gender Bias in Review	% of Case Reversals			
E-Court Reform × Judge is Female	-0.204	-0.395***	-0.493***	-0.496***
	(0.158)	(0.142)	(0.141)	(0.142)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	No	Yes	Yes	Yes
County Court × Month-Year FE	No	No	Yes	Yes
County Court × Non-Criminal Case FE	No	No	No	Yes
Mean Dep Var	0.87	0.87	0.87	0.87
R-squared	0.003	0.036	0.093	0.094
Number of Clusters	52	52	52	52
Observations	98810	98810	98810	98810

Panel B: Religion Bias in Review	% of Case Reversals			
E-Court Reform × Judge is Non-Muslim	0.122	0.015	0.076	0.082
	(0.194)	(0.213)	(0.208)	(0.211)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	No	Yes	Yes	Yes
County Court × Month-Year FE	No	No	Yes	Yes
County Court × Non-Criminal Case FE	No	No	No	Yes
Mean Dep Var	0.87	0.87	0.87	0.87
R-squared	0.003	0.036	0.092	0.094
Number of Clusters	52	52	52	52
Observations	98810	98810	98810	98810

Note: This table examines the impact of the E-Court Reform on judicial bias in review at appellate courts, measured by the probability that lower court decisions are reversed in the high court. The sample comprises cases filed between April 1, 2009, and February 28, 2020. It is restricted to cases assigned to judges who remained in the same court throughout the sample period and who adjudicated cases both before and after the reform, as well as both criminal and non-criminal cases. The dependent variable equals 100 if the case was reversed in the high court and 0 otherwise. Panel A reports results for gender bias, where the key explanatory variable is the interaction between an indicator for post-reform cases in reformed courts and an indicator for whether the lower court judge is female. Judge gender is classified using a bidirectional LSTM trained on Indonesian names (F1-score = 95%). Panel B reports results for religion bias, where the key explanatory variable is the interaction between an indicator for post-reform cases in reformed courts and an indicator for whether the lower court judge is non-Muslim. Judge religion is classified using a bidirectional LSTM trained on Indonesian names (F1-score = 90%). Mean Dep Var is calculated over the full estimation sample in each column. Robust standard errors, clustered at the court level, are reported in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

For Online Publication

Technology and the Production of Justice: System-Wide Effects of Indonesia's E-Court Reform

Sultan Mehmood, Resuf Ahmed and Alexandra Filimonova

Contents

Appendix A: Variable Definitions and Descriptions

Appendix A1. Variable Definitions

Decision Time in Days. Number of days between case initiation and final verdict.

E-Court Reform. Indicator equal to 1 if the case was processed in a month of the e-court reform's implementation in the relevant court, and 0 otherwise.

Non-Criminal Case. Indicator equal to 1 if the case is classified as civil, special civil, state administrative, or religious. Classification follows the official court database.

Level of Process. Categorical variable indicating whether a case was processed at the first-instance or appellate level, based on the official classification.

Region. Province in which the court is located. There are 38 provinces in total.

% of Appeals. Share of first-instance cases that are appealed. Constructed as a binary indicator (1 if appealed, 0 otherwise), multiplied by 100 for presentation.

% of Case Reversals. Share of first-instance cases for which the appeal court reversed the original verdict. Defined analogously to % of Appeals and scaled by 100. Classification is based on GPT-assisted reading of appeal verdicts, using extracted case numbers and reversal status. Appendix C provides further details on the prompt used.

Type of Court. Categorical variable indicating the court type: county, high court, high religious court, high state administrative court, military court, religious court, or state administrative court.

Number of Pages. Total number of pages in the full-text judgment.

Law Citations per Page. Total number of legal references in the judgment divided by the total number of pages.

Facts and Legal Arguments per Page. Number of factual statements and legal arguments in the *Legal Considerations* section, normalized by the number of pages in the judgment.

First Wave. Indicator equal to 1 if the case was processed in a court and month in which the e-court reform had been implemented during the first rollout period (July 2018 to August 2019, inclusive), and 0 otherwise.

Second Wave. Indicator equal to 1 if the case was processed in a court and month in which the reform had been implemented during the second rollout period (August 2019 onward), and 0 otherwise.

Appendix A2. Data Construction

The dataset is organized at the **case level**. For each case, we observe:

- **Dates:** registration and decision;
- **Case characteristics:** case type, decision time, appeal and reversal (which we classified using GPT), text measures;
- **Court information:** judge, court, province, and region.

We combine several datasets:

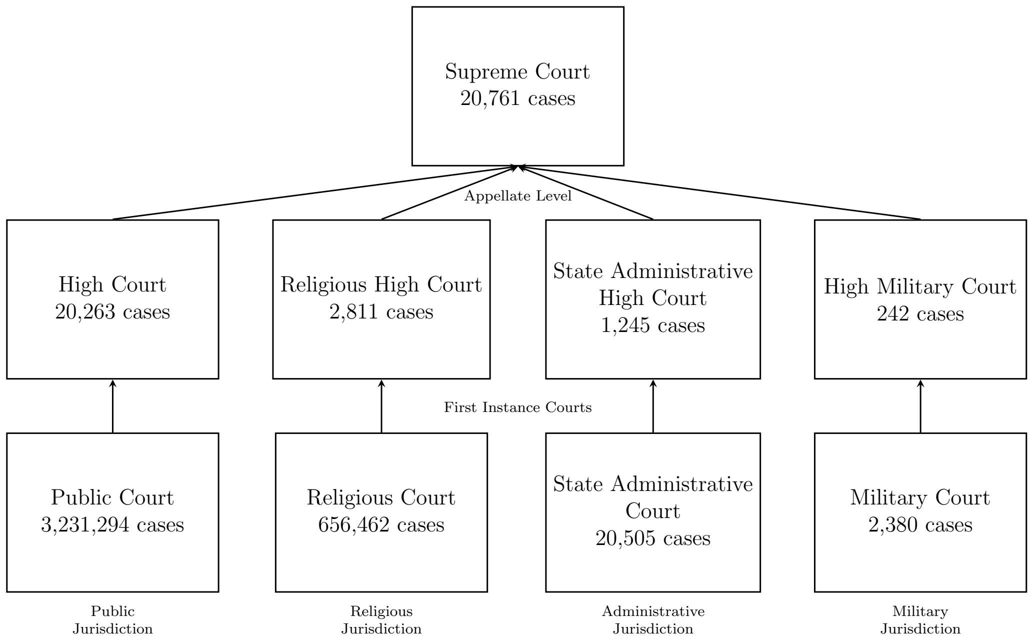
1. **First-instance cases**
    - *Treated courts sample:* 52 courts, 98,810 observations. Includes cases from courts ever treated by the reform, restricted to judges who never move across courts and who adjudicate cases both before and after the reform, and both criminal and non-criminal cases.
    - *Full sample:* 786 courts, 3,191,483 observations. Includes all courts in the dataset, restricted to judges who never move across courts.
2. **Subsample of first-instance cases with textual metrics**

    Contains readability scores, number of pages, statutory citations, and facts per page. The sample is defined as in (1), but restricted to cases for which full judgment texts were successfully parsed.
    - *Treated courts sample:* 15,599 observations.
    - *Full sample:* 92,586 observations.
3. **High-Court cases**

    A high court is considered treated if it is linked to a treated first-instance court.
    - *Treated courts sample:* 32 courts, 24,185 observations. Includes cases from high courts linked to treated first-instance courts, restricted to stationary judges. A high court is considered treated once at least one corresponding first-instance court is treated.
    - *Full sample:* 62 courts, 87,710 observations. Includes all high courts, restricted to judges who never move across courts.

Figure A1: Institutional Background

Panel A: Indonesia's Court Structure

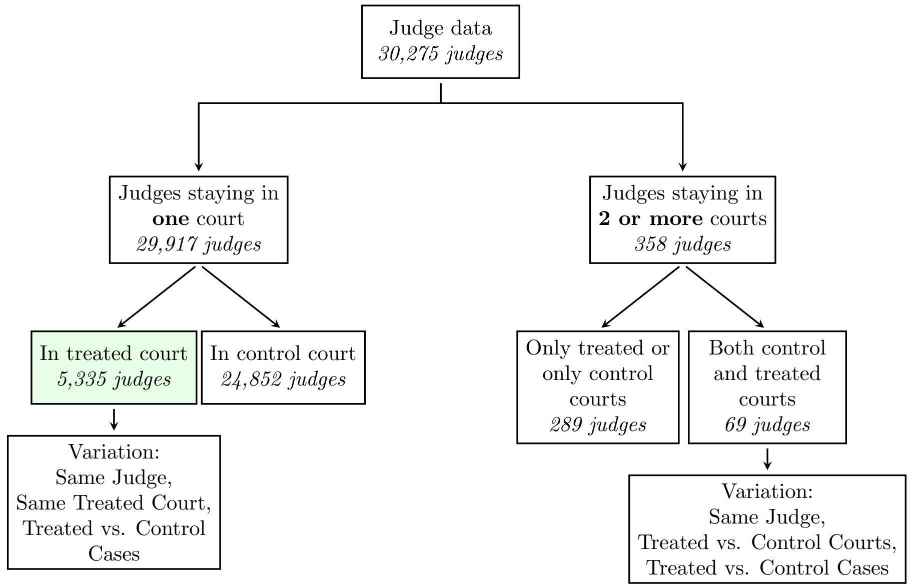


Panel B: E-Court Reform Transition from Paper-Based to Digital Procedures

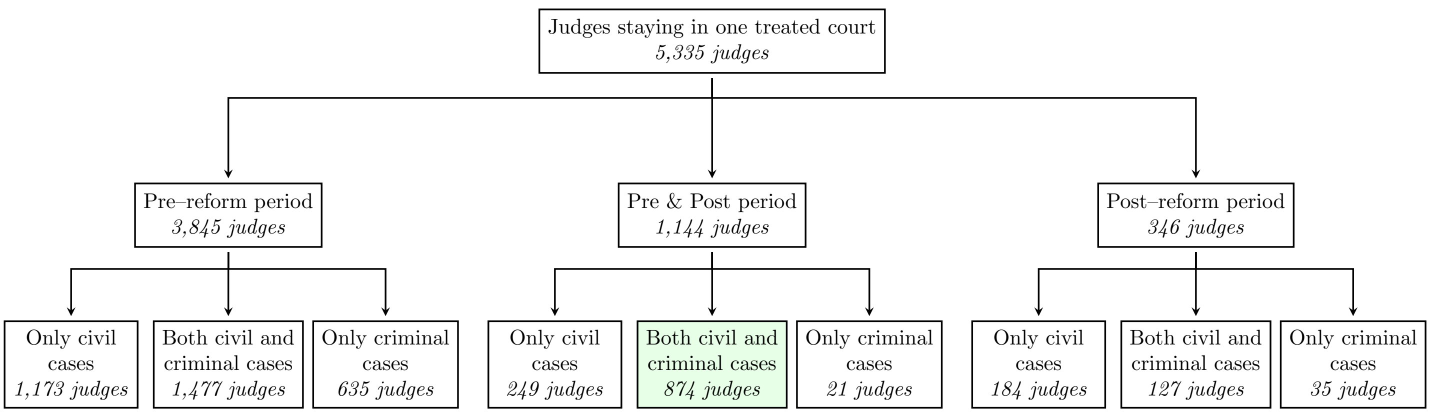
Notes: Panel A illustrates the hierarchy of the Indonesian court system, showing the four main judicial branches: public, religious, administrative and military. The lowest-tier corresponds to municipal courts. They can be subdivided into public or secular courts (Peradilan Umum) that handle civil and criminal disputes; religious courts (Peradilan Agama) that also handle civil and criminal disputes; administrative courts (Peradilan Tata Usaha Negara) that handle disputes with state agencies; and specialized military courts (Pengadilan Militer). The Supreme Court (Mahkamah Agung) serves as the final court of appeal for all judicial streams except constitutional matters. We indicate caseloads for each court in 2020. Panel B illustrates the transition from traditional paper-based and in-person court procedures to digital judicial services introduced under the E-Court reform. Each panel displays the procedural workflow before and after the introduction of electronic systems for filing, litigation, summons, and payments. This figure has been designed using resources Flaticon.com.

Figure A2: E-Court Reform and Sources of Variation

Panel A: Judges in Treated Courts

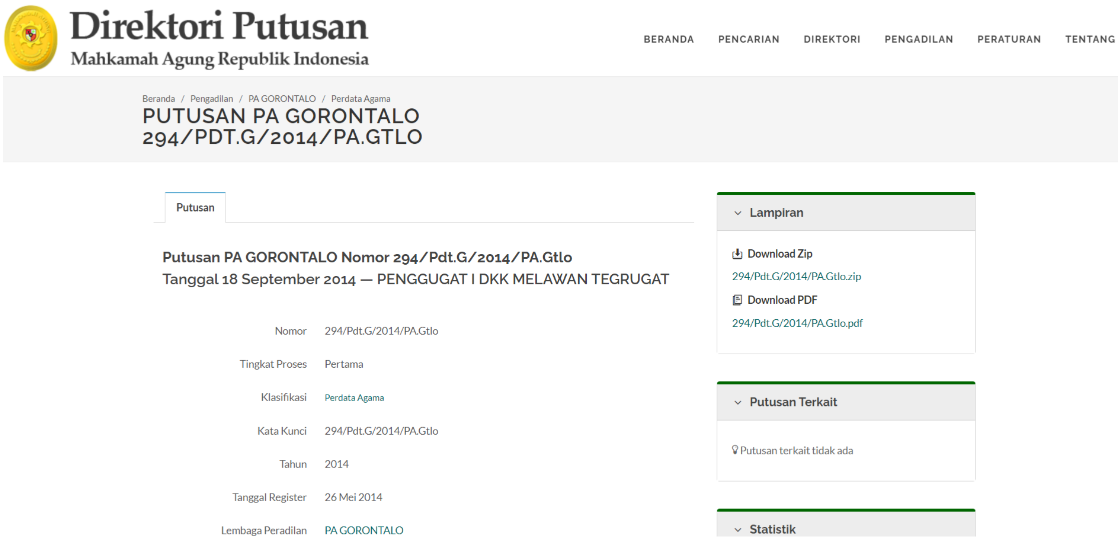


Panel B: Judges Remaining in One Treated Court Over the Sample - Case Type by Time Variation



Note: **Panel A** uses the judge as the unit of observation and summarizes how many judges served in one, or two or more courts over the entire sample period. Each box reports the total number of unique judges in that category. The third row further splits these observations by the treatment status of the courts in which judges served. "Treated" refers to courts ever exposed to the E–Court reform, while "Control" refers to courts never exposed. For judges who served in multiple courts, the classification depends on whether they served exclusively in treated courts or exclusively in control courts, or in a mix of both.
Panel B uses the judge as the unit of analysis and reports how many judges who, over the entire sample, remained in a single treated court and the types of cases they adjudicated across periods. Throughout, "non-criminal" cases are treated cases, while "criminal" cases are control cases. All figures are computed from the sample of all cases filed between April 1, 2009, and February 28, 2020. The final sample consists of 721 judges, because of singleton observations.

Figure A3: Example of Data from the Website



Note: The figure shows a typical case page from the official website of the Supreme Court of Indonesia (https://putusan3.mahkamahagung.go.id/). Each page contains structured metadata at the case level, including the registration date, decision date, case classification, and court identifiers. We utilized this information to construct our dataset and compute measures such as case duration.

Figure A4: Example of a File Containing the Full Text of a Judgment

Note: The screenshot illustrates a section of a full-text judicial decision available on the official website of the Supreme Court of Indonesia (https://putusan3.mahkamahagung.go.id/). We scraped a subset of such judgments, converted them from PDF format into machine-readable text, and analyzed the texts to compute various textual metrics.

Table A1: Types of Cases and Their Descriptions

Type of Case	# of Obs.	Description
Panel A: Non-Criminal Cases (Broadly Defined as Civil)		
Public Civil Case	34,051	Cases in the public (general) civil courts, primarily involving divorce and marital disputes, along with unlawful acts, petitions, contract breaches, land and property conflicts, inheritance matters, and other civil claims.
Religious Civil Case	1,235	Cases under religious (primarily Islamic) family law, dominated by divorce proceedings, along with marriage validation, guardianship, marriage dispensation, inheritance, joint property, polygamy permits, gifts, wills, and other matters governed by religious law.
State Administrative Case	89	Cases concerning disputes with government agencies, including land and property conflicts, civil service and employment matters, licensing and permits, elections, political party issues, taxation, and other administrative decisions.
Special Civil Case	3,222	Specialized civil cases, primarily industrial relations disputes, along with bankruptcy and insolvency, consumer protection, intellectual property rights (copyright, trademarks, patents, industrial design), arbitration, competition law, and political party disputes.
Panel B: Criminal Cases		
General Criminal Case	36,628	General criminal cases, dominated by theft, gambling, assault, defamation, embezzlement, crimes against state security, and fraud, along with murder, traffic offenses, extortion, forgery, and other common criminal offenses.
Military Criminal Case	1	Criminal cases under military jurisdiction, primarily desertion and offenses against decency, along with subordination violations, and occasional cases of corruption, narcotics, weapons, and general crimes committed by military personnel.
Special Criminal Case	29,223	Special criminal cases, dominated by narcotics and psychotropics offenses, along with firearms, corruption, juvenile cases, environmental crimes, and other specific categories such as domestic violence, IT crimes, money laundering, human trafficking, pornography, terrorism, and taxation.

Note: The classification of cases into types and the descriptions of each type are based on metadata provided by the Supreme Court of Indonesia. The descriptions summarize the subcategories of cases listed in the official metadata. The sample includes cases filed between April 1, 2009, and February 28, 2020, in treated courts (ever exposed to the reform) and is restricted to cases assigned to judges who remained in the same court throughout the sample period, were observed both before and after the reform, and adjudicated both criminal and non-criminal cases.

Appendix A3. Additional Institutional Background to Indonesian Court System

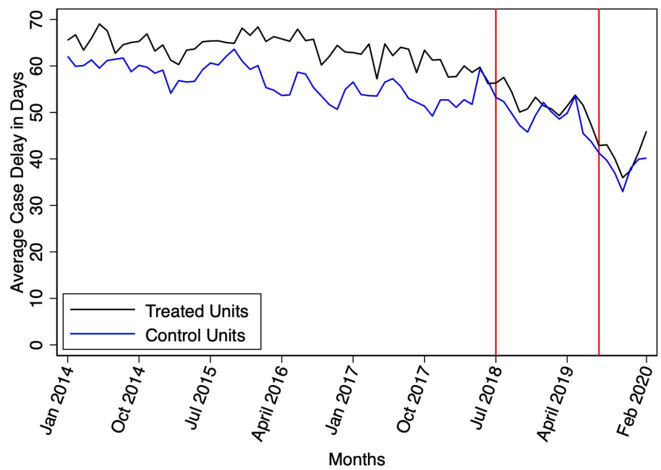
Legal Foundations. Indonesia, an archipelago with more than 17,000 islands, has a history shaped by colonial rule under the Dutch East Indies Company and the Dutch government. After gaining independence in 1945, the country established a unitary system with a centralized government. Indonesia's legal system is a legacy of its colonial past, blending Dutch civil law with elements of customary (adat) law and Islamic jurisprudence. This hybrid legal framework reflects the country's cultural and religious diversity, which has been continuously woven into the fabric of its legal and political institutions (Butt and Lindsey, 2018; Tim and Pausacker, 2020). As is common in many former colonies, Indonesia's legal infrastructure remains grounded in colonial civil law traditions, which are shown to be correlated with long-term legal and economic outcomes across nations (Porta et al., 2008).

Administrative Divisions. Indonesia's administrative framework mirrors its vast geographical and demographic diversity. The country, the fourth most populous in the world, is organized into 38 provinces (*provinsi*), ranging from urban centers like Jakarta to remote areas such as Papua. These provinces are further subdivided into counties (*kabupaten*) and municipalities or districts (*kota*). For simplicity, we refer to these second-tier units collectively as municipalities. As of the most recent census, there are approximately 500 municipalities, each governed by local municipal governments responsible for public services and economic development.

Procedural Practice. Within this institutional and administrative structure, the courts handle a broad mix of disputes, ranging from routine civil matters (such as small-scale contract enforcement, family property disputes, and documentation changes) to more complex commercial and criminal cases. Historically, case processing has relied heavily on paper-based filing, in-person submissions, and manual coordination across court offices, clerks, and judges. This procedural model generated substantial scope for delay, particularly in courts serving large urban populations or remote regions with limited administrative capacity. At the same time, the bulk of the civil docket consists of relatively standardised disputes that follow well-defined procedural templates, creating a natural margin along which improvements in filing, scheduling, and document management can translate into sizeable efficiency gains. The E-Court reform is best understood against this backdrop of heterogeneous caseloads and high fixed administrative costs in moving cases through each procedural stage.

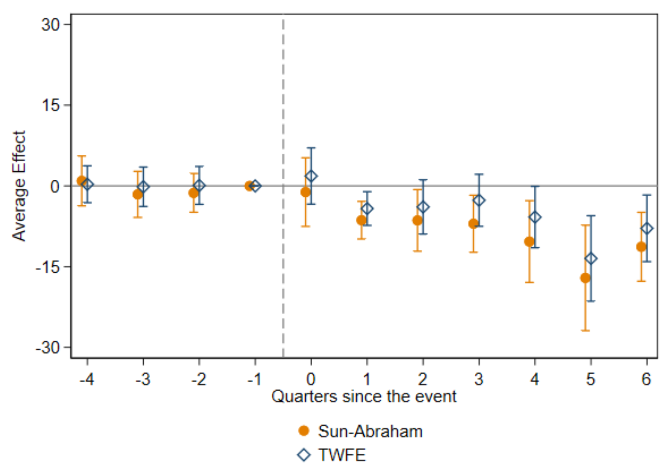
Appendix B: Additional Figures and Tables

Figure B1: Average Case Duration in Days over Time – Treated vs Control Units – Full Sample



Note: The figure plots average case duration (in days) for treated units (non-criminal cases in treated courts) and control units (non-criminal cases in control courts and all criminal cases). The vertical red lines indicate the start of the reform: the first wave in July 2018, and the second wave in August 2019.

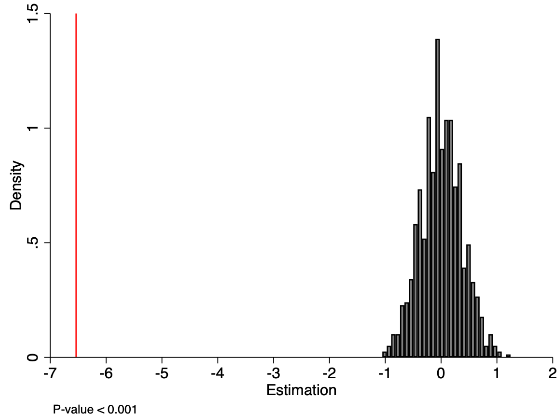
Figure B2: Impact of E-Court Reform on Case Duration - Triple Difference



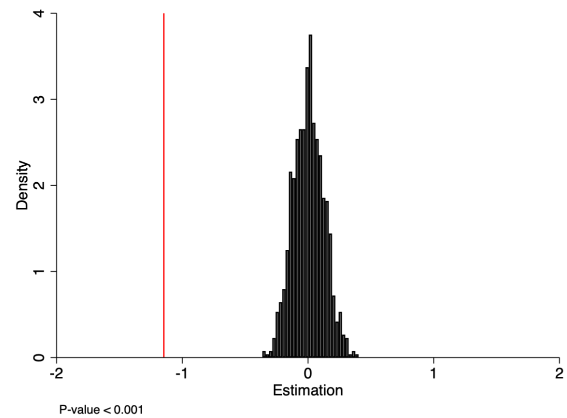
Note: This figure presents event study estimates of the impact of the E-Court reform on case duration (in days). The estimation is based on a triple-differences specification with county court-by-quarter, quarter-by-case-type (civil and criminal cases), county court-by-case-type, and judge fixed effects. The outcome is aggregated at the court-by-judge-by-case-type-by-quarter level. Estimates from Sun and Abraham (2021) are shown as orange circles, and standard two-way fixed effects (TWFE) estimates are shown as blue diamonds. Vertical bars represent 90% confidence intervals. Standard errors are clustered at the county court level. The vertical dashed line at zero indicates the month of reform implementation.

Figure B3: Permutation Inference Test – Treated Courts

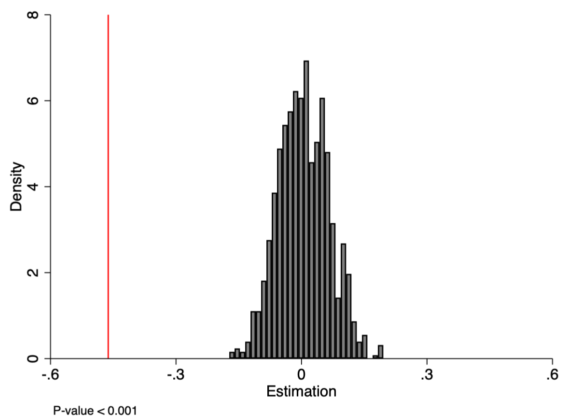
Panel A: Case Duration



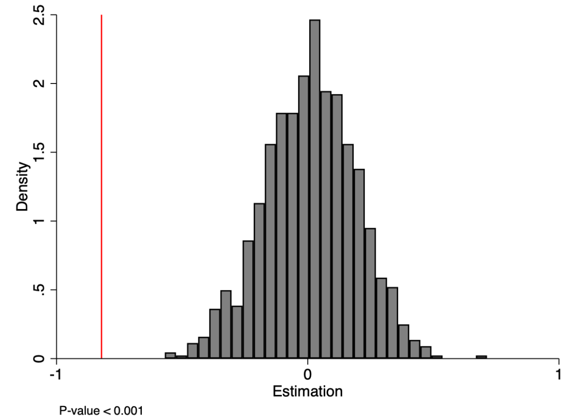
Panel B: % of Appeals



Panel C: % of Reversals

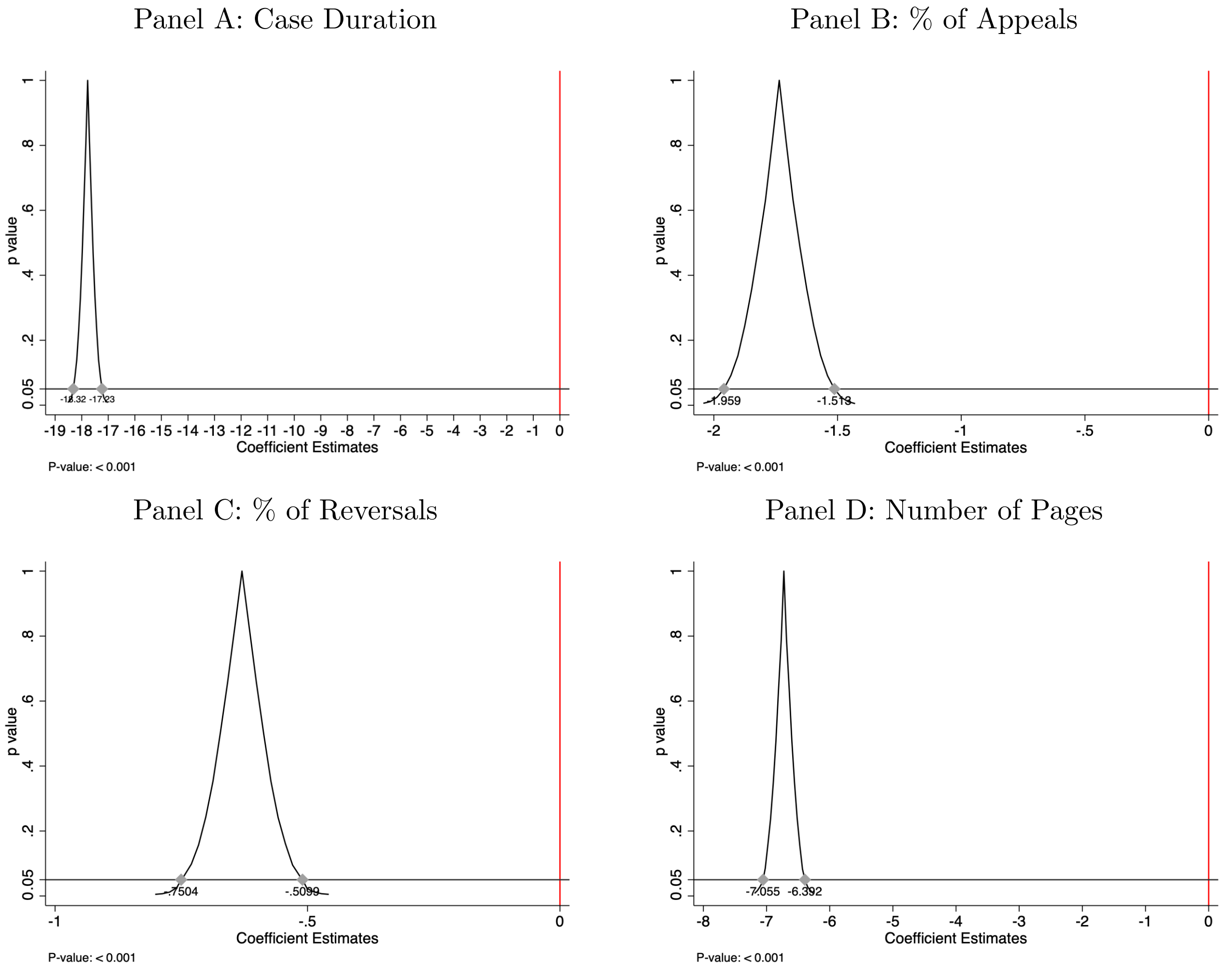


Panel D: Number of Pages



Note: This figure presents the results of a permutation inference test based on 1000 permutations. Each panel shows the estimated probability density function of the treatment effect under the null hypothesis that the reform has no impact on case duration, the share of appeals, the share of reversals and the number of pages in the judgment. All panels use the sample of cases from treated courts assigned to judges who remained in the same court throughout the sample period and who adjudicated cases both before and after the reform, as well as both criminal and non-criminal cases. All regressions include judge, non-criminal, court-by-month, court-by-case type (non-criminal vs. criminal) fixed effects. The vertical line in each panel marks the actual estimated coefficient.

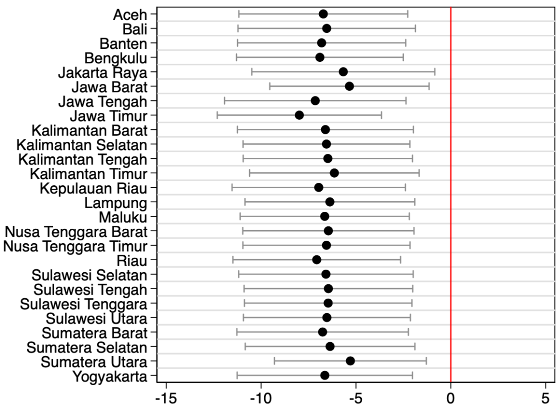
Figure B4: Confidence Intervals by Wild Bootstrap – Treated Courts



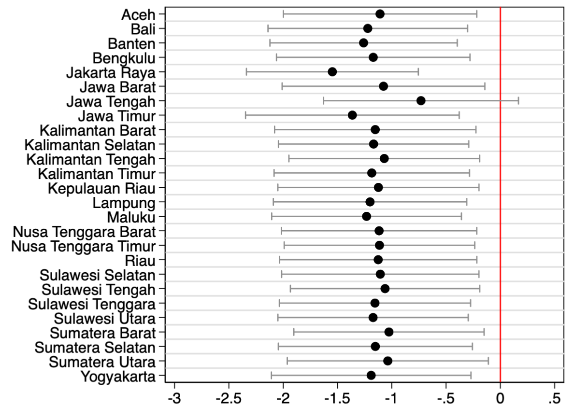
Note: This figure reports estimates obtained using a wild cluster bootstrap procedure of MacKinnon and Webb (2018). All panels use the sample of cases from treated courts assigned to judges who remained in the same court throughout the sample period and who adjudicated cases both before and after the reform, as well as both criminal and non-criminal cases. Each panel presents the histogram of the bootstrapped coefficient distribution for the interaction between the reform and non-criminal cases for case duration, the share of appeals, the share of reversals and the number of pages in the judgment. All regressions include judge fixed effects. Vertical lines indicate zero, and horizontal reference lines indicate the confidence interval boundaries.

Figure B5: Robustness to Dropping Region at a Time – Treated Courts

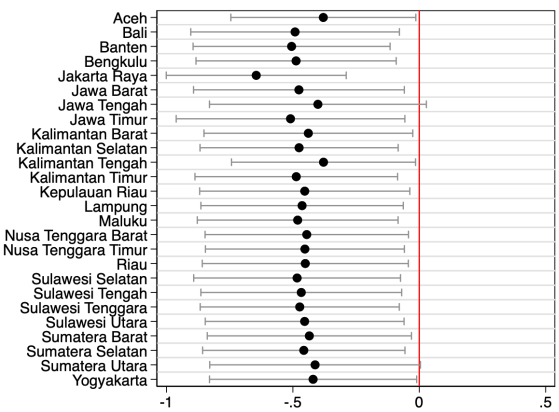
Panel A: Case Duration



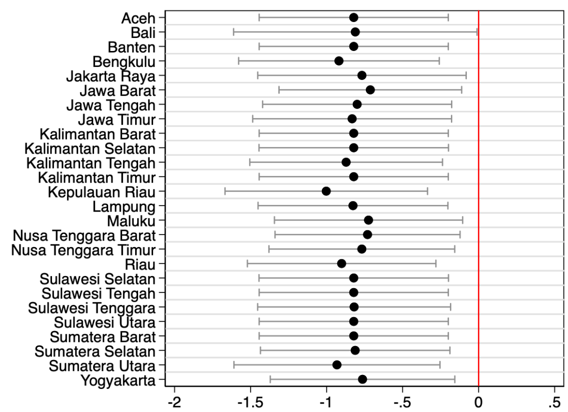
Panel B: % of Appeals



Panel C: % of Reversals



Panel D: Number of Pages



Note: This figure plots estimated coefficients from the baseline specification while sequentially excluding one region at a time. All panels use the sample of cases from treated courts assigned to judges who remained in the same court throughout the sample period and who adjudicated cases both before and after the reform, as well as both criminal and non-criminal cases. Each panel shows how the estimated treatment effect varies across exclusions for case duration, the share of appeals, the share of reversals and the number of pages in the judgment. All regressions include judge, court-by-month-year, case-type (non-criminal vs criminal) and court-by-type fixed effects. Standard errors are clustered at the county-court level. Vertical reference lines indicate zero, and horizontal lines represent confidence intervals. 95% confidence intervals are shown.

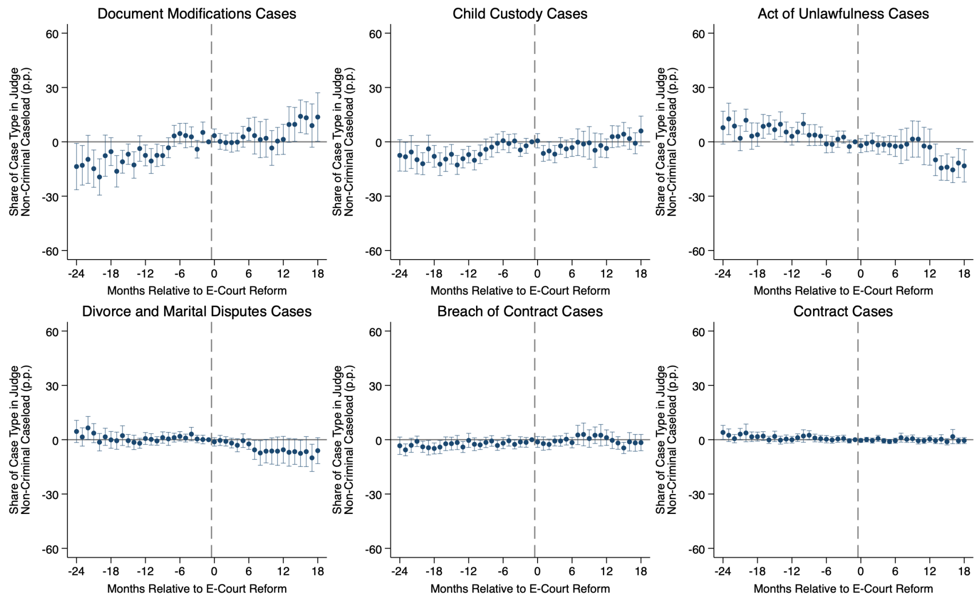
Table B1: Distribution of Courts and Reform Timing across Provinces

Province	# of Treated Courts	# of Untreated Courts	# of Courts	Treatment Date
Aceh	1	45	46	August, 2019
Bali	3	13	16	July, 2018
Bangka Belitung	0	10	10	-
Banten	2	7	9	July, 2018
Bengkulu	1	16	17	August, 2019
Gorontalo	0	11	11	-
Jakarta Raya	11	2	13	July, 2018
Jambi	0	20	20	-
Jawa Barat	7	40	47	July, 2018
Jawa Tengah	7	63	70	July, 2018
Jawa Timur	6	66	72	July, 2018
Kalimantan Barat	1	21	22	August, 2019
Kalimantan Selatan	1	25	26	August, 2019
Kalimantan Tengah	1	22	23	August, 2019
Kalimantan Timur	2	18	20	August, 2019
Kalimantan Utara	0	6	6	-
Kepulauan Riau	3	8	11	July, 2018
Lampung	2	23	25	July, 2018
Maluku	1	15	16	August, 2019
Maluku Utara	0	9	9	-
Nusa Tenggara Barat	1	15	16	August, 2019
Nusa Tenggara Timur	1	29	30	August, 2019
Papua	1	16	17	August, 2019
Papua Barat	0	8	8	-
Riau	2	21	23	August, 2019
Sulawesi Barat	0	8	8	-
Sulawesi Selatan	4	40	44	July, 2018
Sulawesi Tengah	1	17	18	August, 2019
Sulawesi Tenggara	1	18	19	August, 2019
Sulawesi Utara	1	16	17	August, 2019
Sumatera Barat	1	34	35	August, 2019
Sumatera Selatan	1	21	22	July, 2018
Sumatera Utara	3	42	45	July, 2018
Yogyakarta	2	9	11	August, 2019
Total	68	734	802	-

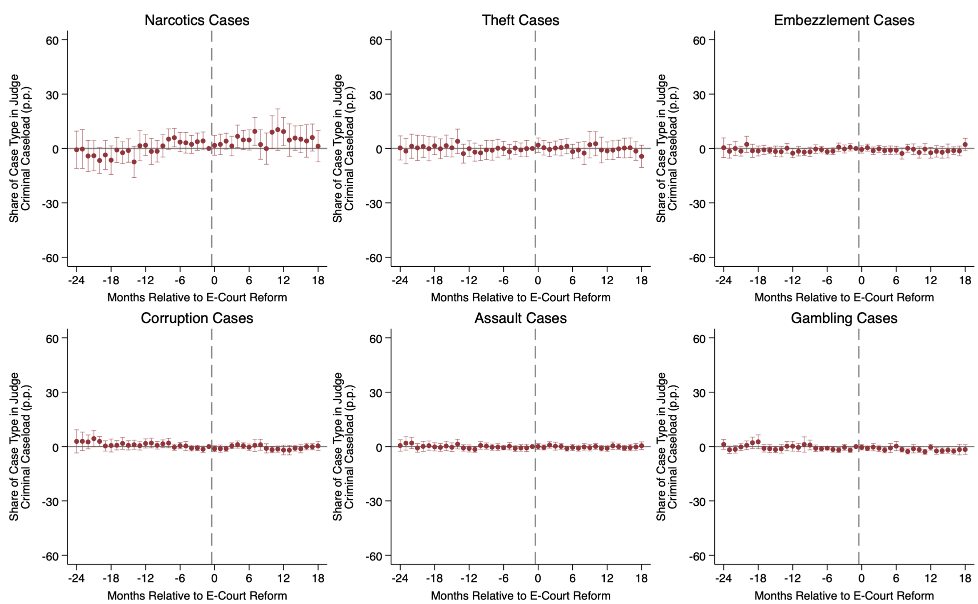
Note: The table is based on our dataset and reports the distribution of courts and reform dates across provinces. The dataset contains 34 provinces with 802 district courts (public, religious, state administrative, and military). Sixteen provinces adopted the E-Court reform in the first wave (July 2018), eleven in the second wave (August 2019), while seven provinces contain never treated courts.

Figure B6: E-Court Reform and Share of Case Types

Panel A: Share of Non-Criminal Case Types Over Time

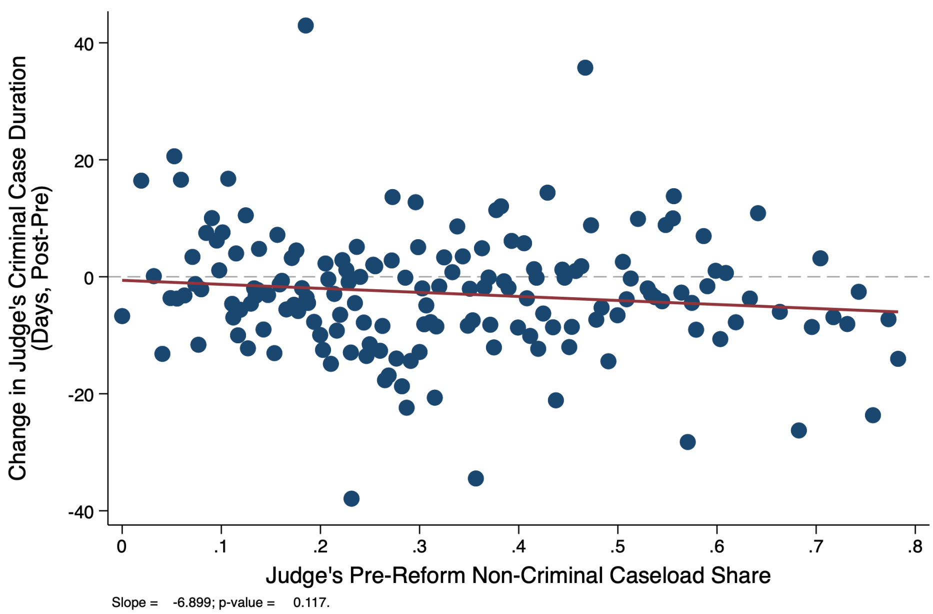


Panel B: Share of Most Frequent Criminal Case Types Over Time



Note: The figure reports event-study estimates of the E-Court Reform and the share of case types over time in treated first-instance courts. Panel A reports the share of non-criminal case types over time, measured as the percentage of a judge's non-criminal monthly caseload belonging to each indicated case type. Panel B reports the share of the most frequent criminal case types over time, measured as the percentage of a judge's criminal monthly caseload belonging to each indicated case type. The six case types in each panel are selected by their pre-reform mean share of total judge-month caseload. Standard errors are clustered at the county-court level, and 95% confidence intervals are shown. The vertical dashed line marks E-Court reform implementation. The sample comprises judge-month observations from cases filed between April 1, 2009, and February 28, 2020, assigned to judges who remained in the same court and adjudicated both criminal and non-criminal cases before and after the E-Court reform.

Figure B7: Criminal Case Duration Changes by Judges' Pre-Reform Non-Criminal Caseload Share



Note: The figure plots changes in criminal-case duration across bins of judges' pre-reform non-criminal caseload share. The horizontal axis is the share of a judge's pre-reform caseload that was non-criminal. Judges are grouped into 200 bins by this pre-reform share. The vertical axis reports the equal-judge-weight average of each judge's post-minus-pre change in mean criminal-case duration within each bin. The fitted line is estimated without bin-size weights. The sample comprises treated first-instance court cases filed between April 1, 2009, and February 28, 2020, assigned to judges who remained in the same court and handled both criminal and non-criminal cases before and after reform implementation.

Table B2: Descriptive Characteristics of Judges, Courts and Cases (Before vs. After Reform)

	Before Reform	After Reform
Panel A: Judge Characteristics		
Number of Judges	8,317	8,680
Gender (Female)	0.266	0.280
Religion (Muslim)	0.833	0.843
Years of experience	2.938	3.451
Caseload per month	9.102	13.938
Panel B: Court Characteristics		
Average judges per month in court	5.509	5.751
Average cases per month in court	50.096	80.102
Year of E-court rollout	-	2018-2019

Note: This table reports summary statistics for the full sample of cases filed between January 2, 2000, and February 28, 2020. The sample is divided into observations from the period before and after the reform implementation. Panel A describes characteristics of the judges, and Panel B describes court-level characteristics. *Number of Judges* is measured as the total count within the 20-month windows immediately preceding and following the reform. All other variables are calculated over the full, respective pre- and post-reform periods. Detailed variable descriptions are provided in Appendix A.

Table B3: Summary Statistics of Main Variables

	Mean	**Median**	**SD**	**Min**	**Max**	**N obs**
Panel A: First-Instance Court Variables – Full Sample						
Case Duration in Days	53.23	38	39.60	0	153	3.15 mln
E-Court Reform	0.015	0	0.122	0	1	3.15 mln
Non-Criminal Cases	0.895	1	0.307	0	1	3.15 mln
% of Appeals	0.455	0	6.727	0	100	3.15 mln
% of Case Reversals	0.168	0	4.089	0	100	3.15 mln
First Wave	0.005	0	0.071	0	1	3.15 mln
Second Wave	0.009	0	0.096	0	1	3.15 mln
Panel B: First-Instance Court Variables – Treated Courts Sample						
Case Duration in Days	47.25	43	35.13	0	153	98,810
E-Court Reform	0.479	0	0.500	0	1	98,810
Non-Criminal Cases	0.367	0	0.482	0	1	98,810
% of Appeals	2.676	0	16.138	0	100	98,810
% of Case Reversals	0.873	0	9.305	0	100	98,810
% Conditional Reversals	32.640	0	46.898	0	100	2,644
First Wave	0.160	0	0.367	0	1	98,810
Second Wave	0.299	0	0.458	0	1	98,810
Panel C: High Court Variables – Full Sample						
Case Duration in Days	43.95	41	22.30	0	108	87,710
E-Court Reform	0.256	0	0.437	0	1	87,710
Non-Criminal Cases	0.617	1	0.486	0	1	87,710
First Wave	0.650	1	0.477	0	1	87,710
Second Wave	0.968	1	0.176	0	1	87,710
Panel D: High Court Variables – Treated Courts Sample						
Case Duration in Days	47.17	43	24.79	0	137	24,185
E-Court Reform	0.305	0	0.460	0	1	24,185
Non-Criminal Cases	0.634	1	0.482	0	1	24,185
First Wave	0.642	1	0.479	0	1	24,185
Second Wave	0.962	1	0.192	0	1	24,185

Note: This table reports mean, median, standard deviation for main variables in our analysis. The unit of analysis is a case. For description of variables see Appendix A. All samples include cases filed between April 1, 2009, and February 28, 2020. First-instance samples are restricted to cases assigned to judges who did not transfer between courts during the sample period. The high-court treated-courts sample follows the estimation sample used in the vertical-spillovers table. Treated courts are courts that were ever affected by the reform. Conditional reversals are calculated only among appealed cases, so this row has a smaller sample.

Table B4: Illustrative Cost Benchmarks for the E-Court Reform

Panel A: Estimated Implementation Expenditure			
Fiscal Year	Estimated Share of Non-Personnel Budget Allocated to Reform	Additional Allocations	Total Expenditure
2018	118,534,248 $\times$ 25% = 29,633,562	–	29,633,562
2019	118,800,000 $\times$ 16% = 19,800,000	17,580,000	37,380,000
2020	166,523,484 $\times$ 16% $\times$ 17% = 4,625,652	–	4,625,652
Total Implementation Expenditure		71,639,214	
Panel B: Illustrative Wage-Equivalent Value of Reduced Case Duration			
Participant Type	Hourly Wage	Implied Value per Case	Implied Value for 100,000 Cases
Litigant (Party)	0.45	$2 \times 0.45 \times 8 \times 7 = 50$	$50 \times 100{,}000 =$ 5,000,000
Judge	12	$3 \times 12 \times 8 \times 7 = 2{,}016$	$2{,}016 \times 100{,}000 =$ 201,600,000
Lawyer	16.5	$2 \times 16.5 \times 8 \times 7 = 1{,}848$	$1{,}848 \times 100{,}000 =$ 184,800,000
Court Clerk	5	$2 \times 5 \times 8 \times 7 = 560$	$560 \times 100{,}000 =$ 56,000,000
Total Implied Wage-Equivalent Value		4,474 per case	447,400,000
Panel C: Comparison with Implementation Expenditure			
Difference Relative to Implementation Expenditure		$447{,}400{,}000 - 71{,}639{,}214 = 375{,}760{,}786$	
Implied Value-to-Expenditure Ratio		$447{,}400{,}000/71{,}639{,}214 = 6.25$	
Illustrative 10-Year Value-to-Expenditure Ratio		$9.084 \times 10^9/(23{,}879{,}738 \times 10) = 38.04$	

continued on next page

Table B4: Illustrative Cost Benchmarks for the E-Court Reform (continued)

Panel D: Four-Hour Comparison with Implementation Expenditure	
Difference Relative to Implementation Expenditure	$223{,}700{,}000 - 71{,}639{,}214 = 152{,}060{,}786$
Implied Value-to-Expenditure Ratio	$223{,}700{,}000/71{,}639{,}214 = 3.12$
Illustrative 10-Year Value-to-Expenditure Ratio	$4.542 \times 10^9/(23{,}879{,}738 \times 10) = 19.02$
Panel E: Five-Day, Eight-Hour Comparison with Implementation Expenditure	
Difference Relative to Implementation Expenditure	$319{,}600{,}000 - 71{,}639{,}214 = 247{,}960{,}786$
Implied Value-to-Expenditure Ratio	$319{,}600{,}000/71{,}639{,}214 = 4.46$
Illustrative 10-Year Value-to-Expenditure Ratio	$6.489 \times 10^9/(23{,}879{,}738 \times 10) = 27.17$

Note: All values are expressed in U.S. dollars. The exchange rate applied is 1 IDR = 0.00006 USD as of October 8, 2025. Panel A reports the estimated costs of the reform. The estimated funds are calculated as 25% in 2018 and 16% in 2019 and 2020 of the annual non-personnel budget of Supreme Court of the Republic of Indonesia (2020), since the e-Court reform constituted one of the 4 main expenditure categories of the Court in 2018 and one of the six main expenditures in 2019 and 2020. For 2020, we use 17% of the annual budget because our data cover only January and February of that year (2 out of 12 months). We also include an additional allocation of 17.6 million that was reported in the Supreme Court's official news release (Supreme Court of the Republic of Indonesia, 2019). Panel B reports the total estimated benefits. We assume that each case involves two parties earning the median Indonesian salary (Yoong and Gil Sander, 2020), three judges, two lawyers, and two court clerks. For judges, lawyers, and clerks we use the median hourly wage from ERI Economic Research Institute (2025). We assume eight-hour workdays and an average of seven days saved per case due to the reform. We aggregate savings for 100,000 cases, corresponding to the aggregated number of cases in our sample. The number is consistent to what Supreme Court of the Republic of Indonesia (2020) report. Panel C presents the net savings and the benefit-to-cost ratio derived from the calculations in Panels A and B. Assuming an average annual cost of USD 23,879,738, a modest annual increase of 15 percent in the number of treated cases (starting from 100,000 cases per year), and constant savings of USD 4,474 per case, we project a 10-year benefit-to-cost ratio of 38.04. Panel D repeats the comparison under a more conservative assumption of four-hour workdays for all participants, implying savings of USD 2,237 per case and a projected 10-year benefit-to-cost ratio of 19.02. Panel E repeats the comparison under an assumption of five days saved with eight-hour workdays for all participants, implying savings of USD 3,196 per case and a projected 10-year benefit-to-cost ratio of 27.17.

Table B5: E-Court Reform and Case Duration: Like-for-Like Case-Type Controls

	Case Duration in Days			
	(1)	(2)	(3)	(4)
E-Court Reform × Non-Criminal Case	-6.947***	-4.665**	-4.996**	-4.351**
	(2.241)	(2.268)	(2.263)	(1.629)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	Yes	Yes	Yes	Yes
County Court × Month-Year FE	Yes	Yes	Yes	Yes
Case Type FE	No	Yes	Yes	Yes
County Court × Non-Criminal Case FE	Yes	Yes	No	No
County Court × Case Type FE	No	No	Yes	No
Judge × Case Type FE	No	No	No	Yes
Mean Dep Var	38.297	38.297	38.297	38.297
R-squared	0.363	0.402	0.421	0.480
Number of Clusters	52	52	52	52
Observations	98810	98810	98810	98810

Note: The table reports like-for-like estimates of the effect of the E-Court Reform on case duration in treated first-instance courts. The dependent variable is the number of days between filing and first-instance decision. The coefficient of interest is the interaction between reform implementation and the non-criminal case indicator; criminal cases serve as the within-court comparison group. Column (1) reports the baseline saturated specification with judge fixed effects, county-court-by-month-year fixed effects, and county-court-by-non-criminal-case fixed effects. Column (2) adds mutually exclusive case-type fixed effects. The case-type bins are other criminal cases, other non-criminal cases, divorce and marital disputes cases, contract cases, narcotics cases, and theft cases. Column (3) replaces county-court-by-non-criminal-case fixed effects with county-court-by-case-type fixed effects. Column (4) absorbs judge-by-case-type fixed effects, comparing cases within the same judge and case-type bin. All columns are estimated on the same case-duration sample as the main treated-courts specification. The sample comprises treated first-instance courts, cases filed between April 1, 2009, and February 28, 2020, and judges who remained in the same court and adjudicated both criminal and non-criminal cases before and after reform implementation. The mean dependent variable is calculated among pre-reform non-criminal cases in each estimation sample. Robust standard errors, clustered at the county-court level, are reported in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table B6: E-Court Reform and Case Duration in Balanced Judge-by-Subtype Cells

	Case Duration in Days			
	(1)	(2)	(3)	(4)
E-Court Reform × Non-Criminal Case	-8.535***	-7.496***	-5.092**	-5.461**
	(1.581)	(1.627)	(2.266)	(2.120)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	No	Yes	Yes	Yes
County Court × Month-Year FE	No	No	Yes	Yes
County Court × Non-Criminal Case FE	No	No	No	Yes
Mean Dep Var	64.2	64.2	64.2	64.2
R-squared	0.065	0.209	0.307	0.320
Number of Clusters	49	49	49	49
Observations	44587	44587	44587	44587

Note: The table reports estimates of the effect of the E-Court Reform on case duration after imposing a like-for-like judge-by-case-subtype restriction. The dependent variable is case duration in days, measured as the number of days between filing and first-instance decision. The starting sample is the common treated-courts sample from the saturated case-duration specification, restricted to cases filed between April 1, 2009, and February 28, 2020, assigned to judges who remained in the same court and adjudicated both criminal and non-criminal cases before and after reform implementation. Within that sample, the code retains only detailed case-subtype cells that the same judge handled at least once before and at least once after the E-Court reform. For example, for divorce cases, the restriction keeps only judge-by-divorce cells for judges who handled at least one divorce case before reform and at least one divorce case after reform; judge-by-divorce cells observed only on one side of the reform are excluded. This restriction compares post-reform changes among case subtypes observed for the same judge on both sides of the reform. The coefficient of interest is the interaction between E-Court reform implementation and the non-criminal case indicator; criminal cases serve as the comparison group. All columns include the non-criminal case indicator. Column (2) adds judge fixed effects. Column (3) adds county-court-by-month-year fixed effects. Column (4) additionally includes county-court-by-non-criminal-case fixed effects. The mean dependent variable is calculated over all pre-reform cases in each column's estimation sample. Robust standard errors, clustered at the county-court level, are reported in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table B7: E-Court Reform and Conditional Case Reversals in Treated Courts

	% of Conditional Case Reversals			
	(1)	(2)	(3)	(4)
E-Court Reform × Non-Criminal Case	-0.037	3.326	6.120	3.287
	(4.150)	(4.260)	(6.700)	(8.554)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	No	Yes	Yes	Yes
County Court × Month-Year FE	No	No	Yes	Yes
County Court × Non-Criminal Case FE	No	No	No	Yes
Mean Dep Var	33.160	33.160	33.160	33.160
R-squared	0.010	0.221	0.444	0.452
Number of Clusters	46	46	46	46
Observations	2117	2117	2117	2117

Note: The table reports estimates of the effect of the E-Court Reform on case reversals conditional on appeal. The dependent variable equals 100 if an appealed first-instance decision was reversed on appeal and 0 otherwise. The sample is restricted to appealed cases from treated first-instance courts, filed between April 1, 2009, and February 28, 2020, assigned to judges who remained in the same court and adjudicated both criminal and non-criminal cases before and after reform implementation. The coefficient of interest is the interaction between E-Court reform implementation and the non-criminal case indicator; criminal cases serve as the within-court comparison group. All columns include the non-criminal case indicator. Column (2) adds judge fixed effects. Column (3) adds county-court-by-month-year fixed effects. Column (4), the preferred saturated specification, additionally includes county-court-by-non-criminal-case fixed effects. Mean Dep Var is calculated over the full estimation sample in each column. Robust standard errors, clustered at the county-court level, are reported in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table B8: Vertical Spillovers to High Courts: Pre-Reform Exposure Measures

	Case Duration in Days			
Panel A: Exposure = High-Court Pre-Reform Non-Criminal Caseload Share				
Pre-reform Exposure × Post × Non-Criminal Cases	-0.114***	-0.123***	-0.121***	-0.138***
	(0.032)	(0.022)	(0.021)	(0.033)
Mean Pre-reform Exposure	45.43	45.43	45.43	45.43
R-squared	0.291	0.502	0.521	0.658
Panel B: Exposure = Treated First-Instance Pre-Reform Non-Criminal Caseload Share				
Pre-reform Exposure × Post × Non-Criminal Cases	-0.036	-0.081***	-0.104***	-0.103***
	(0.044)	(0.027)	(0.022)	(0.037)
Mean Pre-reform Exposure	47.56	47.56	47.56	47.56
R-squared	0.285	0.500	0.520	0.657
Panel C: Exposure = Treated First-Instance Pre-Reform Non-Criminal Appeal Inflow Share				
Pre-reform Exposure × Post × Non-Criminal Cases	0.050*	-0.035**	-0.075***	-0.064**
	(0.029)	(0.017)	(0.013)	(0.029)
Mean Pre-reform Exposure	29.53	29.53	29.53	29.53
R-squared	0.285	0.498	0.518	0.656
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	No	Yes	Yes	Yes
County Court × Non-Criminal Case FE	No	No	Yes	Yes
County Court × Month-Year FE	No	No	No	Yes
Mean Dep Var	40.78	40.78	40.78	40.78
Number of Clusters	30	30	30	30
Observations	15923	15923	15923	15923

Note: This table reports high-court spillover estimates using three alternative pre-reform exposure measures. Panel A defines exposure as each high court's own pre-reform non-criminal caseload share. Panel B defines exposure as the pre-reform non-criminal caseload share among treated first-instance courts in the corresponding high-court jurisdiction. Panel C defines exposure as the share of pre-reform non-criminal appellate inflow from treated first-instance courts among all pre-reform non-criminal appellate inflow in the corresponding high-court jurisdiction. In all panels, the coefficient reports the differential post-reform change in high-court non-criminal case duration associated with a one-percentage-point higher pre-reform exposure. Untreated high courts are not assigned an artificial post date and therefore do not enter this specification. The sample includes high-court cases filed between April 1, 2009, and February 28, 2020, and excludes the Supreme Court. Case duration is right-trimmed at the 95% confidence interval. Judges remain in the same court throughout the sample period, and high courts are observed before and after the reform with both criminal and non-criminal cases during the sample period. Robust standard errors, clustered at the county court level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B9: Impact of E-Court Reform on Case Duration in Days by Pre-Reform Statutory Citation Complexity

	Case Duration in Days	
E-Court Reform × Non-Criminal Case × Statutory Citation Complexity	8.907***	7.888***
	(2.486)	(2.305)
Judge FE	Yes	Yes
Non-Criminal Case FE	Yes	Yes
County Court × Month-Year FE	Yes	Yes
County Court × Non-Criminal Case FE	No	Yes
Mean Dep Var	63.25	63.25
R-squared	0.287	0.300
Number of Clusters	52	52
Observations	58327	58327

Note: The dependent variable is case duration in days, right-trimmed at the 95% confidence interval. The key explanatory variable is the interaction between the E-Court reform indicator, a non-criminal case indicator, and pre-reform statutory citation complexity. Statutory Citation Complexity measures the density of legally substantive reasoning and citation-related content in the judgment text and is averaged by case subtype over the full pre-reform period. The estimation sample is restricted to cases whose subtype has statutory citation complexity observed in the pre-reform period. The sample is drawn from the main treated-courts case-duration sample. Robust standard errors clustered at the county court level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B10: Task Reallocation from Non-Criminal Exposure under E-Court Reform

Panel A: E-Court Reform on Case Duration on Criminal Cases				
	Case Duration in Days			
	(1)	(2)	(3)	(4)
Pre-Reform Non-Criminal Share × Post × Criminal Case	-0.431***	-0.327***	-0.350***	-0.405***
	(0.135)	(0.112)	(0.113)	(0.132)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	No	Yes	Yes	Yes
County Court × Month-Year FE	No	No	Yes	Yes
Judge × Month-Year FE	No	No	No	Yes
Mean Dep Var	47.25	47.25	47.25	47.25
Mean Pre-Reform Non-Criminal Share	40.47	40.47	40.47	40.47
R-squared	0.064	0.257	0.338	0.497
Number of Clusters	52	52	52	52
Observations	98810	98810	98810	98810

Panel B: Criminal and Total Cases Resolved by Judge-Month				
	Criminal Cases Resolved		Total Cases Resolved	
	(1)	(2)	(3)	(4)
Pre-Reform Non-Criminal Share × Post	-0.012	-0.002	-0.003	0.003
	(0.011)	(0.008)	(0.015)	(0.012)
Judge FE	Yes	Yes	Yes	Yes
Month-Year FE	Yes	Yes	Yes	Yes
County Court × Month-Year FE	No	Yes	No	Yes
Mean Dep Var	2.21	2.21	3.76	3.76
Mean Pre-Reform Non-Criminal Share	43.51	43.51	43.51	43.51
R-squared	0.266	0.420	0.388	0.578
Number of Clusters	50	50	50	50
Observations	26541	26541	26541	26541

Note: This table tests whether the E-Court Reform reallocates judge capacity from treated non-criminal cases to untreated criminal work. The exposure measure is the judge's pre-reform non-criminal share, defined as the percentage of the judge's pre-reform cases that were non-criminal, measured before the reform and then fixed. Panel A uses the full treated-courts sample and estimates whether criminal case duration changes differentially after reform for judges with higher predetermined non-criminal exposure. The coefficient of interest is the interaction between pre-reform non-criminal share, the post-reform indicator, and the criminal-case indicator. Column (4) adds judge-by-month-year fixed effects to the fixed effects included in column (3), so identification comes from within-judge-month comparisons between criminal and non-criminal cases. Panel A columns are estimated on the same common sample as the main treated-courts case-duration table. Panel B collapses the data to judge-by-month cells and tests whether judges with higher predetermined non-criminal exposure resolve more criminal cases after reform; total resolved cases are reported as a secondary outcome. Judge-month cells with zero closures are included within each judge's observed active spell. The sample is restricted to treated first-instance courts, cases filed between April 1, 2009, and February 28, 2020, and judges who remain in the same court and adjudicate both criminal and non-criminal cases before and after the reform. Robust standard errors are clustered at the county-court level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B11: Effects on Treated and Untreated Cases

	Case Duration in Days		
	(1)	(2)	(3)
E-Court Reform	7.811***	-1.235	-2.488**
	(2.191)	(1.484)	(1.208)
E-Court Reform × Non-Criminal Case	-27.767***	-9.513***	-6.610***
	(2.861)	(2.666)	(2.524)
Non-Criminal FE	Yes	Yes	Yes
Judge FE	No	Yes	Yes
County Court × Non-Criminal Case FE	No	No	Yes
Mean Dep Var	53.380	53.380	53.380
R-squared	0.004	0.169	0.172
Number of Clusters	786	786	786
Observations	3191483	3191483	3191483

Note: The table reports estimates of the effect of the E-Court Reform on case duration in the full sample of treated and control first-instance courts. The E-Court Reform row reports the post-reform change for criminal cases in treated courts relative to untreated courts. The interaction row reports the additional post-reform change for non-criminal cases relative to criminal cases. County-court-by-month-year fixed effects are not included because they absorb the court-month-level E-Court Reform indicator. All columns include non-criminal case fixed effects. Column (2) adds judge fixed effects, and Column (3) adds county-court-by-non-criminal-case fixed effects. Robust standard errors, clustered at the county-court level, are reported in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table B12: Sample Selection Diagnostics: Region-Level Characteristics

Treated Regions vs Other Regions	Treated	Others	Difference (Treated–Others)
	1.265	0.469	0.796***
Population (in millions, 2015)	(0.128)	(0.028)	(0.130)
			p-value < 0.01
	8.085	12.533	-4.447***
Share of Population in Poverty, 2016	(0.618)	(0.354)	(0.708)
			p-value < 0.01
	13.044	9.581	3.464***
Expenditure (in thousands, 2016)	(0.464)	(0.098)	(0.471)
			p-value < 0.01
	0.531	0.521	0.009
Share of Votes for Inc. President, 2019	(0.026)	(0.012)	(0.029)
			p-value $= 0.747$
	5.429	5.158	0.271
Unemployment Rate, 2016	(0.269)	(0.099)	(0.285)
			p-value $= 0.342$
	6.261	4.003	2.258**
# of Businesses (in thousands, 2018)	(0.884)	(0.295)	(0.926)
			p-value $= 0.015$
	0.050	0.023	0.028***
# of Markets (in thousands, 2018)	(0.010)	(0.002)	(0.010)
			p-value $= 0.004$
	0.021	0.004	0.017***
# of Hospitals (in thousands, 2018)	(0.003)	(0.000)	(0.003)
			p-value < 0.01
	0.330	0.196	0.134***
# of Violent Acts (in thousands, 2018)	(0.047)	(0.016)	(0.049)
			p-value $= 0.007$
Number of Regions	51	342	
Number of Clusters (Regions)			393

Note: This table presents region-level comparisons used to assess sample selection. A region is classified as treated if it contains at least one judge from the treated court–that is, a judge who does not move across courts, sits in a treated court both before and after the reform, and adjudicates both criminal and non-criminal cases. All other regions form the comparison group and contain only judges from the full sample outside this treated group. All variables are measured at the region level at a single point in time. Missing values are imputed using the median. Standard errors clustered at the region level are reported in parentheses. The averages and differences are calculated at the region level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B13: Sample Selection Diagnostics: Case-Level Characteristics – Non-Criminal Types

Treated-Judge Cases vs Other Cases from the Full Sample	Treated	Others	Difference (Treated–Others)
	37.001	90.315	-53.314***
% Non-Criminal Cases	(0.149)	(0.016)	(2.416)
			p-value < 0.01
	2.600	68.647	-66.046***
% Divorce and Marital Disputes	(0.049)	(0.025)	(1.567)
			p-value < 0.01
	0.615	1.163	-0.548**
% Family Property Cases	(0.024)	(0.006)	(0.238)
			p-value = 0.022
	19.701	2.173	17.528***
% Document Modifications Cases	(0.123)	(0.008)	(2.195)
			p-value < 0.01
	11.107	6.994	4.113***
% Child Custody Cases	(0.097)	(0.014)	(1.407)
			p-value = 0.004
	0.533	0.363	0.170
% Contract Cases	(0.022)	(0.003)	(0.150)
			p-value = 0.259
	2.063	0.171	1.891***
% Act of Unlawfulness	(0.044)	(0.002)	(0.138)
			p-value < 0.01
	1.293	0.040	1.253**
% Debt Accord	(0.035)	(0.001)	(0.614)
			p-value = 0.042
	0.764	0.035	0.729***
% Breach of Contract	(0.027)	(0.001)	(0.100)
			p-value < 0.01
	0.117	0.003	0.113
% Bankruptcy Cases	(0.011)	(0.000)	(0.108)
			p-value = 0.295
	0.050	0.009	0.040*
% Consumer Rights Cases	(0.007)	(0.001)	(0.022)
			p-value = 0.066
Number of Cases	104,716	3,362,592	
Number of Clusters (Courts)			802

Note: This table compares the composition of cases adjudicated by treated-group judges with all other cases in the sample. The treated group includes cases handled by judges who do not move across courts, sit in a treated court both before and after the reform, and adjudicate both criminal and non-criminal cases. The comparison group includes all remaining cases with judges who also do not move across courts. All variables are measured as percentages at the case level. Standard errors clustered at the county court level are reported in parentheses. The averages and differences are calculated at the case level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B14: E-Court Reform and Case Duration – All Courts (with all Untreated Courts)

Panel A: Case Duration

	Case Duration in Days			
	(1)	(2)	(3)	(4)
E-Court Reform × Non-Criminal Case	-19.956***	-10.249***	-10.330***	-6.947***
	(2.651)	(2.162)	(2.705)	(2.221)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	No	Yes	Yes	Yes
County Court × Month-Year FE	No	No	Yes	Yes
County Court × Non-Criminal Case FE	No	No	No	Yes
Mean Dep Var	53.380	53.380	53.380	53.380
R-squared	0.003	0.169	0.242	0.245
Number of Clusters	786	786	786	786
Observations	3191483	3191483	3191483	3191483

Panel B: Decision Quality and Contestability Measures

	% of Appeals		% of Case Reversals	
	(1)	(2)	(3)	(4)
E-Court Reform × Non-Criminal Case	-2.555***	-1.154**	-0.748***	-0.451**
	(0.425)	(0.447)	(0.137)	(0.196)
Non-Criminal FE	Yes	Yes	Yes	Yes
Judge FE	Yes	Yes	Yes	Yes
County Court × Month-Year FE	Yes	Yes	Yes	Yes
County Court × Non-Criminal Case FE	No	Yes	No	Yes
Mean Dep Var	0.449	0.449	0.166	0.166
R-squared	0.201	0.205	0.146	0.148
Number of Clusters	786	786	786	786
Observations	3191483	3191483	3191483	3191483

Note: The table reports estimates of the effect of the E-Court Reform in the full sample. Panel A reports case duration, measured as the number of days between filing and first-instance decision. Panel B reports appellate-review outcomes: whether a first-instance decision was appealed and whether it was reversed on appeal. The full sample comprises treated and control first-instance courts and cases filed between January 2, 2000, and February 28, 2020. Judges remain in the same court throughout the sample period; treated-court judges are observed before and after reform implementation with both civil and criminal cases. The coefficient of interest is the interaction between E-Court reform implementation and the non-criminal case indicator; criminal cases serve as the within-court comparison group. In Panel A, all columns include the non-criminal case indicator, Column (2) adds judge fixed effects, Column (3) adds county-court-by-month-year fixed effects, and Column (4), the preferred saturated specification, additionally includes county-court-by-non-criminal-case fixed effects. In Panel B, Columns (1) and (3) include non-criminal case fixed effects, judge fixed effects, and county-court-by-month-year fixed effects; Columns (2) and (4) additionally include county-court-by-non-criminal-case fixed effects. Mean Dep Var is calculated over the full estimation sample in each column. Robust standard errors, clustered at the county-court level, are reported in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table B15: Impact of E-Court Reform on Court Case Processing Outcomes

	Cases Filed			
E-Court Reform × Non-Criminal Case	4.902**	14.242***	-2.919	-4.131
	(1.935)	(2.030)	(2.704)	(3.062)
County Court FE	Yes	Yes	Yes	Yes
Non-Criminal Case FE	No	Yes	Yes	Yes
Month-by-Year FE	No	No	Yes	Yes
Province Specific Trends	No	No	No	Yes
Mean Dep Var	21.29	21.29	21.29	21.29
R-squared	0.171	0.216	0.461	0.497
Number of Clusters	54	54	54	54
Observations	4905	4905	4905	4905

Note: This table reports the effect of the E-Court Reform on a case processing outcome using monthly court-level data aggregated by court and case type (criminal vs. non-criminal). The dataset is structured at the month–court–case-type level and includes all months in which a case of a given type was filed or closed in a particular court. The dependent variables is Cases Filed (total number of cases filed). The main explanatory variable is an interaction between an indicator for the post-reform period and an indicator for non-criminal cases. The sample comprises cases filed between April 1, 2009, and February 28, 2020, in treated courts (ever exposed to the reform) and is restricted to cases assigned to judges who remained in the same court throughout the sample period, were observed both before and after the reform, and adjudicated both criminal and non-criminal cases. Robust standard errors, clustered at the county court level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

For Remaining Supplementary Tables and Figures in Appendix B, including additional Appendix C, please follow the link HERE.