# Impact of free legal search on rule of law: Evidence from Indian Kanoon*

Sandeep Bhupatiraju
World Bank

Daniel L. Chen
Toulouse School of Economics

Ritesh Das
World Bank

Shareen Joshi
Georgetown University

Peter Neis
Toulouse School of Economics

June 9, 2023

**Abstract**

Access to legal information is limited in many parts of the world. Can digital platforms offering free legal search reduce market-level constraints on economic development? We estimate the impact of Indian Kanoon, a free legal search engine, using a generalized difference-in-differences empirical strategy. We find that the staggered rollout of Kanoon was associated with a 1-2% increased likelihood of case resolutions and doubling of the number of appeals, which are also less likely to be dismissed by the courts. It affected the finances of firms with positive impacts on assets and negative impacts on audit fees and bad debts.

JEL Codes: K00, K15, D83, L86, R11

Keywords: Legal Information, Legal Aid, Common Law, Digital Platforms, Event Study, India

*We are grateful to the World Bank and Georgetown University's Summer grant opportunity for financial support. Daniel L. Chen also acknowledges IAST, TSE-Partnership, and ANITI funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EUR-0010. We thank Sushant Sinha, the founder of Indian Kanoon for sharing data and his experiences of establishing this platform. For research assistance, we thank Nixon Joshi, Shilpa Rao, Sunny Anand, and Shashank Singh. We are also grateful to Shibani Ghosh, Vijayendra Rao, Vikram Raghavan, Vasujith Ram, Nicholas Robinson, Mrinal Satish, and Nayantara Vohra for very helpful advice, critical insights and very stimulating discussions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of the World Bank, its Board of Executive Directors, or the governments they represent.

# 1 Introduction

Equal justice for all - a goal endorsed by all the world's nations in 2015 - requires all citizens to have timely access to legal information.[1] For most people, however, legal information remains expensive, unreliable, and difficult to understand (United Nations 2016). Policies that address these barriers and improve access have recently been shown to have notable impacts. A randomized controlled trial in South Africa, for example, showed that the provision of free legal search information to firms allowed them to make better hiring decisions and raised employment by 12% (Bertrand and Crépon 2021). Similarly, the provision of information on the number of pending cases in Pakistani state courts resulted in greater citizen trust in formal institutions (Acemoglu et al. 2020).

What happens when all citizens get access to a free online legal search platform? We examine this question in the context of India, the world's largest common law country where access to legal information services is particularly expensive for citizens and court inefficiencies have been shown to have an adverse impact on markets (Greenleaf et al. 2013, Chemin 2009; Chemin 2012).

Specifically, we study the impact of the release of Indian Kanoon, a free legal search engine that was launched in 2008 by Sushant Sinha, an overseas graduate student without any coordination with the government of India. He had observed that "finding most applicable sections from hundreds of pages of law documents is too daunting for common people" and intended for Kanoon to "bring the knowledge of law to the common people". Kanoon emphasized keyword searches and tight integration of court judgments with laws and with prior judgments to allow automatic determination of the most relevant clauses and court judgments (Iyengar 2010). Today, Kanoon is widely regarded as a "first-stop" in a search for legal information in India. According to estimates by Sinha himself, the website receives 2.9 million search queries and 1.6 million sessions per month and the average time spent per page is six minutes, suggesting people actually read the legal text.[2]

We use an event study framework that exploits the staggered rollout of the platform across different states to assess its effects on cases, courts, and firms. We scrape all available cases from the Kanoon

[1] Goal 16 of the Sustainable Development Goals commits to "Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels https://sdgs.un.org/goals/goal16).

[2] We are grateful to Sushant Sinha for this information for the period June 2010–June 2020.

website and code the resulting corpus for judgment dates, judges, courts, litigants, and citations. We then scrape the official online records of the Indian high courts (widely referred to as the eCourts High Court website), which contain additional metadata such as filing dates to measure overall court efficiencies. We also calculate the same efficiency measures for district courts using our scrape from the eCourts district court website. Additionally, we identify appeals from district courts in high courts by merging our scrape of the eCourts district courts website with the data from the eCourts High Court website. We also identify firms featured in eCourts cases and link these to the Prowess firms database, which provides financial indicators from firms' balance sheets.

To estimate the impact of Kanoon on cases, courts and firms, we estimate a two-way fixed effects model using the methods of Sun and Abraham (2021). We find that the launch of Kanoon is associated with an increase in the speed of processing cases and a slight increase in the number of filings per court-month. It is also associated with a discernible increase in the number of decided cases in the year after the rollout. Most notably, the number of appeal cases nearly double after the launch of Kanoon. We also see a 1-3% increase in the likelihood of an appeal case being withdrawn and a 2-6% decline in the likelihood of a case being dismissed after the rollout of Kanoon. This suggests that the better availability of legal information may have not only increased the number of appeals, but also their quality, for they are more likely to be heard within the court.

Turning to Prowess, we conduct a panel analysis of firms' financial indicators. We find that the emergence of Kanoon reduced search costs and had an overall positive impact on firm balance sheets. Kanoon had sizeable impacts on assets and reductions of bad debt, audit fees, and bank guarantees, whether we look at firms that filed at least one case or at all firms. This reinforces the findings of earlier studies showing an increase in employment resulting from the availability of free legal search (Bertrand and Crépon 2021). The magnitude of the effects conjectures the potential positive complementarities as citizens increasingly get access to free legal search.

Taken together, these results suggest that the release of Kanoon reduced the cost of legal information to the Indian population, reduced bottlenecks, increased efficiency, and contributed to improvements in the quality of legal research in the courts. These improvements had in turn a significant effect on the economy and citizens' life outside of court. This case study offers valuable insights on the

importance of making legal information more accessible to all stakeholders in justice systems and on the value of free and accessible information in general.

We contribute to three key strands of literature. First, we contribute to the body of work exploring the links between the functionality of the legal system and economic development (Djankov et al. 2003; Ponticelli and Alencar 2016; Lichand and Soares 2014; Visaria 2009; Kondylis and Stein 2018; Chemin 2020). We find support for the broadly held view that incomplete information about laws and regulations has both direct and indirect costs. Our findings also reinforce previous evidence from randomized trials highlighting the causal relationship between access to legal information and economic outcomes (Bertrand and Crépon 2021).

Second, economic theory has long ago shown that the reduction in the cost of gathering information improves decision-making efficiency (Stigler 1961; Diamond 1971; Varian 1980). Empirically, the recent literature confirmed that reductions in search costs increase overall market efficiency (Goldfarb and Tucker 2019). We show that the reduction of search costs for legal information can alleviate market-level constraints to economic development with substantial positive impacts on firm assets and negative impacts on bad debt.

Third, we contribute to the literature on the consequences of procedural formalities in the judiciary on economic outcomes (Djankov et al. 2003; Chemin 2009). The causal impact of a drop in the cost of information and the improved access to legal information for all stakeholders on outcomes such as courts' productivity and firms' performance remains poorly understood. We examine this question in the case of India, a context where the barriers to accessing information are particularly high and courts are significantly inefficient (Rao 2021).

The remainder of this paper is structured as follows. Section 2 presents background information on the Indian legal system and the general context of this study. Section 3 presents our research hypotheses. Section 4 presents an overview of our data and some descriptive statistics of the samples that are used for analysis. Section 5 presents our empirical strategy. Section 6 presents the results of our analysis. Section 7 explores some mechanisms that may explain our results (or not). The final section concludes.

# 2 Background

## 2.1 The Indian Legal System

India's legal system goes back several centuries (Jois 2004). Its modern system of justice, however, is grounded in the common law justice system that was established by the British colonial administration in the 19th century (Baxi 1982; Galanter 1963). The current judicial system is significantly shaped by the Constitution of India, which was written after Indian independence in 1947.

The court system of India comprises a hierarchical structure with the Supreme Court of India at the top, the high courts as the next tier and subordinate courts at district, municipal and village levels. The justice system is an integrated system, which means that decisions made by higher courts are binding on the lower courts. Appendix Table A1 gives an overview of the 25 Indian high courts, their establishment date, and the states and union territories they have jurisdiction over.

While the Indian judiciary commands a high level of public trust, it has been increasingly criticized for a growing backlog of cases, lengthy delays in outcomes and inefficiency (Krishnaswamy and Swaminathan 2019). There are currently more than 5.9 million pending cases at the high courts, even though their average rate of disposal between 2015 and 2019 was about 1.8 million cases per year.[3]

Long and complex proceedings discourage citizens from seeking justice and distort markets (Krishnan et al. 2014; Rao 2021; Chemin 2012). Chemin (2009) for example, demonstrates that 430 high court amendments to the Code of Civil Procedure between 1971 and 1996 lengthened trials and these delays affected credit markets, agricultural development, and manufacturing performance across the country.

## 2.2 Legal information

In the past twenty years, Indian courts have adopted new technologies to bring greater transparency into the system. This began around 2000 when the Supreme Court of India began to publish some of its cases online. In 2005 the Chief Justice of India, R.C. Lahoti, established the Information Technology and Judicial Reform Cell to promote the introduction of online legal information and services at all

[3] National Judicial Data Grid (NJDG), https://njdg.ecourts.gov.in/hcnjdgnew/, accessed on March 4, 2022.

courts in the country. The first phase of the program, which lasted from 2007 to 2015, focused on the computerization of courts with the installation of hardware, network infrastructure, and software that could provide basic case related services to litigants and lawyers. In the years that followed, cases were uploaded on the websites of the Supreme Court, high courts, district courts and tribunals.[4] The top panel of Figure 1 gives an overview of when cases of the different high courts became first available on their websites.

To upload cases on their websites, states relied heavily on the National Informatics Centre (NIC), a public sector corporation that is responsible for hosting, maintaining and updating the websites of constitutional, central, state and local government agencies across India. This organization has maintained the websites of the Election Commission, Planning Commission, tax authorities, and the Securities Exchange Board of India (SEBI). Iyengar (2010) notes that considerable legal information was provided at each of these sites through the complete texts of applicable legislation, subordinate legislation, administrative rulings, reports, census data, application forms, etc. This served as the foundation for the emergence of the Kanoon platform.

### 2.2.1 The Emergence of IndianKanoon.org

The development of IndianKanoon.org began in the summer of 2007 and was publicly announced on 4 January 2008. The entire effort was led, financed and managed by Sushant Sinha, a graduate student in the Computer Science Department at the University of Michigan.

The project began informally. Sinha's self-stated goal was to "bring the knowledge of law to the common people".[5] He relied minimally on physical infrastructure or hardware and utilized free and open-source software for the purpose of scraping cases from court websites and building a searchable database that could be helpful for legal education and research (Iyengar 2010).[6]

[4]The second phase of the program, launched in 2015 intends to improve the experience of litigants, lawyers and other stakeholders in the courts through provisions such as information in local languages, applications for mobile phones, kiosks in court complexes, the delivery of certified copies of documents via electronic platforms and the deployment of ePayment Gateways for making deposits, payment of court fees, fines, etc. Throughout this period however, cases have been uploaded to the court websites.

[5]See `indianKanoon.org/about.html`, accessed April 1, 2022.

[6]Iyengar (2010) interviewed Sushant Sinha in 2010 and documented a heavy reliance on a database in *Postgres*. This was favored for its' "inbuilt search functionality, inverted index and ranking functions" (Iyengar 2010). When existing packages were inadequate for his needs, Sushant Sinha developed patches for the broader community of software developers. His

In his reflections from a personal blog, as well as the objectives stated on the website itself, two key priorities emerged in the initial phases of the project. First, great effort was made to ensure that information about laws was accessible to a broad range of stakeholders. In his own words, Sinha argues that "acts are very large and in most scenarios just a few sections of laws are applicable". As a result, "finding the most applicable sections from hundreds of pages of law documents is too daunting for common people". He placed considerable emphasis on the ease of keyword searches, particularly for terms that are of interest to a broad audience.

A second priority was to illuminate how laws are interpreted by the courts. He argues that "laws are often vague and one needs to see how they have been interpreted by the judicial courts". On the website, laws and judgments were thus separately maintained. Legal documents were broken down into smallest possible clauses. A tight integration of court judgments with laws and with prior judgements allows automatic determination of the most relevant clauses and court judgments.

The coverage of Indian Kanoon was extended in waves. When the website was launched in January 2008, only cases from the Supreme Court and the texts of central (federal) legislations were featured. By 2009, judgments of 10 high courts and 17 tribunals had also been uploaded. The full text of India's Constituent Assembly debates, Law Commission reports and the full texts of central legislation were also added throughout this time. The timing of the uploading of cases on the website is summarized in the bottom panel of Figure 1.

The public response to the site appeared to be overwhelmingly positive from a brief study of users (Iyengar 2010) and the comments received by Sinha himself. Users were enthusiastic to have a legal resource that is accessible, user-friendly and free (Iyengar 2010). The site also has some interesting features: the data is searchable, pages have links to posts or other writings on the internet that refer to the cases and there are cross-links within judgements to cited cases. These innovations significantly enhance a user's experience of accessing the law, reducing search costs for relevant information substantially.

There are, however, also some limitations. Unlike officially curated cases, Kanoon does not provide a case note, it cannot be officially cited as a source and it is not accepted by judges during official

efforts contributed to an improvement in the 'headline citation' functionality of Postgres, which facilitates the retrieval of contextual information associated with search queries.

proceedings. Web scraping errors are not manually corrected, making it somewhat less reliable than paid databases.

To date, there has been no rigorous evaluation of the impact of the Kanoon platform. Sinha continues to maintain the site, without formal funding or any formal organizational structure. He has become an advocate for freedom of information in India and protecting the rights of citizens to get information from their government.

### 2.2.2 Other Databases

Indian Kanoon was not the only electronic source of data at the time it was released. The increased online availability of legal data from the courts facilitated a proliferation of electronic resources for Indian legal research. A detailed list is available in Appendix Section A1 of this paper. These resources were different than Kanoon in several ways: they were expensive, difficult to access and had only partial coverage of Indian law. Sinha described these other databases as follows:

> Until very recently, most law resources in India were provided by libraries or Websites that charged a significant amount of money. In effect, they prohibited access to a significant portion of the population that wanted to look into legal issues. The average time spent per page on the Indian Kanoon Website is six minutes; this shows that most users actually read the legal text, and apparently find it easier to understand than they had previously expected.[7]

Since 2015, there has also been an effort to curate all available Indian laws as well as their amendments in a single repository. These are available at http://www.indiacode.nic.in/. This website, which has been functional since 2017, includes all central Acts and subordinate legislation. The portal presents the complexity of the chain of laws, starting from the "parent" act to the subordinate legislation. This website however, remains difficult to access for those who do not have a legal background. Sinha emphasizes that "lawyers are often accustomed to using these interfaces, and of course understand these technical legal terms" but "requiring prior knowledge of this kind of technical legal

[7] https://blog.law.cornell.edu/voxpop/tag/indiankanoon/, accessed on June 23, 2022.

information as a prerequisite for performing a search raises a big barrier to access by common people" (Sinha, 2022).

In summary, Indian Kanoon has been the *only* free electronic resource that was readily available and accessible to the people of India since 2007. It is widely regarded as a first-stop in the search for legal information, not just for lawyers but lay citizens. Lawyers often browse this site to curate data and then turn to paid databases for adding details and formal citations prior to presenting their work in an official capacity.[8]

# 3 Research Hypotheses

We postulate that the release of Kanoon brought a sudden, substantial and entirely exogenous reduction in the cost of searching for legal information in India. This affected litigants, courts and markets. The immediate effects are likely to be driven by the changes in behavior of stakeholders who have already filed cases. As stakeholders adjust to the new technology and form new expectations however, long-term impacts are likely to become evident. Given that the mean age of decided cases in the high courts (district courts) is about 1.37 (2.13) years (Table 1), we expect the effects to take at least three years to unfold.

## 3.1 Individual Cases

Though the rulings on the Kanoon website could be helpful to litigants in all areas of India's justice system, the most immediate beneficiaries are likely private citizens who had filed cases or were appealing cases at the high courts.[9]

Easier access to legal information can immediately alter both expectations of success as well as the

[8] A cursory examination of India legal research guides at libraries across the United States, particularly libraries of prominent law schools (Harvard University, Yale University, Georgetown University and others) found Kanoon remains prominently listed in the recommended research platforms with a note that the service is free, easy to search and requires no formal registration.

[9] High courts in India have four areas of jurisdiction: (1) original jurisdiction (i.e. the authority to hear certain types of cases that cannot be heard at subordinate courts), (2) appellate jurisdiction (i.e. the authority to hear appeal cases by any subordinate court), (3) advisory jurisdiction (i.e. the review of cases sent to it by government departments, (4) legislature or the governor) and judicial review (i.e. the review of any judgment or order developed by any subordinate court). Only the first two areas of jurisdiction feature citizen cases.

probability of winning a case. Citizens who use the platform may be able to make clearer and stronger arguments in the proceedings. This may make it easier for the judge to make a decision on the case, and thus reduce the time for resolution. On the other hand, however, it can also increase the complexity of the case and thus increase the time taken for judges to review. Since these two mechanisms have opposing effects on the time taken to resolve cases, the ultimate impact of Kanoon is an empirical question.

Since high court rulings serve as binding precedents for the lower courts, we are also likely to see the impact of this platform on the lower courts. Here too petitioners who gain access to legal information can improve the quality of their arguments. This could result in higher quality decisions, fewer appeals and fewer reversals at the high court. In the immediate aftermath of the availability of the platform however, when cases have already been filed, Kanoon may cause more cases to be appealed thus raise the number of appeals at the high court.

## 3.2 Courts

Easier access to legal information can directly affect aggregated performance indicators for courts. We will consider a range of outcomes: the number of filings, the number of decisions, the number of pending cases, the age of the pending cases, the age of decided cases and the rate of clearance of cases in both the high courts and the subordinate district courts.

## 3.3 Firms

Easier access to legal information should have benefits not only for stakeholders within the legal system, but also the broader set of economic actors who rely on this information for economic activity. Firms operating in India face a complex economic, regulatory, and legal landscape for doing business (Bloom et al. 2013; Bertrand and Crépon 2021). Prior to the arrival of a free database like Indian Kanoon, accountants and managers at firms largely relied on lawyers to provide guidance on accounting practices. With the arrival of a free database however, all employers were likely able to access the full body of laws that affect their balance sheets and their business at large. On the basis of this,

we expect the rollout of Kanoon to have favorable impacts on all measures of firm financial status. There is also, of course, the possibility that the arrival of Kanoon enabled firms and customers to increase litigation – this effect however, will be measured in our analysis of the productivity of the courts themselves.

# 4 Data and Descriptive Statistics

## 4.1 IndianKanoon.org

We scraped publicly available data on Indian Kanoon for the years 2005–2015. This resulted in a corpus of 5,632,421 cases. This corpus includes both criminal and civil cases. We coded the cases for the date of judgment, court name, judge name(s), party names, advocates, and cases cited within the judgment. We also added other useful data not readily available on the website, such as the dates on which individual judgments were uploaded on the website.[10]

Figure 2 presents a summary of the data on cases that are found on the Kanoon database by high court, grouped by the dates of the Kanoon rollout, which are depicted in solid vertical lines on each graph. This figure documents the first stage with a tremendous increase in cases available around the dates of the roll out.

## 4.2 eCourt High Courts

We scraped summary data from the eCourts portal to obtain summary information on case types, filing dates, and decision dates (if a given case has been decided), as well as the names of the presiding judge, plaintiff, and respondent. We exported case lists from each eCourt website of an Indian high court, and analyzed cases filed (opened) and cases decided (closed) as a function of their varying Kanoon rollout dates. This gave us a sample of 11,894,096 cases.

Figure 3 presents the total number of filings in the eCourts system for each of the four sets of high courts and the corresponding rollout dates that pertained to them. The figure demonstrates a striking

[10]We greatly appreciate the support of Sushant Sinha, the founder of Indian Kanoon, in procuring this data. We further thank him for providing additional insight into search engine mechanics and design.

level of variability in the resolutions across the four groups of courts, as well as over time. Here too, we see some overlap between the timings of the rollout and the steep increase in the number of cases in the system. In panel (a), at the top-left of Figure 3 we see that the rollout data precedes the increase in the number of cases, while in panel (d), at the bottom right, the rollout date occurs in the midst of this increase.[11]

## 4.3 eCourt District Courts

The Kanoon platform did not include records from the 2,800 district courts of India. Given that Kanoon may have been most useful to stakeholders in less affluent regions of India however, we can expect impacts in the district courts: decisions to open or not open a court case affect the district court system, where cases in the trial of first instance are simpler than the cases on appeal in the higher courts. We include all cases from these courts which were open at any time between 2000 and 2020 in our analysis. We examine the total numbers of cases filed and decided as a sum of district court cases within each high court's jurisdiction.

Figure 4 depicts the number of filings in these data. We note that there is a jump in the number of cases filed in 2010 for many states, and a gradual increase in the number of cases filed for many states post-2012. As noted previously, this is consistent with the previous observation - these increases occurred very close to the time of the Kanoon rollout.[12]

## 4.4 Prowess

Prowess is a database that curates financial information of nearly 34,000 listed private and public companies in India. It covers nearly all companies on the National Stock Exchange and Bombay Stock Exchange.[13] These firms account for more than 70% of industrial output and 75% of corporate taxes

[11] In the appendix, we present additional data on case resolutions. We find that the overall patterns are quite similar to the total number of filings shown here.

[12] Case resolutions are presented in appendix Figure A4, and these too show a similar pattern as the number of filings in the district courts.

[13] These are registered companies that disclose their financial statements according to the 1956 Companies Act. Initially, the companies had to meet one of the following conditions to be included in the database: either the firm needed to have a turnover of at least 2.5 crore rupees, or the firm's annual reports must be available for at least two years before the date of updating.

collected by the Indian government. The database has been widely used in academic analysis (Goldberg et al. 2010). The data is collected, supplied and continuously updated by the Centre for Monitoring the Indian Economy, an independent, non-governmental research organization. The database is compiled from firms' audited Annual Reports and information supplied to the Ministry of Corporate Affairs, as well as company filings with stock exchanges and the prices of securities listed on the main stock markets in the case of publicly traded corporations.

We conduct our analysis on two samples of these firms. First, we use the full sample, regardless of whether or not the firm has pending cases at the courts. Second, we examine the sub-sample of firms that have active cases (either pending or new) during the time-period 2006-2015. To do this we identify all the firm Corporate Identification Numbers (CIN) - unique firm identification numbers - in the Prowess database and then restrict the sample of the eCourts data to the cases where these entities are litigants in the cases.[14]

## 4.5 Data Construction

Our main analysis relies on the samples drawn from the eCourt High Courts and ecourt District Courts data. The timing of the rollout of the Kanoon platform (Figure 1) together with the filing dates and judgement dates from these records allow us to estimate the impact of the platform on cases as well as court efficiency.

### 4.5.1 Appeal Cases

India's high courts have the jurisdiction to hear appeals from both civil and criminal cases from the lower (district) courts.[15] We identify these cases in our data. We begin by linking the records from the lower to the upper courts (see Appendix Section A4.1 for a detailed description of this process).

[14]The CIN number is a 21 digit alpha-numeric number that is given by the Registrar Of Companies of various states under the Ministry of Corporate Affairs (MCA). The CIN number is typically used to track all the activities of an enterprise after its registration by the government of India. This number contains the identity of an organization and additional information such as the type of company, the founding date of the company, the state code and the types of exemptions that a company is subjected to.

[15]Civil cases that were decided by district judges and Munsif courts can be appealed at the high court. Criminal cases that were decided by district judges, judges at Munsif courts and the sessions courts can also be heard as long as the sentence for imprisonment is seven years or more. High Courts can also hear appeals from the orders of the tribunal.

In the district courts data, for each state we construct an estimate of the total number of cases that are registered and decided in every year-month between January 2001 and December 2018. Of the cases that are registered, we then count the number of these cases that are appealed at a high court at any time in the sample period. On average 0.3% of the cases that are registered are appealed in the high courts within 1 year.[16] We repeat this for the number of cases that are decided. An average of 0.75% of decided cases are appealed within 1 year.[17] From this we calculate the share of the total registrations and decisions per state, year and month which are appealed. These are referred to as % Appealed by reg. date and % Appealed by dec. date respectively.

An appeal filed by a litigant moves through the pipeline of justice in several stages (see Appendix Figure A12). We rely on the official language of the courts themselves to label these stages.[18] When the case is first filed at the high court, a case is marked as "Admitted" (the paperwork is in order and the case is deemed suitable for processing at the high court and listed to the court for hearing/decision) or "Rejected" (the application is not in the jurisdiction of the court - territorial or in content - or the paperwork has defects of non-curable nature). If a case was admitted, it can be heard by the high court and when the bench takes a (whatever kind of) decision the appeal is said to be "Disposed". A disposal can be of different natures: the appeal can be "Allowed" (accepted) or "Dismissed" by the bench or "Withdrawn" by the petitioner.[19] gives and overview over these different stages and potential outcomes of an appeal case.

### 4.5.2 Firms

We similarly linked the courts in the e-courts databases to firms in the Prowess database. Of the 83.8 million cases in the ecourts database, we found 1.16 million cases that matched to firms with distinctive CINO numbers. Of these 718,000 cases have CINO numbers that matched with firms in the Prowess database. This corresponds to 9246 unique Prowess firms – this corresponds to 18% of all the firms in

[16] 0.33% cases are appealed within 2 years and 0.35% cases are appealed within 5 years.

[17] 0.82% cases are decided within 2 years and 0.91% cases are decided within 5 years.

[18] These are official terms in the eCourts system that are applied by the e-filing administrator at the time of the review of the paperwork in the case (see https://ecourts.gov.in/ecourts_home/static/manuals/efiling-User-manual.pdf)

[19] Theoretically, the decision to withdraw can be made by the appellant at any time, even at the administrative stage before being admitted.

the Prowess data.

Summary statistics of all key variables from these different datasets are presented in Tables 1 and 2. Appendix Table A1 gives an overview of the exact Kanoon rollout dates and the data coverage for high courts and each state's district courts.

# 5 Empirical Strategy

Our empirical strategy begins by estimating the before-after difference in the outcome variables between cases at the courts where Kanoon was introduced and those where it was not. A court is considered treated from the date on when its cases are uploaded on Kanoon. We refer to this event as the "Kanoon rollout date" for that court. On this date, all cases that were filed *prior* to the rollout date *and* available for review on the court websites are available in an accessible format. For all the dates that follow, cases are added to the websites on the date that they appear on the court websites. Though Kanoon now covers all the courts of India, we restrict our attention to the set of courts where the rollout first occurred and the rollout date is clearly known.[20]

We estimate the following two-way fixed effects (TWFE) model at the case level:

$$Y_{ict} = \alpha + \sum_{j=-4}^{5} \beta_j \text{Kanoon}_{jct} + \delta X_{ct} + \varepsilon_{ict} \tag{1}$$

where $Y_{ict}$ is an observed outcome of case $i$ at court $c$ filed in year-month $t$. $Kanoon_{jct}$ is a set of dummy variables that take value 1 if court $c$ had cases uploaded on Kanoon at $j$ number of years before or after time $t$ (and 0 otherwise). We omit the treatment dummy for the 12 months before the filing of a case ($Kanoon_{-1ct}$). $X_{ct}$ includes year, court, court*year and month fixed effects. We estimate equation 1 using OLS regression. Standard errors are clustered at the court-level.

We examine several types of outcome variables ($Y_{ict}$). We begin by examining the time taken to case resolution: *Resolved* is a dummy variable equal to one if the case is resolved (and 0 otherwise); *Resolved < 1 Year* is a dummy variable equal to one if the case is resolved in less than one year after its filing (and 0 otherwise). We also look at the parties contesting the cases. The variable *Government*

[20] We thank Sushant Sinha for all the details of the rollout in these initial courts.

is also a dummy variable whose value equals one if the petitioner or the respondent is a government organization (and 0 otherwise).

Next, we adapt equation 1 in order to estimate the impact of Kanoon on aggregate measures of the performance of courts:

$$Y_{ct} = \alpha + \sum_{j=-4}^{5} \beta_j \text{Kanoon}_{jc} + \delta X_{ct} + \varepsilon_{ct} \tag{2}$$

where $Y_{ct}$ is an observed outcome of high court $c$ in year-month $t$, and the remaining variables are defined as for equation 1. We consider a broad range of measures of court-level efficiency: the number of filings, the number of resolved cases, the number of pending cases, the backlog, the mean age of decided cases, the mean age of pending cases, the clearance rate and the time taken to disposal.

Similarly, we calculate the same efficiency parameters at the state-year-month level, aggregating over all district courts in the respective state. Here Kanoon rollout is defined with respect to the high court which has jurisdiction over the state.

We also consider the progression of cases from the lower district court to the upper high court in the form of appeals. For these cases, we examine how many of the cases from district courts in a state-year-month were allowed (*Allowed*), appealed (*Appealed*), dismissed (*Dismissed*), disposed (*Disposed*), overturned (*Overturned*) or withdrawn (*Withdrawn*).

Lastly, we inspect the downstream impact Kanoon had on firms. For this, we perform an event study analysis on a panel of Indian firms. We consider the following specification:

$$Y_{fcy} = \alpha + \sum_{j=-4}^{5} \beta_j \text{Kanoon}_{jcy} + \delta_c + \gamma_y + \sigma_f + \varepsilon_{fcy} \tag{3}$$

where $Y_{fcy}$ is a financial indicator of firm $f$ based under the jurisdiction of high court $c$ for financial year $y$. $\text{Kanoon}_{jcfy}$ are dummy variables equal to one if the financial year $y$ coincides with the $j$'th year pre- (or post) Kanoon rollout. $\delta_c$, $\gamma_y$ and $\sigma_f$ are court, financial year and firm fixed effects respectively.

We estimate this equation in two ways. First, we examine the full sample of firms. Next we estimate the sample of firms that actually have filed cases, as identified in the Kanoon data.

The coefficients of interest to us in all three settings (case-level, aggregated court-level and firm-level regressions) are $\beta_j$ where $j = -4, \ldots, 5$ where we exclude $j = -1$ and interpret the results relative

to this baseline. In order to interpret the coefficients $\beta_1, \ldots, \beta_5$ as the average treatment effects on the treated (ATT) of the introduction of the Kanoon platform on the outcome variables, we rely on the parallel trends assumption, i.e. in the absence of treatment, the difference between treated and untreated observations remains constant over time. We also assume that the court-level average treatment effects are homogeneous across treated courts and over time. We discuss the possible concerns with these assumptions, and our proposed solutions, below.

## 5.1 Econometric Challenges

The specifications above eliminate some specific sets of confounding factors in the impact of Kanoon. The inclusion of court and year fixed-effects rules out the possibility that Kanoon's observed impact is induced by court-specific factors that evolve over time, such as the adoption of internet and technology systems. This is important considering that the Indian judiciary is a single, integrated, common-law system with a single set of laws and operating procedures throughout the country. This framework also allows us to rule out the role of temporal trends such as macroeconomic fluctuations, changes in internet regulations, digital privacy laws, etc. To the extent that such factors might affect all courts in a similar way, year fixed effects allow us to rule out such concerns.

Causal identification of the impact of Kanoon on outcomes, however, hinges on the independence of Kanoon's rollout timeline. This assumption would be violated if the Kanoon rollout for a state's high court coincided with a substantive change in law in this state, a change in the judicial functioning of the state, or if it followed closely the digitization of the state's courts itself. If every time a high court digitized and started to publish case level data online, Kanoon followed closely and included the court in its database, we could not extract the relative weights of the causal effect of Kanoon and the causal effect of digitization.

This concern does not apply to our study for several reasons. First, the Kanoon rollout was independent of the timing of the rollout of the content on high court websites (Figures 1 and 2). Second, the eCourts web hosting program was launched after the years that we are focusing on in this analysis. The cases uploaded in the first 5 years of the digitization initiative were those that had been previously decided and placed in the public domain.

That being said, the order of expansion across states may be correlated with other confounding factors. We thus take additional steps to ensure the robustness of our findings. First, we assess the presence of unforeseen or unmeasurable confounding variables by examining pre-trends. More specifically, we assess the validity of the parallel trends assumption in a dynamic model that includes four years prior to the launch of Indian Kanoon in a given jurisdiction. We analyze the association between the launch of Indian Kanoon with outcome variables that are normalized relative to one year prior to the arrival of Indian Kanoon in that jurisdiction. Second, we conduct placebo exercises suggested by De Chaisemartin and d'Haultfoeuille (2020). The Placebo estimates are obtained by assuming that instead of the treatment happening at t, it occurred at time t-k where k $\in \{1,2,3,4\}$. We present these placebo estimates in the Appendix. Both sets of analyses lend confidence to our causal inference.

Another challenge to causal inference is the issue of heterogeneous treatment effects. Recent literature in econometrics has demonstrated that TWFE models such as ours can deliver consistent estimates only under relatively strong assumptions about homogeneity in treatment effects. In our case, we might suspect that Kanoon might have stronger or weaker effects when employed in courts with quite different underlying characteristics. This could happen if legal system participants in poorer areas rely more heavily on Kanoon, or if Kanoon is more effective in providing timely information in jurisdictions where high quality, low-cost internet services are commercially available. Given that jurisdictions assigned different Kanoon rollout dates are in fact poorer or richer, and have different markets for web access, the TWFE model cannot identify the causal effect of Kanoon on the 'average' Indian court jurisdiction.

To address this identification problem, we implement a re-weighting at each time interval of all groups which (at that time) have yet to be treated against groups which have been treated. We follow the methodology proposed in Sun and Abraham (2021). This method involves estimating the underlying weights on cohort-specific average treatment effects with auxiliary regressions to remove contamination from spillover effects from earlier time-periods (Sun and Abraham 2021).[21] This method is similar to other proposed corrections in recent literature (see for example, Borusyak and Jaravel 2017, Callaway and Sant'Anna 2020, De Chaisemartin and d'Haultfoeuille 2020 and Goodman-Bacon

[21] We use the STATA package entitled "eventstudyweights" to conduct this analysis (Sun and Abraham 2021)

2021). Our element-wise difference-in-differences (DID) estimator is consequently an unbiased and consistent estimator for the causal average treatment on the treated (ATT) of each group of courts treated with Kanoon at the same points in time ('cohort'). By weighting these estimators based on their cohort's share of the untreated at each point in time, we report the causal effect of Kanoon in the context of interest, without needing to assume away heterogeneous treatment effects.

# 6 Results

We present and interpret our baseline estimates of the causal effect of Kanoon in three broad areas: (1) The efficiency of courts; (2) Outcomes of cases from lower courts that are appealed at higher courts; and (3) Financial status of firms.

We mainly present the results from estimation of the element-wise difference-in-differences (DID) estimator (Sun and Abraham 2021). We present the results that examine pre-trends in a set of placebo exercises suggested by De Chaisemartin and d'Haultfoeuille (2020) in the Appendix.

## 6.1 Impact on Efficiency of Courts

### 6.1.1 High Courts: Case Analysis

Table 3 presents results of the two-way fixed effects estimators on case-level data for 14 high courts. We use well-defined Kanoon rollout dates, i.e. the date on which the cases from that court were made available on the platform, to estimate its impact. We start with a simple specification and build up to our preferred specification with all fixed-effects and controls for the first dependent variable, a dummy variable that measures whether a case in our data is resolved (columns 1-4). The results suggest that cases that were filed during the first and third year after the inclusion of their high court on Kanoon are significantly more likely to be resolved compared to cases which are filed during the year before the Kanoon rollout. These effects are modest and represent a 1-3 percentage point increase in case resolution compared to the year before the Kanoon rollout. Interestingly, the effect three years after Kanoon rollout is larger than the significant effect in year 1.

In column 5 of Table 3 we investigate the impact on fast resolutions, which we define as cases being resolved during the first year after their filing in the high court. Here, we observe a statistically significant effect only for cases filed during the year of the Kanoon rollout for that court. There, the likelihood of a case to be resolved within one year is increased by 2.2 percentage points, and this effect is significant at the 10% level. Besides this immediate effect, there are no statistically significant effects in the longer run.

Looking at columns 4 and 5 together, we observe an immediate but only short term impact of a high court's Kanoon rollout on the probability of being resolved during the first year after their filing but a medium term impact on the overall probability of being resolved. Lastly, the results in column 6 suggest that the emergence of the Kanoon platform does not change the likelihood of the government being a petitioner or respondent in the case.

#### 6.1.2 High Courts: Court Efficiency

Next we estimate the impact of Kanoon at the high court level (Equation 2). We consider eight outcomes: the number of filings, the number of decisions, the number of pending cases at the end of a month, the number of pending cases at least one year old at the end of a month (backlog), the average age of decided cases during a month, the average age of all cases pending at the end of the month, the number of cases decided divided by the number of cases filed during a month (clearance rate) and the number of disposed cases decided by the number of filings during a month (disposition time). As described in the last section, we use the technique from Sun and Abraham (2021) to adjust the weights in a context of potentially heterogeneous treatment effects. We calculate all these different court efficiency parameters at the high court-year-month level and define the treated observations of a court as all months following the Kanoon rollout of this court.

Coefficients of interest are presented in Figure 5. While most estimates are either noisy or show no impact of Kanoon on the aggregate court efficiency of high courts, we observe three results. First, the number of filings increases slightly, but not significantly after the Kanoon rollout. Second, the number of decisions increases in the year following the rollout. And lastly, backlog increases steadily over time after the rollout although these effects appear not to be statistically significantly different from zero.

The increase in filings and backlog creates a pronounced decline in the clearance rate in the aftermath of the Kanoon rollout.

Since we find some evidence of pre-trends in the analysis of the mean age of decided and pending cases (lower panel, Figure 5), we exclude consideration of these results. Placebo tests for the event study estimations are presented in Appendix Figure A5 and rule out the presence of pre-trends in the three years prior to the rollout of Kanoon.

In results not shown here (included in Appendix A6), we also examine the impact of Kanoon on a set of outcome variables that capture the citation networks of these cases on India Kanoon. In the Indian common law system prior cases from high courts and the supreme court represent binding presents for the courts. Therefore, as it is the case for citing laws, citing these prior cases is an essential part of judicial reasoning and part of the written decisions. We examine whether the Kanoon rollout changed which previously decided cases are cited by high court cases. We observe almost no statistically significant impacts on any measures of citations in the cases uploaded to Kanoon.

## 6.2 District Courts

To study whether and how high courts' Kanoon rollout impacted district court efficiency, we apply again the method from Sun and Abraham (2021). We calculate the same court efficiency parameters as we used for the high courts (filings, decisions, pending cases, backlog, average age of decided cases, average age of pending cases, clearance rate and disposition time) at the state-year-month level. Here, a state is defined as treated when the high court which has jurisdiction over this state is included on Kanoon.

The results of these regressions are presented in Figure 6. Note that the number of filings shows a small initial dip but then a steady increase in the first four years after the launch of Kanoon, and these effects are statistically significant in the third and fourth year. After this, the effect decreases again and has a large confidence interval which includes zero. The coefficients for decided cases in Figure 6 show that the number of monthly decided cases drops in the first two years but then reverts back to the initial levels and increases to achieve its largest level after the third year of the rollout.

One of the most noteworthy result of Figure 6 is the statistically significant increase in pending

cases and backlog (pending cases older than one year) in the year of the rollout and an even larger effect in the following year. In the following years the estimated effect size is oscillating between the point estimates of the first two ears, although they have then large confidence intervals including zero.

We also note an increase in the mean age of decided cases throughout the rollout period, though the confidence intervals include 0. The mean age of pending cases increases in the first two years before dropping and becoming close to 0 in the later years after the rollout.

We see some evidene of pre-trends in the clearance rate and disposition time in these graphs. We thus do not rely on them for interpretation. Placebo tests for the event study estimations are presented in Appendix Figure A6 and rule out the presence of pre-trends in the three years prior to the rollout of Kanoon.

In summary, we see strong impacts of the Kanoon platform on the efficiency of district courts, though the immediate impacts two years after the rollout are quite different than the longer term impacts, suggesting that there was a period of transition. Next, we explore the impacts of the platform on a specific category of category of category of cases that are shared between them: appeals.

## 6.3 Appeals from District Courts at High Courts

As discussed earlier, we identify the cases that were filed in the district courts and then later appealed in the high courts during our period of study. We link the records from the lower to the upper courts and identified appealed cases in each state, year and month.

We first examine the likelihood that a district court case is appealed in the high court. Figure 7 presents estimated coefficients and 95% confidence intervals for $\beta_j$s from a slight modification of Equation 2. We construct estimates of the efficiency of appeals in two different ways. In the top row of Figure 7, efficiency is defined by the difference in months between a state-year-month observation based on the registration date of a case in the district court and the month of Kanoon rollout of the state's high court. In the bottom row, the year and month of the district court case is based on the decision date of the case in the district court. $Kanoon_{-1c}$ (the dummy variable which is equal to one for the 12 months before Kanoon rollout) is omitted and estimates should be interpreted relative to this baseline.

Results in Figure 7 suggest that the number of appeals and the number of decisions related to these appeals is steadily rising over the years after the launch of Indian Kanoon. Specifically, the number of appeals doubles by the third year of the launch of Indian Kanoon when we define the start of the process as the registration date in the lower court. The effect size, an increase of 0.5 percentage points, is considerable considering that only 1% of all cases are appealed at the high courts in our sample (Table 2). This effect is more modest when we use the decision date at the lower court. The difference in the results based on the use of registration dates and decision dates is likely driven by the increase in the number of appeals filed after the release of the platform.

Next we present the results for the dependent variables that measure the outcomes of the appeal. We measure these in conditional percentages (Figure 8).[22] Results with the counts are presented in the Appendix (Figure A7). Three results are particularly noteworthy. First, in the first three years that follow the launch of Kanoon, there is a statistically significant decrease in the percentage of admitted appeals, after which it appears to revert back to the original levels. Second, there is also a statistically significant increase, about 2 to 4 percentage points in magnitude, in the percent of disposed appeal cases that were withdrawn. This effect peaks a year after the rollout but remains higher than the pre-rollout levels throughout the study period. Finally, in Figure 8 we also see a slight decrease in cases being dismissed, though this effect does not become statistically significant till the fifth year of the rollout. This reduction of dismissals is noteworthy. It suggests that the better availability of legal information may have not only increased the number of appeals, but also their quality, for they are more likely to be heard within the court and result in an overturning of the lower court decision.

These results suggest that the availability of free legal information via Kanoon may improve the capacity of appellants to understand and interpret the proceedings of the court, conduct their own legal research, and file appeals. There effort however, may not be sufficient to ensure that the appeals are actually admitted in the courts. Conditional on being admitted however, litigants do appear to decide against the continuation of the legal process more than they did before they had access to the Kanoon platform. Finally, Kanoon is associated with a higher likelihood of cases from the lower courts being

[22] Conditional in the sense that the denominator for each case status is the last node in Figure A12. For the variables % Admitted and % Rejected, the denominator is the number of decided appeal cases from that state-year-month. For % Disposed, the denominator is the number of admitted appeals. For the three variables % Withdrawn, % Allowed and % Dismissed, it is the number of disposed cases.

overturned.

We also conducted the placebo tests for all the event study estimations above as suggested by De Chaisemartin and d'Haultfoeuille (2020). Results are presented in Figures A8, A9 and A10. The results suggest the absence of the pre-trends for all the outcome variables.

## 6.4 Impacts on Firms

Next we explore the impact of Kanoon on Indian firms. We linked firm-level data to the legal data by matching the eCourt case level data to the firms in the Prowess database. We restricted our sample to firms for which we have financial information for all years 2006-2015 and, therefore, abstract from entry and exit decisions of firms. Several financial parameters might be impacted by the decline in the cost of legal information to conduct the event study analysis; assets, income, legal charges, audit fees, bank guarantees and bad debts. Financial variables are measured for financial years and are standardized using an inverse hyperbolic sine transformation.[23] We define as $j = 0$ the first full financial year after the Kanoon rollout of the high court in a firm's jurisdiction. To address the econometric concerns raised earlier, we once again run the event studies using the corrections from Sun and Abraham (2021). The estimation controls for firm and year fixed effects. Standard errors are clustered at the state level.

We run our analysis on two different samples of the matched data. First we estimate the impact of Kanoon on all firms, regardless of whether they had any cases within our considered time frame. We refer to this as a "general equilibrium" effect.[24] Next we restrict our analysis to firms having at least one case at the courts during the 2006-2015 time period. We refer to this as the "partial equilibrium" analysis. Figure 9 presents both general and partial-equilibrium results.

Results of our estimations are presented in Figure A13. We present the coefficients that correspond to the estimated $\beta_j$s ($Kanoon_{jc}$ in Equation 2). These are defined relative to the first full financial year after the Kanoon rollout of the state's high court. $Kanoon_{-1c}$ (the dummy variable which is equal to one if the Kanoon rollout falls into the financial year) is omitted and estimates should be interpreted

[23] A financial year is from April 1st to March 31st of the following year.

[24] Labelling these regressions as general equilibrium is a slight but common misuse of language, as we do not allow for entry or exit of firms. We think the label is still informative, as we want to see the effect on firms which are not directly involved in legal charges and therefore impacted by Kanoon.

relative to this baseline.

Note that in both general and partial equilibrium settings, the Kanoon rollout had large and significant effects on firms' financial variables. We observe similar effects in magnitude in both settings for assets, legal charges, bank guarantees and bad debts. The effect for legal charges grows over time, during the first financial year after Kanoon rollout legal charges increase about 10% and five years later even by around 80%. The effect on assets is steadily increasing over the first three years after Kanoon rollout before levelling off at an longer term effect between 20% and 40%.[25] Bank guarantees and bad debts, on the other hand, see a sharp and persisting decline after the Kanoon rollout.

For firms' income, we observe a much larger negative effect for the partial equilibrium results (in Figure 9 these results are represented as squares) compared with the general equilibrium results (in Figure 9 these results are represented as circles), where, again, the effect is increasing in magnitude over time. On the other hand, for audit fees, the partial equilibrium effect are close to zero in all post-rollout periods. However, when including all firms in the regression, audit fees decline significantly from the third year after Kanoon rollout on.

In Appendix Figure A13, we present as additional results the estimation of the impact on court-level efficiency measures for cases involving firms from the Prowess data. We also present placebo tests for the partial and general equilibrium estimations and the court-level efficency estimates (see Appendix Tables A14, A15, and A16). These results strengthen our confidence in the absence of pre-trends.

# 7 Discussion

Overall, these results suggest that the reduction of search costs through the introduction of free legal information had a positive impact on the balance sheet of firms. Case-level analysis suggests that individual cases were more likely to be resolved immediately after the rollout. Though there were not statistically significant impacts on the overall measures of efficiency at the high courts, there is also a pronounced increase in the number of filings at the district courts. We also see a significant impact in the number of appeals, both in absolute numbers and the percentage of cases that are appealed for at

[25] To get the %-effect of an coefficient estimated as $x$ we calculate $(e^x - 1) \times 100$.

least the cohort of cases that we construct using the registration date. We also see that fewer appeals are admitted and more appeals are withdrawn during the proceedings. Finally, we see significant impacts on the financial status of firms. The launch of Kanoon increases assets and legal charges of firms. It decreases income, audit fees, bank guarantees and bad debts.

These results are consistent with previous literature. Bloom et al. (2013) demonstrate that firms operating in India face a complex economic, regulatory, and legal landscape for doing business. Bertrand and Crépon (2021) find that providing firms in South Africa with information about labor regulation via newsletters and access to a specialized website resulted in a 12% increase in employment in just six months.

While our event-study framework does not provide insights into the causal mechanisms at play, we conjecture that free legal search impacted Indians through several channels. First, at the level of cases, impacts are likely to be different for cases that were were *already* filed prior to the launch of the platform than the cases that were filed for the first time after the platform was available. Cases in the former category will be immediately impacted since litigants who gain new legal information may update their beliefs and expectations, and choose to either withdraw their case or to settle out of court (See for instance Woodruff, Sadka, and Seira (2020) for how more accurate information decreases the length of cases in court and increases their rate of settlements). This would increase the speed of resolution for the cases that are already in the system.

For those who are considering the submission of new cases after the launch of the platform however, access to legal information can change the set of cases present in the court. The availability of legal information can empower new litigants who were previously unaware of their rights under the law to file cases. It may also help litigants improve the quality of their paperwork or legal arguments. Such newly filed cases may have a higher likelihood of being resolved during the first year. Even if Kanoon does not affect the set of cases, free legal information may help litigants better prepare for hearings, anticipate counter-arguments and thus resolve their legal issues faster (potentially via the channel of withdrawals and settlements). While we are unable to draw any direct inference on whether the platform changed the mix of cases at the court, it is noteworthy that we observe no change in the likelihood of the Indian state being a petitioner or a respondent in the cases after the Kanoon platform

was launched (Table 3). There may however, still be a shift in the types of cases filed between private parties at the courts. This will be a critical area of further research.

Second, we interpret the results on court efficiency. There is no statistically significant impact on the efficiency of the high courts, yet one of the most significant results in this paper is the impact of Kanoon on the filing of appeals at the high courts. We noted an increase in the number of filings, in number of cases as well as percentages of cases that are appeals that was constructed using the registration date in the court system. Moreover, we also see a significant decrease in the share of admitted cases and an increase in the share of withdrawals of disposed appeals. These results suggest that the effect may be driven by the litigants: greater access to legal information affects their decisions to file appeals but at the cost of less well prepared appeals which do not pass the administrative stage. However, litigants seem to be able to learn from Kanoon and to withdraw a case if they feel that the proceedings are not going in their favor. In other words, Kanoon may have enabled the stakeholders to make better arguments and better argue their cases in India's common law system but does not help them understand the administrative complexity of the system, an issue that has been previously highlighted in the literature on India (Chemin 2009; Chemin 2012; Goldberg et al. 2010).

Third, to interpret the impacts on efficiency of the district courts, here we must keep in mind that the cases from the lower courts are not included in the Kanoon database. What is important here is that legal information became available to stakeholders in these courts - all decisions from the high courts are binding on the lower courts in India's justice system. Access to Kanoon likely provided all parties contesting a case with better access to information. This likely initially increased the time needed to settle a case, resulting in more pending cases and greater backlog in the system. Eventually, however, greater efficiency led to an improvement in the clearance rate. Other explanations are also possible for these effects. For example, access to Kanoon may lead to overconfidence of the parties, which can decrease the number of out of court settlements, which consequently increases the number of open cases.

Lastly, we discuss the results on firms. We find that new access to legal information has a pronounced positive impact on the assets and legal charges of firms. It decreases income, audit fees, bank guarantees and bad debts. How may we explain these effects? Here we emphasize that the availability

of free legal information may have enabled people other than lawyers, such as managers, accountants and other stakeholders, to access critical legal information. Prior to the launch of Indian Kanoon, these individuals may not have had easy access to legal information and relied exclusively on lawyers for guidance on accounting practices. With the arrival of this database however, they were likely able to independently and efficiently access the full body of laws that affect their businesses.

Our attempts to examine the citation practices within cases (Appendix A6) does not suggest any significant influence of the Kanoon platform. A deeper analysis of the cases themselves – the types of cases, the mix of criminal and civil cases, and the quality of the legal arguments – is an important area for further research. A deeper analysis of the usage of the Kanoon platform itself is also an important area of inquiry.

# 8 Conclusion

Legal information can be difficult to access, expensive and technically challenging (United Nations 2016). Achieving equal justice for all requires policies to overcome these barriers. We examine the impact of Indian Kanoon – a free legal search engine – that was implemented in India in 2008 by an overseas graduate student. The engine had the explicit goal of bringing "the knowledge of law to the common people". Emphasis was placed on keyword searches and tight integration of court judgments with laws and with prior judgements to allow automatic determination of the most relevant clauses and court judgments.

We use an event study framework that exploits the staggered rollout of the platform across different states to assess its effects on cases, courts and firms. Our working sample consists of all cases that are present in the Kanoon database that are linked to official cases scraped from the publicly available e-Courts system. We estimate the two-way fixed effects model adjusting for the fact that the control group for the later-treated is already treated using the methods of Sun and Abraham (2021). We examine the impact on the Indian judiciary and firms.

We find that the launch of Kanoon is associated a 1–3% increase in the number of resolved cases. Though there were not statistically significant impacts on the overall measures of efficiency at the high

courts, the number of appeal cases from lower courts that are heard at the high court nearly double after the launch of Kanoon. We also see a 3–6% increase in the likelihood of an appeal case being withdrawn and a 3–6% decline in the likelihood of a case being dismissed after the rollout of Kanoon. There is also a pronounced increase in the number of filings at the district courts, even though the website only publishes cases from the high courts. Finally, we see significant impacts on the financial status of firms. The launch of Kanoon increases assets and legal charges of firms. It decreases income, audit fees, bank guarantees and bad debts.

We interpret these effects as evidence that increased access to legal information through the Kanoon platform had widespread effects in India. Specifically, it reduced bottlenecks, increased efficiency and also contributed to improvements in the quality of legal research in the courts. This case study has relevance far beyond India – it suggests that improving access to legal information for all stakeholders in justice system can have broad impacts on markets and society at large.

# References


Acemoglu, Daron et al. (2020). "Trust in state and nonstate actors: Evidence from dispute resolution in Pakistan". In: *Journal of Political Economy* 128.8, pp. 3090–3147.

Baxi, Upendra (1982). *The Crisis of the Indian Legal System. Alternatives in Development: Law*. Stranger Journalism.

Bertrand, Marianne and Bruno Crépon (2021). "Teaching labor laws: Evidence from a randomized control trial in South Africa". In: *American Economic Journal: Applied Economics* 13.4, pp. 125–49.

Bloom, Nicholas et al. (2013). "Does management matter? Evidence from India". In: *The Quarterly journal of economics* 128.1, pp. 1–51.

Borusyak, Kirill and Xavier Jaravel (2017). "Revisiting event study designs". In: *Available at SSRN 2826228*.

Callaway, Brantly and Pedro HC Sant'Anna (2020). "Difference-in-differences with multiple time periods". In: *Journal of Econometrics*.

Chemin, Matthieu (2009). "Do judiciaries matter for development? Evidence from India". In: *Journal of Comparative Economics* 37.2. Special Symposium in Honor of Padma Desai, pp. 230–250. ISSN: 0147-5967.

— (2012). "Does court speed shape economic activity? Evidence from a court reform in India". In: *The Journal of Law, Economics, & Organization* 28.3, pp. 460–485.

— (2020). "Judicial efficiency and firm productivity: Evidence from a world database of judicial reforms". In: *Review of Economics and Statistics* 102.1, pp. 49–64.

De Chaisemartin, Clément and Xavier d'Haultfoeuille (2020). "Two-way fixed effects estimators with heterogeneous treatment effects". In: *American Economic Review* 110.9, pp. 2964–96.

Diamond, Peter A (1971). "A model of price adjustment". In: *Journal of economic theory* 3.2, pp. 156–168.
Djankov, Simeon et al. (2003). "Courts". In: *The Quarterly Journal of Economics* 118.2, pp. 453–517.
Galanter, Marc (1963). "Law and caste in modern India". In: *Asian Survey*, pp. 544–559.
Goldberg, Pinelopi Koujianou et al. (2010). "Imported intermediate inputs and domestic product growth: Evidence from India". In: *The Quarterly journal of economics* 125.4, pp. 1727–1767.
Goldfarb, Avi and Catherine Tucker (2019). "Digital economics". In: *Journal of Economic Literature* 57.1, pp. 3–43.
Goodman-Bacon, Andrew (2021). "Difference-in-differences with variation in treatment timing". In: *Journal of Econometrics*.
Greenleaf, Graham et al. (2013). "Challenges for Free Access to Law in a Multi-Jurisdictional Developing Country: Building the Legal Information Institute of India". In: *Journal of National Law University Delhi* 1.1, pp. 63–93.
Iyengar, Prashant (2010). "Free Access to Law in India-Is it Here to Stay?" In: *Available at SSRN 1778820*.
Jois, Rama (2004). *Legal and Constitutional History of India: Ancient, Judicial and Constitutional System*. Universal Law Publishing.
Kondylis, Florence and Mattea Stein (2018). "The speed of justice". In: *The Review of Economics and Statistics*, pp. 1–46.
Krishnan, Jayanth K et al. (2014). "Grappling at the Grassroots: Access to justice in India's Lower Tier". In: *Harv. Hum. Rts. J.* 27, p. 151.
Krishnaswamy, Sudhir and Siddharth Swaminathan (2019). "Public Trust in the Indian Judiciary: The Power to Transform". In: *A Qualified Hope: The Indian Supreme Court and Progressive Social Change*. Ed. by Gerald N. Rosenberg, Sudhir Krishnaswamy, and Shishir (Editors) Bail. Comparative Constitutional Law and Policy. Cambridge University Press, pp. 123–140.
Lichand, Guilherme and Rodrigo R Soares (2014). "Access to justice and entrepreneurship: Evidence from Brazil's special civil tribunals". In: *The Journal of Law and Economics* 57.2, pp. 459–499.
Ponticelli, Jacopo and Leonardo S Alencar (2016). "Court enforcement, bank loans, and firm investment: evidence from a bankruptcy reform in Brazil". In: *The Quarterly Journal of Economics* 131.3, pp. 1365–1413.
Rao, Manaswini (2021). *Courts Redux: Micro-evidence from India*. Tech. rep. Working Paper.
Stigler, George J (1961). "The economics of information". In: *Journal of political economy* 69.3, pp. 213–225.
Sun, Liyang and Sarah Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". In: *Journal of Econometrics* 225.2, pp. 175–199.
United Nations (Oct. 2016). *Global Study on Legal Aid, Global Report*. Tech. rep. United Nations.
Varian, Hal R (1980). "A model of sales". In: *The American economic review* 70.4, pp. 651–659.
Visaria, Sujata (2009). "Legal reform and loan repayment: The microeconomic impact of debt recovery tribunals in India". In: *American Economic Journal: Applied Economics* 1.3, pp. 59–81.
Woodruff, Christopher, Joyce Sadka, and Enrique Seira (2020). "Information and Bargaining through Agents: Experimental Evidence from Mexico's Labor Courts". In:

# 9 Figures and Tables

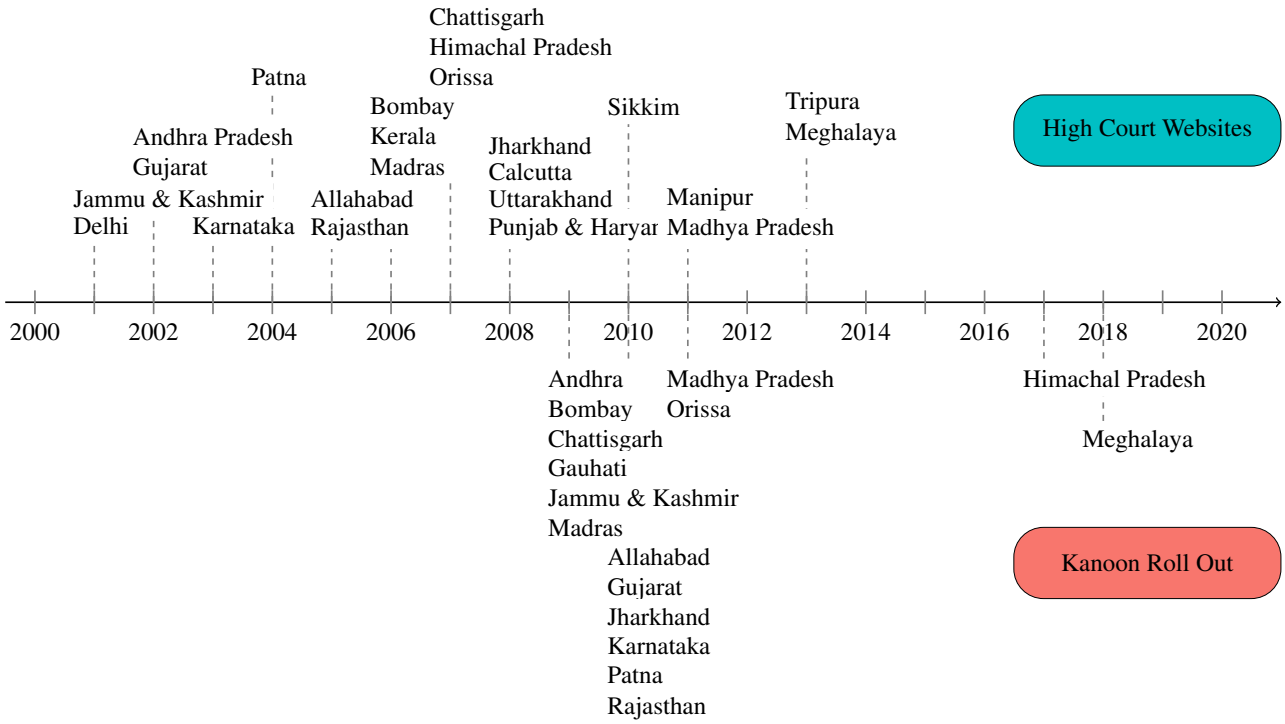


Figure 1: Roll Out Years for High Court Websites (top) and Kanoon (bottom)

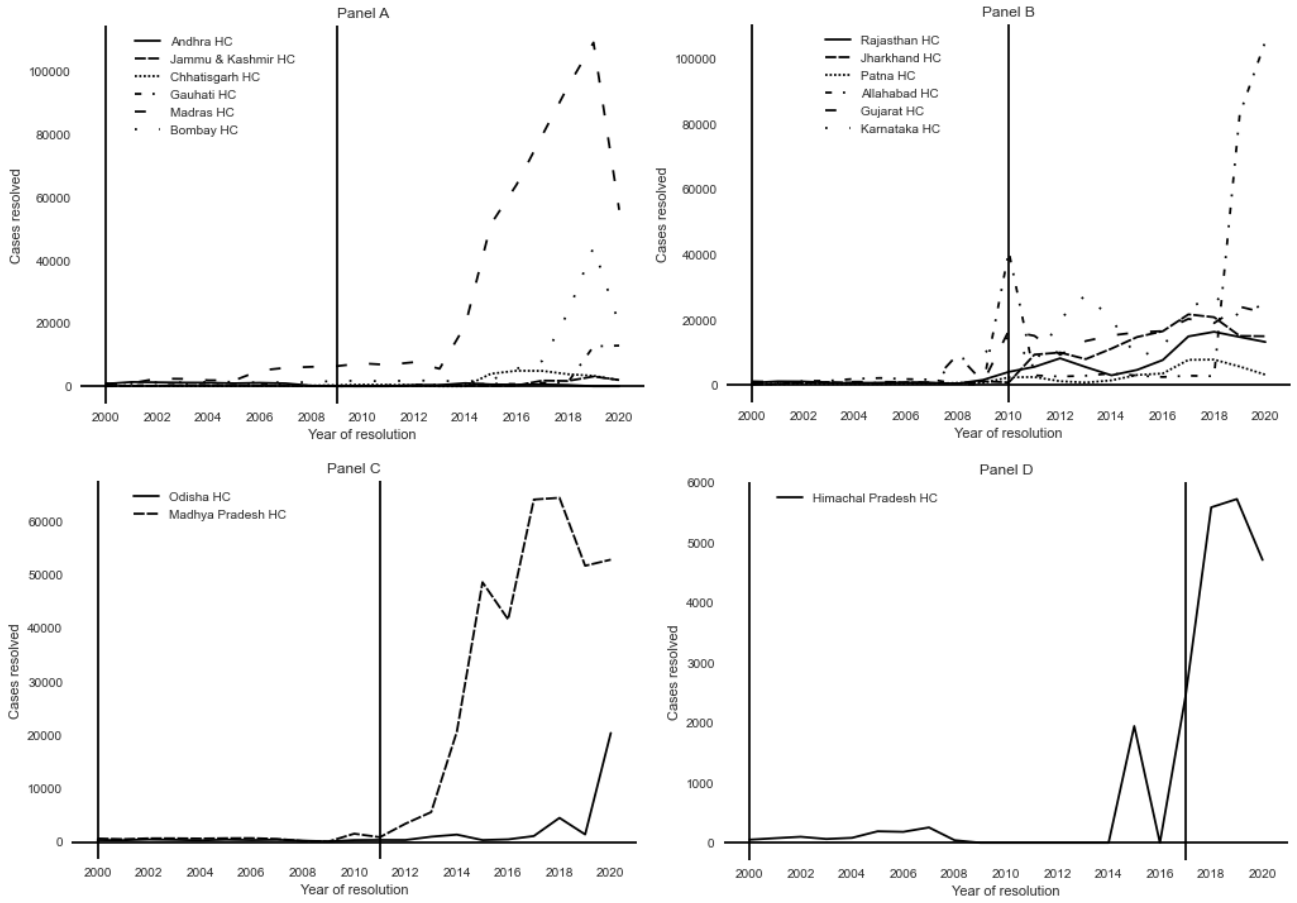


Figure 2: Number of Documents on Indian Kanoon per Publication Year and High Court

*Notes:* This figure counts the number of available documents (Judgements, Orders, Laws) available on Indian Kanoon in March 2021 by the year of the documents' publication dates of the document and by high court. Each panel assembles all high courts rolled out on Kanoon in the same year, which is represented by a solid vertical line.

Table 1: Summary Statistics

| | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| *eCourt High Courts Data (Case Level)* | | | | | |
| Resolved Cases | 11,894,096 | 0.76 | 0.43 | 0 | 1 |
| Resolved within 1 year | 11,894,096 | 0.50 | 0.50 | 0 | 1 |
| *eCourt High Courts Data (State-Month Level)* | | | | | |
| Number of Filings (Thsd.) | 5,006 | 2.76 | 3.43 | 0.00 | 54.84 |
| Decided cases (Thsd.) | 4,853 | 2.14 | 3.76 | 0.00 | 182.75 |
| Pending cases (Thsd.) | 5,292 | 452.14 | 473.52 | 0.20 | 2,602.00 |
| Backlog cases (Thsd.) | 4,691 | 28.68 | 48.42 | 0.00 | 438.30 |
| Mean age of Pending cases (Years) | 4,979 | 0.04 | 0.01 | -0.00 | 0.08 |
| Mean age of Decided cases (Years) | 4,853 | 1.37 | 1.28 | -12.00 | 10.50 |
| Disposition time (Years) | 4,853 | 16.74 | 99.31 | 0.01 | 1,455.06 |
| Clearance rate | 4,853 | 0.01 | 0.10 | 0.00 | 5.83 |
| *eCourt District Courts Data (State-Month Level)* | | | | | |
| Number of Filings (Thsd.) | 6,500 | 11.92 | 21.27 | 0.00 | 217.72 |
| Decided cases (Thsd.) | 5,259 | 10.96 | 22.17 | 0.00 | 354.24 |
| Pending cases (Thsd.) | 6,612 | 396.59 | 714.21 | 0.00 | 5,286.06 |
| Backlog cases (Thsd.) | 6,612 | 301.64 | 568.18 | 0.00 | 4,123.98 |
| Mean age of Pending cases (Years) | 6,612 | 3.78 | 1.44 | 0.60 | 7.85 |
| Mean age of Decided cases (Years) | 5,259 | 2.13 | 1.48 | 0.00 | 21.70 |
| Disposition time (Years) | 5,259 | 22.73 | 99.03 | 0.00 | 2,541.55 |
| Clearance rate | 5,259 | 0.53 | 0.51 | 0.00 | 7.76 |

Table 2: Summary Statistics for Appeal Cases

| | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| *Appeals of DC Cases* | | | | | |
| Registrations (Thsd.) | 2,808 | 16.55 | 22.32 | .01 | 115 |
| Appeals (by Reg. Date) | 2,808 | 89.13 | 176.82 | 0 | 1,129 |
| % Appealed (by Reg. Date) | 2,808 | 0.01 | 0.01 | 0 | .074 |
| Decisions (Thsd.) | 2,808 | 12.95 | 23.59 | 0 | 355 |
| Appeals (by Dec. Date) | 2,808 | 92.17 | 203.74 | 0 | 1,377 |
| % Appealed (by Dec. Date) | 2,453 | 0.01 | 0.03 | 0 | 1 |
| *Case Status of DC Appeals (by Reg. Date)* | | | | | |
| Rejected | 2,808 | 2.59 | 8.38 | 0 | 95 |
| Withdrawn | 2,808 | 2.11 | 6.12 | 0 | 59 |
| Allowed | 2,808 | 44.30 | 97.92 | 0 | 706 |
| Dismissed | 2,808 | 15.45 | 31.76 | 0 | 204 |
| Disposed | 2,808 | 14.29 | 38.86 | 0 | 335 |
| % Rejected | 1,946 | 0.04 | 0.08 | 0 | 1 |
| % Withdrawn | 1,946 | 0.08 | 0.13 | 0 | 1 |
| % Allowed | 1,946 | 0.90 | 0.15 | 0 | 1 |
| % Dismissed | 1,946 | 0.35 | 0.25 | 0 | 1 |
| % Disposed | 1,946 | 0.30 | 0.29 | 0 | 1 |
| *Case Status of DC Appeals (by Dec. Date)* | | | | | |
| Rejected | 2,808 | 0.52 | 2.45 | 0 | 29 |
| Withdrawn | 2,808 | 0.19 | 1.42 | 0 | 44 |
| Allowed | 2,808 | 8.87 | 39.11 | 0 | 372 |
| Dismissed | 2,808 | 4.19 | 18.29 | 0 | 165 |
| Disposed | 2,808 | 2.10 | 9.74 | 0 | 94 |
| % Rejected | 374 | 0.05 | 0.08 | 0 | 1 |
| % Withdrawn | 374 | 0.04 | 0.11 | 0 | 1 |
| % Allowed | 374 | 0.93 | 0.11 | 0 | 1 |
| % Dismissed | 374 | 0.44 | 0.28 | 0 | 1 |
| % Disposed | 374 | 0.18 | 0.20 | 0 | 1 |

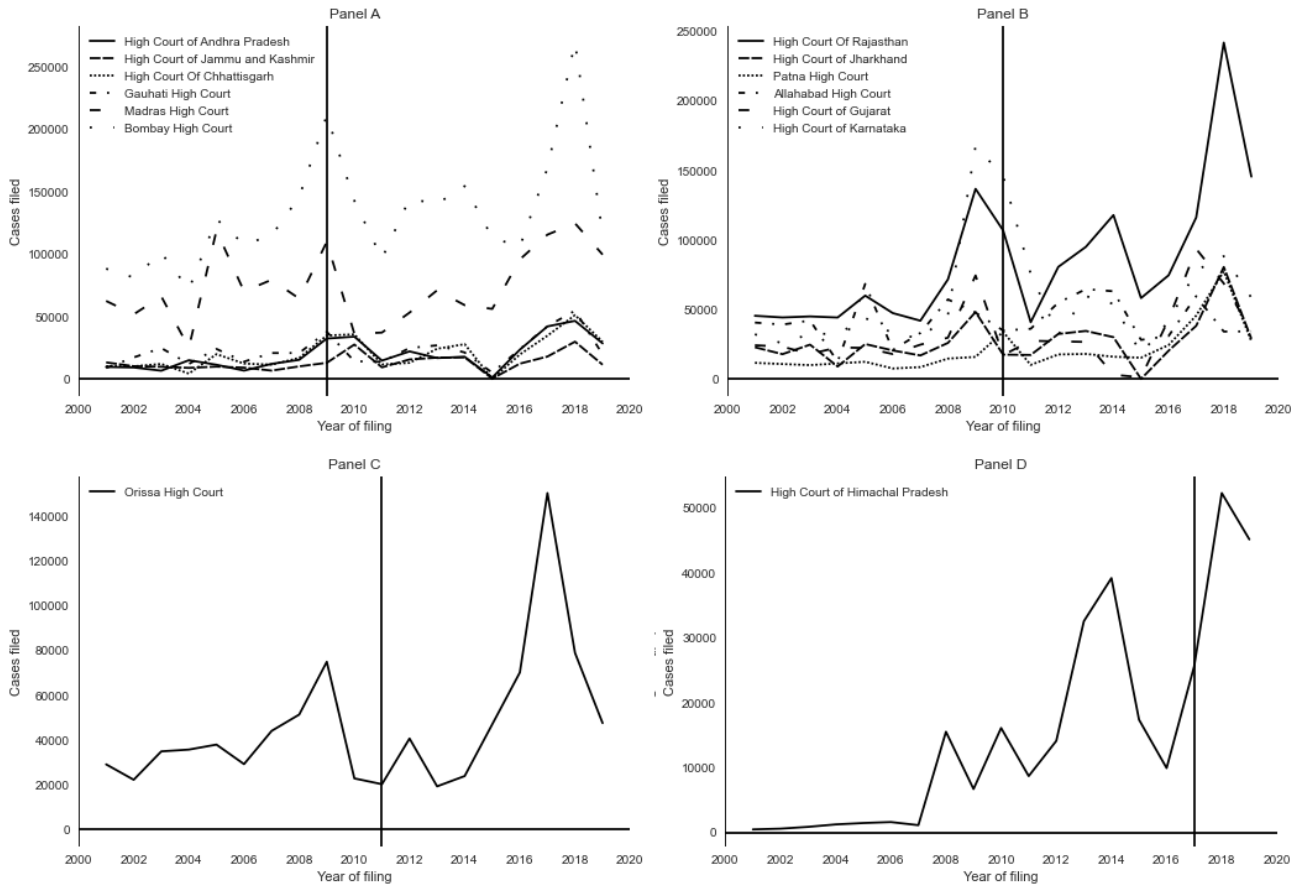


Figure 3: Number of Filings in High Courts per Year and High Court

*Notes:* This figure counts the number of filings per high court and year of filing for cases available on eCourts High Court in November 2020. Each panel assembles all high courts rolled out on Kanoon in the same year, which is represented by a solid vertical line.

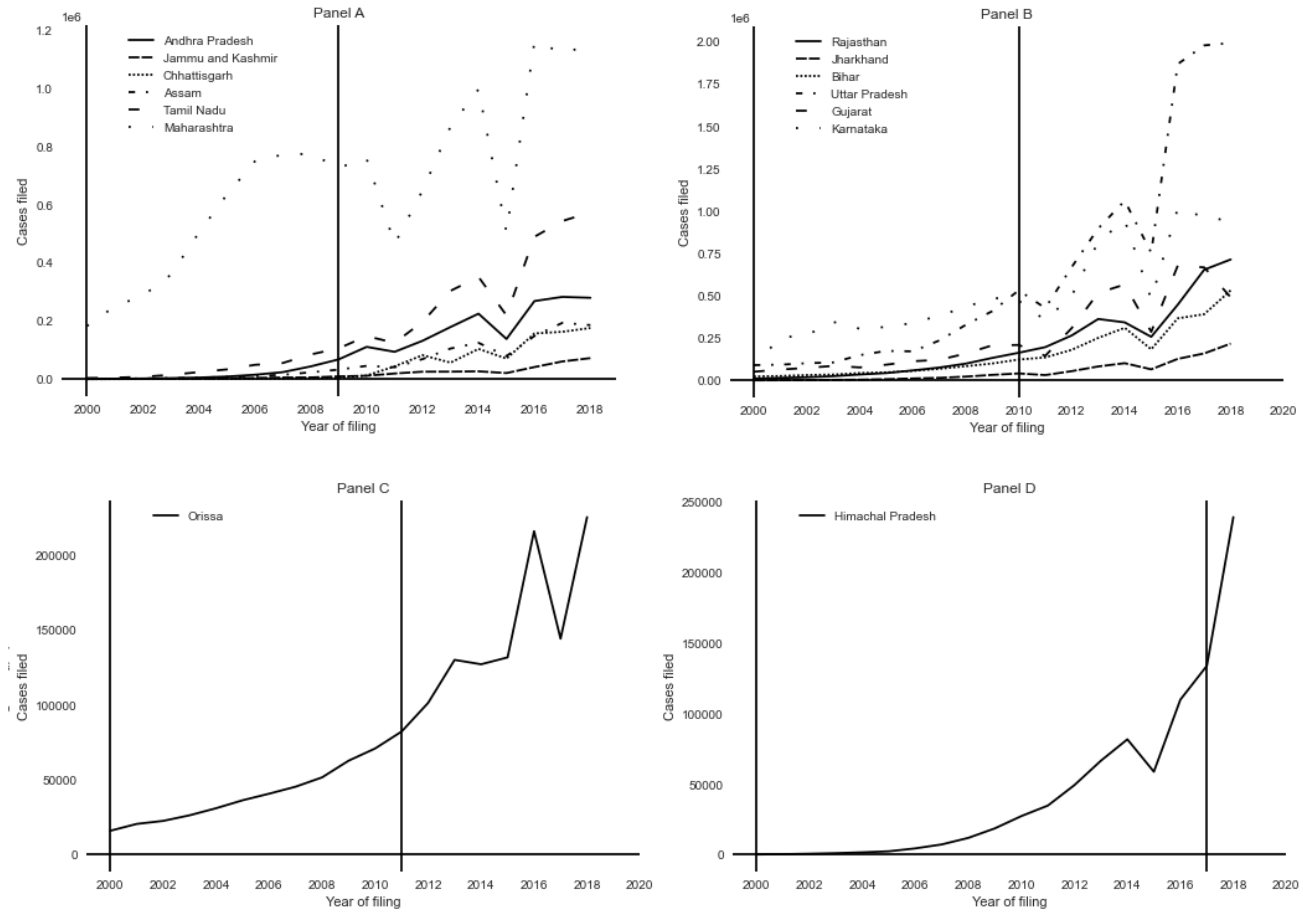


Figure 4: Number of Filings in District Courts by Year and State

*Notes:* This figure counts the number of filings in district courts per state and year of filing for cases available on eCourts District Court in May 2019. Each panel assembles all states which high courts were rolled out on Kanoon in the same year, which is represented by a solid vertical line.

Table 3: Effects on High Court Cases' Characteristics and Outcomes

| | Resolved | | | | Resolved $<$ 1 Year | Government |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Year 3 before Rollout | 0.15*** | 0.0010 | -0.0029 | -0.0083 | 0.0064 | -0.018 |
| | (0.0281) | (0.0187) | (0.00712) | (0.00768) | (0.0160) | (0.0139) |
| Year 2 before Rollout | 0.14*** | -0.00092 | -0.0075 | -0.012 | 0.0043 | -0.00066 |
| | (0.0251) | (0.0123) | (0.00669) | (0.00745) | (0.0125) | (0.00589) |
| Year of Rollout | 0.13*** | 0.026 | 0.0064 | 0.0092 | 0.022* | 0.010 |
| | (0.0311) | (0.0218) | (0.00689) | (0.00597) | (0.0117) | (0.0142) |
| Year 1 after Rollout | 0.094*** | 0.022 | 0.014*** | 0.017*** | 0.012 | 0.0055 |
| | (0.0250) | (0.0246) | (0.00408) | (0.00423) | (0.0213) | (0.0181) |
| Year 2 after Rollout | 0.048 | 0.0066 | 0.0048 | 0.0081 | -0.028 | 0.0091 |
| | (0.0492) | (0.0448) | (0.0215) | (0.0197) | (0.0267) | (0.0191) |
| Year 3 after Rollout | 0.063 | 0.033 | 0.023** | 0.028*** | -0.0096 | 0.00059 |
| | (0.0362) | (0.0352) | (0.00889) | (0.00769) | (0.0241) | (0.0136) |
| Court FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | | Yes | Yes | Yes | Yes | Yes |
| Court x Year FE | | | Yes | Yes | Yes | Yes |
| Month FE | | | | Yes | Yes | Yes |
| Adj. R2 | 0.06 | 0.15 | 0.18 | 0.18 | 0.07 | 0.11 |
| N | 12,791,187 | 12,791,187 | 12,791,187 | 12,791,187 | 12,791,187 | 12,791,187 |

Note: This Table presents estimation results of $\beta_j$ in Equation 1. The first three columns build up to the final specification, which includes Court, Year, Court-Year and Month fixed effects. The regressions are run on all cases available on eCourts High Courts in November 2020. "Resolved" is a dummy variable equal to one if the case is resolved any time before the collection of our data. The variable "Resolved $<$ 1 Year" is equal to one if the case is resolved in less than one year after its filing at the high court. "Government" equals one if the petitioner or the respondent is a government organization. The explanatory variables ($Kanoon_{jct}$ in Equation 1) are defined by the difference in days between a case's filing date and the exact Kanoon rollout date of the respective high court. "Year 1 before Rollout" (e.g. the 365 days before Kanoon rollout) is omitted and estimates should be interpreted relative to this baseline. Standard errors are clustered at the high court level. Stars *, ** and *** indicate that the p-value is below 0.1, 0.05 and 0.01 respectively.

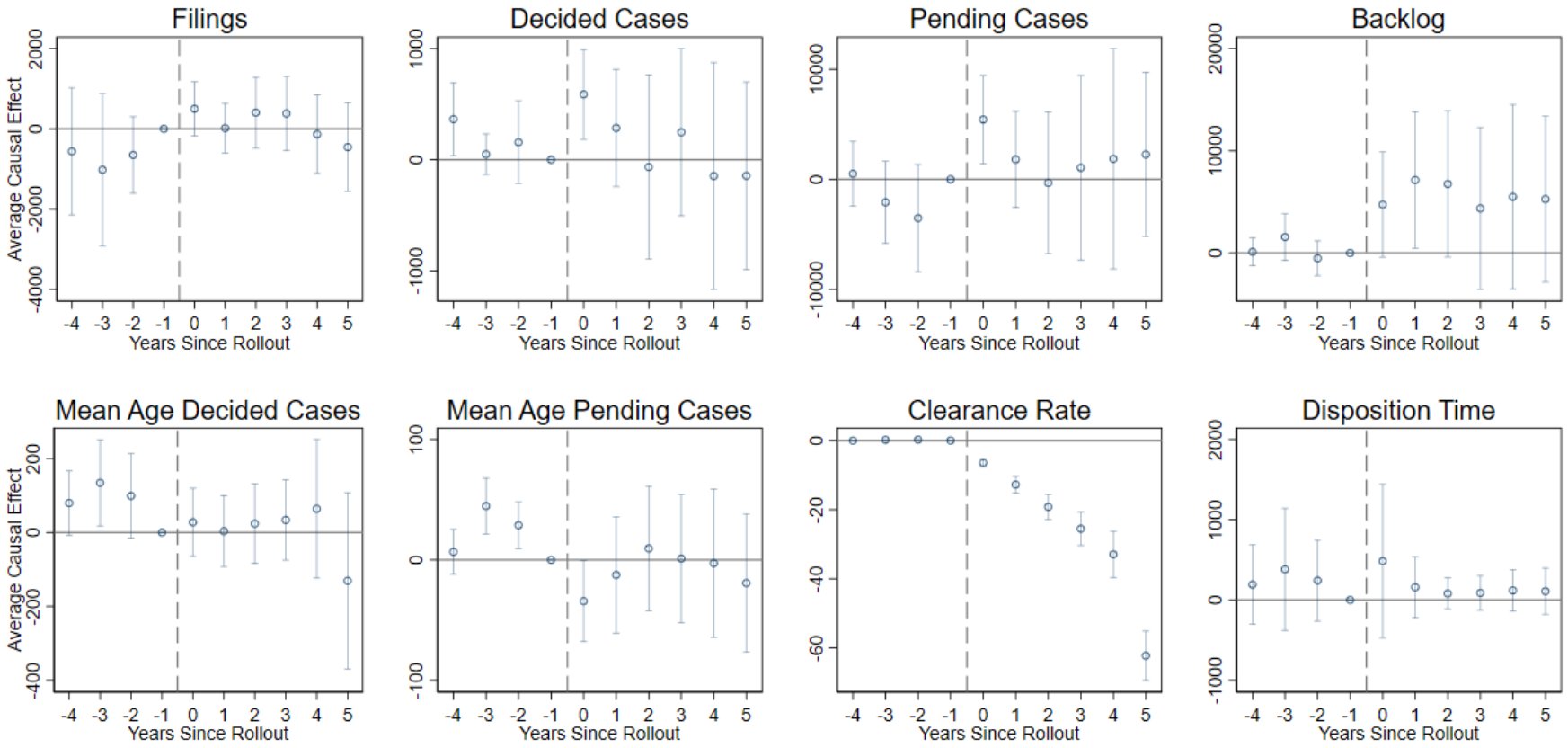


Figure 5: Impact of Kanoon Rollout on Aggregate Efficiency Measures of High Courts

*Notes:* This figure presents the estimated effects Kanoon had on aggregate high court efficiency. It displays estimated coefficients and 95% confidence intervals for $\beta_j$s from Equation 2, with a different outcome variable for each panel. Efficiency measures are calculated at the high court-year-month level using eCourts High Courts data for the years 2005-2016. The variables corresponding to the estimated $\beta_j$s ($Kanoon_{jc}$ in Equation 2) are defined by the difference in months between a high court-year-month observation and the month of Kanoon rollout of the respective high court. $Kanoon_{-1c}$ (the dummy variable which is equal to one for the 12 months before Kanoon rollout) is omitted and estimates should be interpreted relative to this baseline. The estimation uses the algorithm of Sun and Abraham (2021), where Himachal Pradesh (rollout 04/2017) is defined as control cohort. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the high court level.

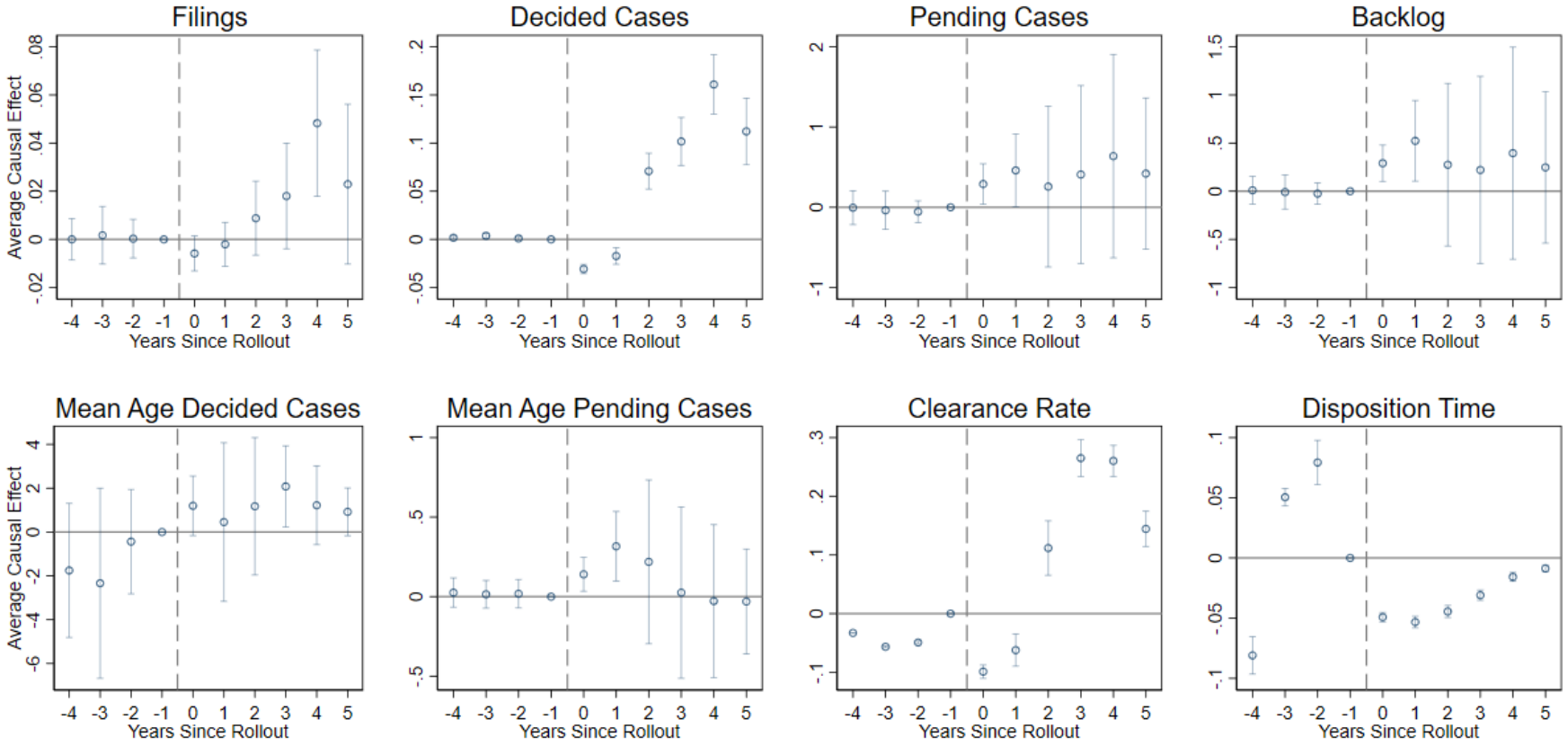


Figure 6: Impact of Kanoon Rollout on Aggregate Efficiency Measures of District Courts

*Notes:* This figure presents the estimated effects Kanoon had on district court efficiency. It displays estimated coefficients and 95% confidence intervals for $\beta_j$s from a slight modification of Equation 2. Efficiency measures are calculated at the state-year-month level using eCourts District Court data for the years 2005-2016. Each panel presents results for a different efficiency measure as outcome variable. The variables corresponding to the estimated $\beta_j$s ($Kanoon_{jc}$ in Equation 2) are defined by the difference in months between a state-year-month observation and the month of Kanoon rollout of the state's high court. $Kanoon_{-1c}$ (the dummy variable which is equal to one for the 12 months before Kanoon rollout) is omitted and estimates should be interpreted relative to this baseline. The estimation uses the algorithm of Sun and Abraham (2021), where Himachal Pradesh (rollout 04/2017) is defined as control cohort. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

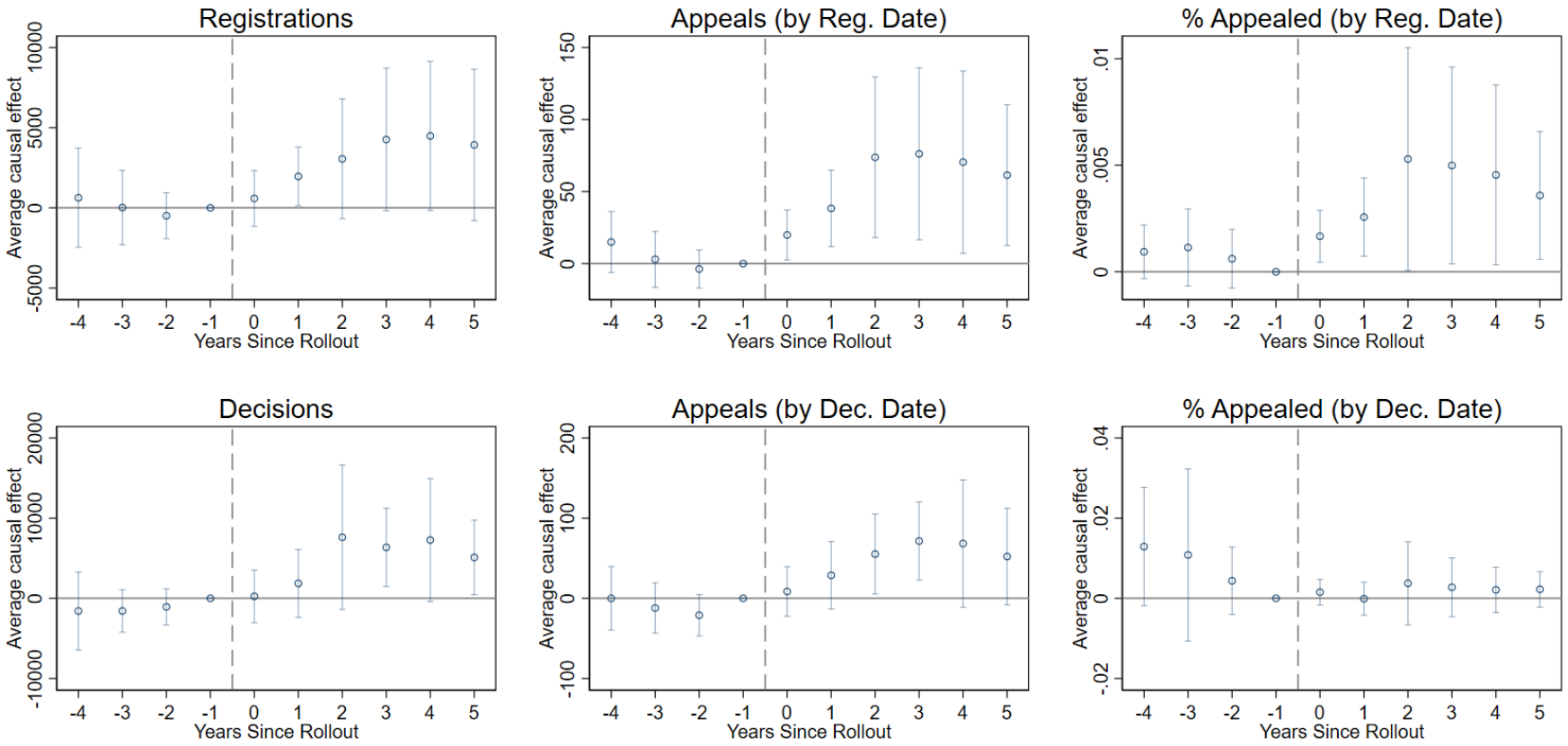


Figure 7: Impact of Kanoon on Appeals of District Court Cases in High Courts

*Notes:* This figure presents the estimated effects Kanoon had on the likelihood that a district court case is appealed in the high court. It displays estimated coefficients and 95% confidence intervals for $\beta_j$s from a slight modification of Equation 2. Efficiency measures are calculated at the state-year-month level using eCourts District Court data for the years 2005-2016. Each panel presents results for a different efficiency measure as outcome variable. The variables corresponding to the estimated $\beta_j$s ($Kanoon_{jc}$ in Equation 2) are defined differently for the top and the bottom row. For the top row, it is defined by the difference in months between a state-year-month observation based on the registration date of a case in the district court and the month of Kanoon rollout of the state's high court. In the bottom row, the year and month of the district court case is based on the decision date of the case in the district court. $Kanoon_{-1c}$ (the dummy variable which is equal to one for the 12 months before Kanoon rollout) is omitted and estimates should be interpreted relative to this baseline. The estimation uses the algorithm of Sun and Abraham (2021), where Himachal Pradesh (rollout 04/2017) is defined as control cohort. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

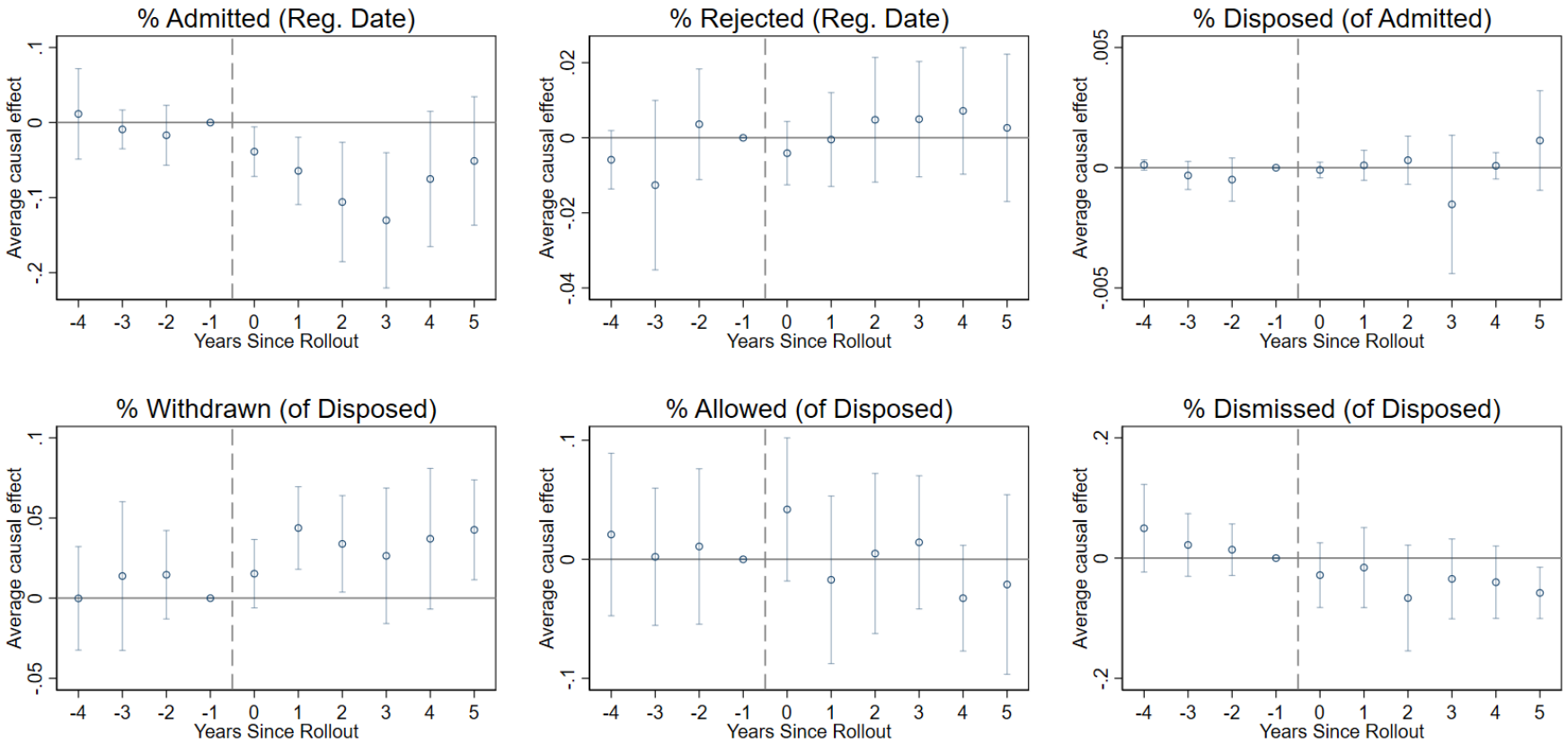


Figure 8: Impact of Kanoon on Appeal Outcomes of District Court Cases in High Courts (in %)

*Notes:* This figure presents the estimated effects Kanoon had on the likelihood that a district court case is appealed in the high court. It displays estimated coefficients and 95% confidence intervals for $\beta_j$s from a slight modification of Equation 2. Efficiency measures are calculated at the state-year-month level using eCourts District Court data for the years 2005-2016. Each panel presents results for a different efficiency measure as outcome variable. For % Admitted and % Rejected the denominator is the number of decided appeal cases from that state-year-month. For % Disposed the denominator is the number of admitted appeals and for % Withdrawn, % Allowed and % Dismissed it is the number of disposed cases. The variables corresponding to the estimated $\beta_j$s ($Kanoon_{jc}$ in Equation 2) are defined by the difference in months between a state-year-month observation based on the registration date of a case in the district court and the month of Kanoon rollout of the state's high court. $Kanoon_{-1c}$ (the dummy variable which is equal to one for the 12 months before Kanoon rollout) is omitted and estimates should be interpreted relative to this baseline. The estimation uses the algorithm of Sun and Abraham (2021), where Himachal Pradesh (rollout 04/2017) is defined as control cohort. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

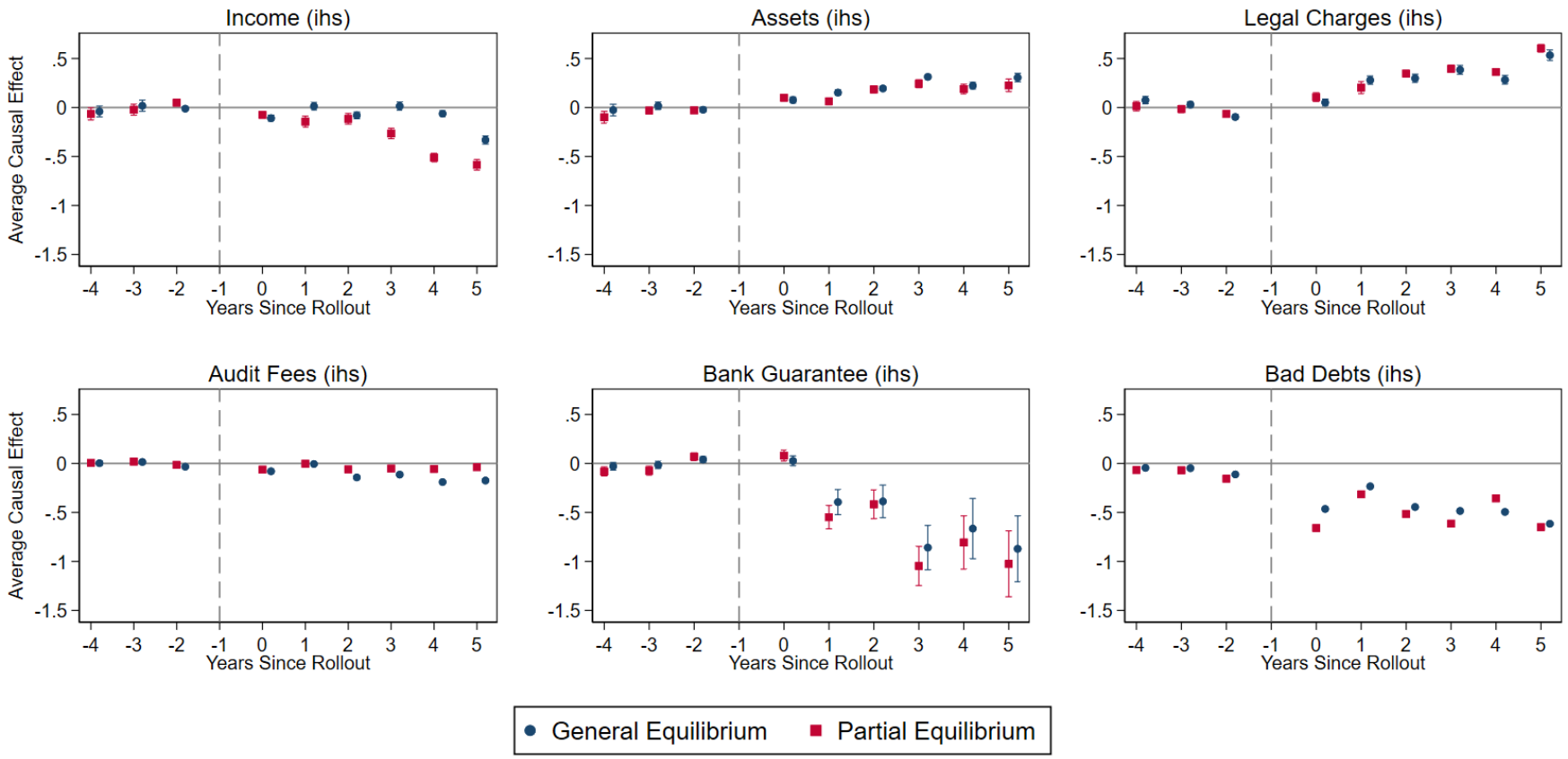


Figure 9: Effects of Kanoon rollout on Firm Financials

*Notes:* This figure shows the impact of Kanoon rollout on various firm outcomes for all firms (General Equilibrium) and firms having at least one case during the considered time frame (Partial Equilibrium). It displays estimated coefficients and 95% confidence intervals for $\beta_j$s from a slight modification of Equation 3. Firm outcomes are measured for financial years (April 1st to March 31st of the following year) and obtained from the Prowess data for the years 2006 - 2015. Each outcome variable is a measure of financial performance and is transformed with the inverse hyperbolic sine function. The variables corresponding to the estimated $\beta_j$s ($Kanoon_{jc}$ in Equation 2) are defined relative to the first full financial year after the Kanoon rollout of the state's high court. $Kanoon_{-1c}$ (the dummy variable which is equal to one if the Kanoon rollout falls into the financial year) is omitted and estimates should be interpreted relative to this baseline. The estimation uses the algorithm of Sun and Abraham (2021), where Himachal Pradesh (rollout 04/2017) is defined as control cohort. The estimation controls for firm and year fixed effects. Standard errors are clustered at the state level.

# Appendix

Table A1: State and HCs in India and Data Availability.

| High Court | Established | State | State Created | Population | Kanoon | eCourt HC | eCourt DC |
|---|---|---|---|---|---|---|---|
| Allahabad HC | 1866-03-17 | Uttar Pradesh | 1950-01-26 | 19,9812,341 | 2010-07-08 | 2001-2019 | 2000-2018 |
| Andhra Pradesh HC[a,b] | 2019-01-01 | Andhra Pradesh | 1956-11-01 | 4,9506,799 | 2009-01-18 | 2001-2019 | 2000-2018 |
| | | Telangana (until Jan 2019) | | | | | 2000-2018 |
| Bombay HC | 1862-08-14 | Goa | 1987-05-30 | 145,8545 | 2009-01-18 | x | 2000-2018 |
| | | Maharashtra | 1960-05-01 | 112,374,333 | | | 2000-2018 |
| | | Dadra & N. H. & D. & D. (UT) | 2020-01-26 | 586,956 | | | x |
| Calcutta HC | 1862-07-02 | West Bengal | 1950-01-26 | 91,276,115 | x | x | 2000-2018 |
| | | Andaman & Nicobar Islands (UT) | 1956-11-01 | 380,581 | | | x |
| Chhattisgarh HC | 2000-11-01 | Chhattisgarh | 2000-11-01 | 25,545,198 | 2009-01-18 | 2001-2019 | 2000-2018 |
| Delhi HC | 1966-10-31 | Delhi (UT) | 1956-11-01 | 16,787,941 | x | x | 2000-2018 |
| Gauhati HC[c] | 1948-03-01 | Arunachal Pradesh | 1987-02-20 | 1,383,727 | 2009-01-18 | 2001-2019 | x |
| | | Assam | 1950-01-26 | 31,205,576 | | | 2000-2018 |
| | | Mizoram | 1987-02-20 | 1,097,206 | | | 2000-2018 |
| | | Nagaland | 1963-12-01 | 1,978,502 | | | x |
| | | Meghalaya (until Mar 2013) | | | | | 2000-2018 |
| | | Manipur (until Mar 2013) | | | | | 2000-2018 |
| | | Tripura (until Mar 2013) | | | | | 2000-2018 |
| Gujarat HC | 1960-05-01 | Gujarat | 1960-05-01 | 60,439,692 | 2010-06-13 | 2001-2019 | 2000-2018 |
| Himachal Pradesh HC | 1971-01-25 | Himachal Pradesh | 1971-01-25 | 6,864,602 | 2017-04-15 | 2001-2019 | 2000-2018 |
| Jammu & K. and L. HC | 1928-03-26 | Jammu and Kashmir (UT) | 2019-10-31 | 12,258,433 | 2009-01-18 | 2001-2019 | 2000-2018 |
| | | Ladakh (UT) | 2019-10-31 | 290,492 | | | x |
| Jharkhand HC | 2000-11-15 | Jharkhand | 2000-11-15 | 32,988,134 | 2010-06-13 | 2001-2019 | 2000-2018 |
| Karnataka HC | 1905-02-26 | Karnataka | 1956-11-01 | 61,095,297 | 2010-06-13 | 2001-2019 | 2000-2018 |
| Kerala HC | 1956-11-01 | Kerala | 1956-11-01 | 33,406,061 | x | 2001-2019 | 2000-2018 |
| | | Lakshadweep (UT) | 1956-11-01 | 64,473 | | | x |
| Madhya Pradesh HC | 1936-01-02 | Madhya Pradesh | 1950-01-26 | 72,626,809 | 2011-01-29 | x | 2000-2018 |

Continuation of Table A1

| High Court | Established | State | State Created | Population | Kanoon Rollout | eCourt HC | eCourt DC |
|---|---|---|---|---|---|---|---|
| Madras HC | 1862-08-15 | Tamil Nadu | 1956-11-01 | 72,147,030 | 2009-01-18 | 2001-2019 | 2000-2018 |
| | | Puducherry (UT) | 1962-08-16 | 1,247,953 | | | x |
| Manipur HC[c] | 2013-03-25 | Manipur | 1972-01-21 | 2,855,794 | x | 2001-2019 | 2000-2018 |
| Meghalaya HC[c] | 2013-03-23 | Meghalaya | 1972-01-21 | 2,966,889 | 2018-10-21 | 2005-2014 | 2000-2018 |
| Orissa HC | 1948-04-03 | Odisha | 1950-01-26 | 41,974,218 | 2011-01-26 | 2001-2019 | 2000-2018 |
| Patna HC | 1916-09-02 | Bihar | 1950-01-26 | 10,409,9452 | 2010-06-13 | 2001-2019 | 2000-2018 |
| Punjab and Haryana HC | 1947-08-15 | Chandigarh (UT) | 1966-11-01 | 1,055,450 | x | x | 2000-2018 |
| | | Haryana | 1966-11-01 | 25,351,462 | | | 2000-2018 |
| | | Punjab | 1966-11-01 | 27,743,338 | | | 2000-2018 |
| Rajasthan HC | 1949-06-21 | Rajasthan | 1950-01-26 | 68,548,437 | 2010-06-13 | 2001-2019 | 2000-2018 |
| Sikkim HC | 1975-05-16 | Sikkim | 1975-05-16 | 61,0577 | x | 2001-2009 | 2000-2018 |
| Telangana HC[a,b] | 2019-01-01 | Telangana | 2014-06-02 | 35,193,978 | x | 2001-2019 | 2000-2018 |
| Tripura HC[c] | 2013-03-23 | Tripura | 1972-01-21 | 367,3917 | x | 2010-2018 | 2000-2018 |
| Uttarakhand HC | 2000-11-09 | Uttarakhand | 2000-11-09 | 10,086,292 | x | 2001-2019 | 2000-2018 |

[a] 2 June 2014: bifurcation of Andhra Pradesh into Telangana and Andhra Pradesh: common HC renamed to "HC of Judicature at Hyderabad"

[b] 1 January 2019: Separation of HCs: "Andhra Pradesh HC" for Andhra Pradesh and "Telangana HC" for Telangana

[c] Until March 2013, Gauhati HC had jurisdiction over Meghalaya, Manipur, Tripura. "Manipur HC" established 25 March 2013, "Meghalaya HC" established 23 March 2013, "Tripura HC" established 23 March 2013

## A1 What was Kanoon's Competition? An Overview of Other Electronic Legal Databases

As noted in the paper, Kanoon was not the only electronic source of legal informaton in India. Several other databases were developed during the same time-period as Indian Kanoon. Some of these are described below:

**All India Reporter** The All India Reporter (AIR) is one of the oldest and most respected publishers of decisions from the Indian Supreme Court as well as various State high courts. With more than 17 journals and more than 1 million subscribers, it curates, edits, prints and disseminates digests, commentaries and analyses of key cases that are heard at the courts of India.

**Manupatra** This paid subscription database includes both primary sources (judicial opinions, statutes and other legislative materials, administrative agency materials, etc.) and secondary sources (including treatises and law journals). This company first launched its products in August 2001. The launch however, was via CD ROM. Additional media formats such as the online database, e-mail services, books and journals were added over time.

**SCC Online** This paid subscription database includes cases from a wide variety of Indian courts, including the Supreme Court, the Privy Council, high courts, district courts, and tribunals and commissions. It also includes selected case law from other jurisdictions in the region, including Bangladesh, Malaysia, Pakistan, and Sri Lanka, and from several African jurisdictions. SCC online also includes other Indian legal materials: acts and rules, articles, secondary sources, treaties, and more. It was introduced in 2010 with limited coverage that expanded over the next three years.

**LII of India** part of the Free Access to Law Movement, also provides an integrated search platform for primary and secondary sources from over a hundred Legal Information Institute (LII) databases of other countries and territories. This project was established in 2010 with the coordinated efforts of the Asian Legal Information Institute (AsianLII) project, funded by AusAID, and its Commonwealth Legal Information Institute project, funded by the Australian Research Council (ARC). Disruptions in funding however, prevented the database from reaching scale till late in 2012, when it was formally launched at the LII of India.

## A2 Additional Descriptive Statistics

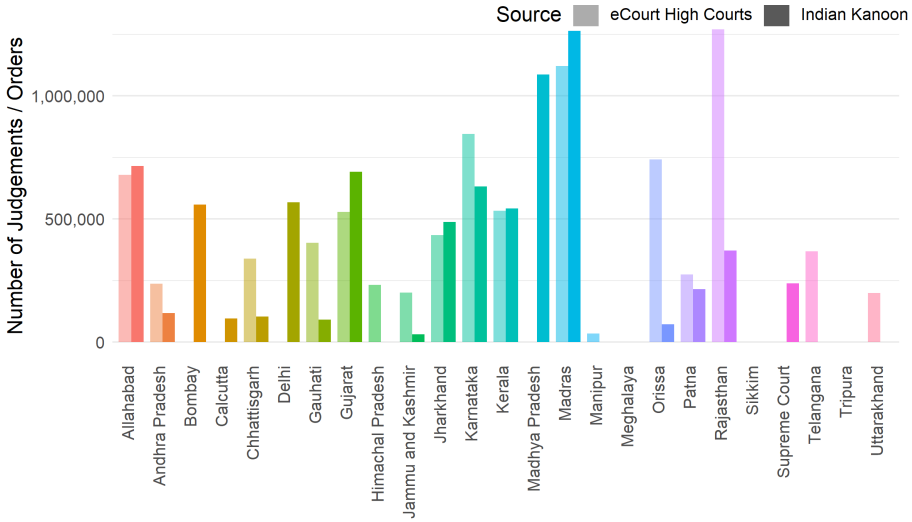


Figure A1: Number of Dispositions (eCourt High Court) and Judgements and Orders (Indian Kanoon) per high court for the years 2001-2019.

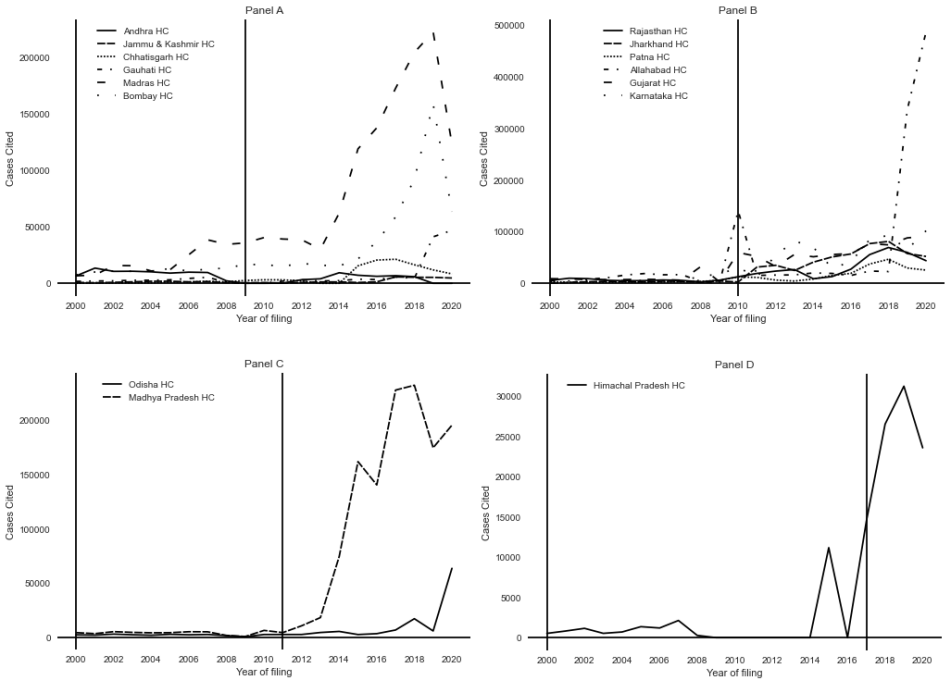


Figure A2: Total citations per year for different rollout years

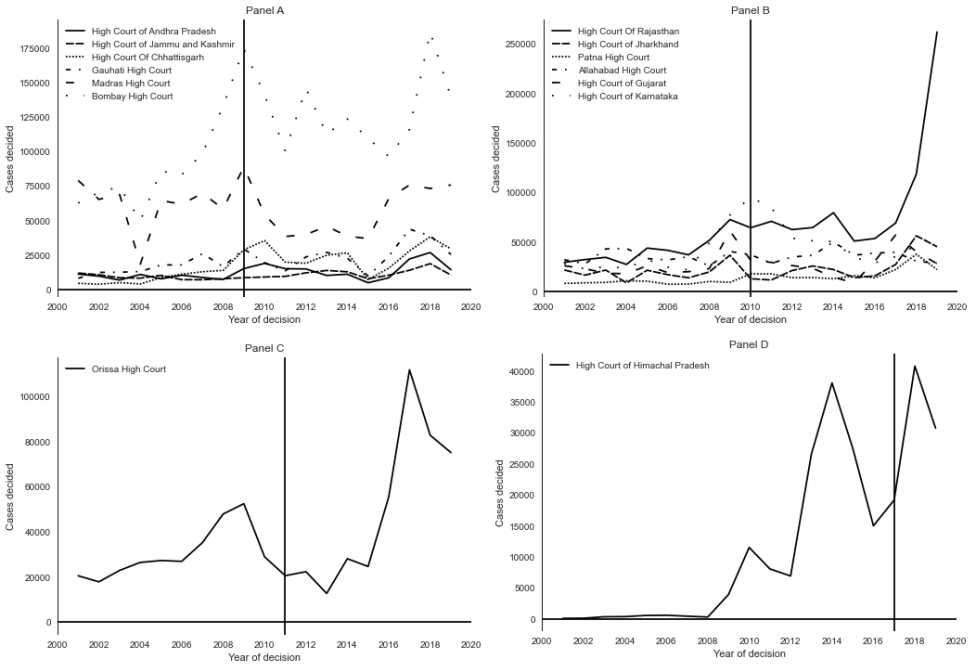


Figure A3: Total number of resolutions per year by rollout years (eCourts High Court)

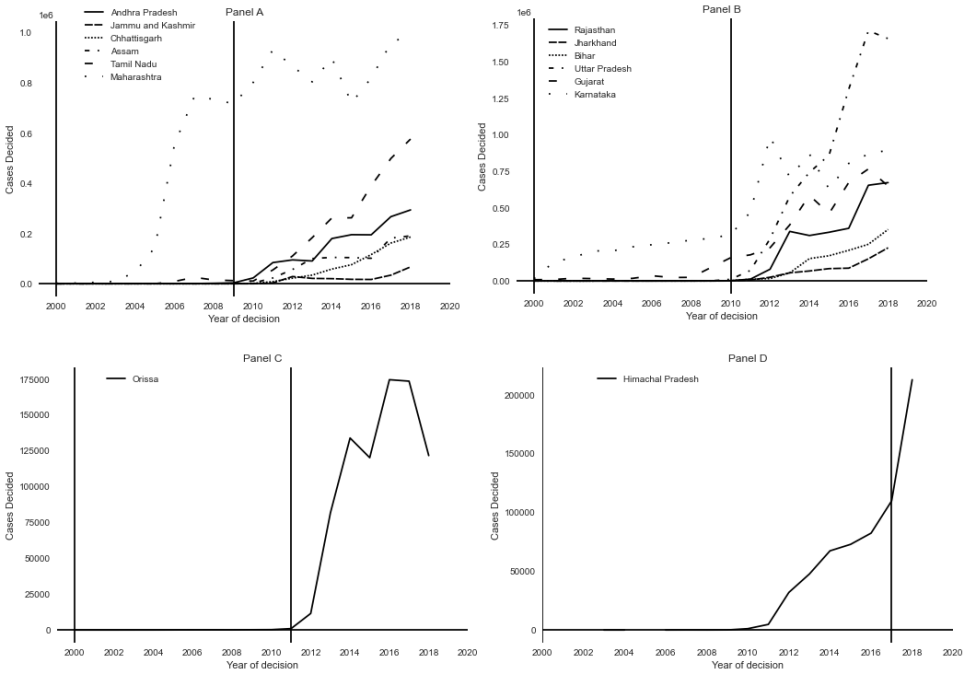


Figure A4: Total number of resolutions per year by rollout years (eCourts District Court)

# A3 Additional Results on Court Efficiency

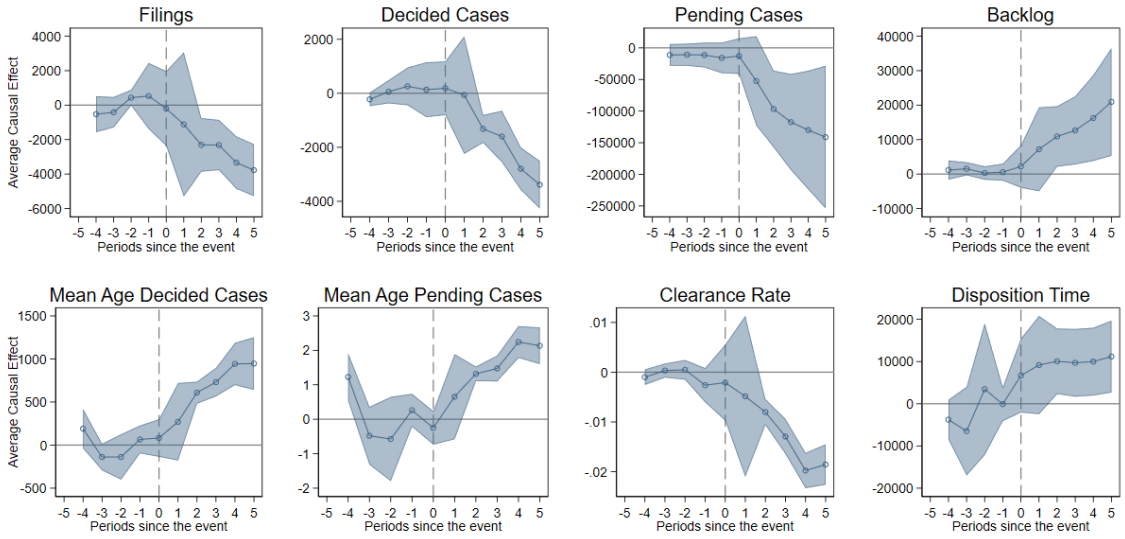


Figure A5: Placebo test to check pre-trends of the Aggregate Efficiency Measures of High Courts

*Notes:* This figure shows the results of the Placebo test for checking pre-trends as suggested by Chaisemartin and D'Houltfoeuille, 2020. The Placebo tests are constructed assuming that, for tests whose treatment actually happens at t, treatment occured at the time t-k, for k in (1,2,3,4). All variables are calculated at the state-year–month level. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

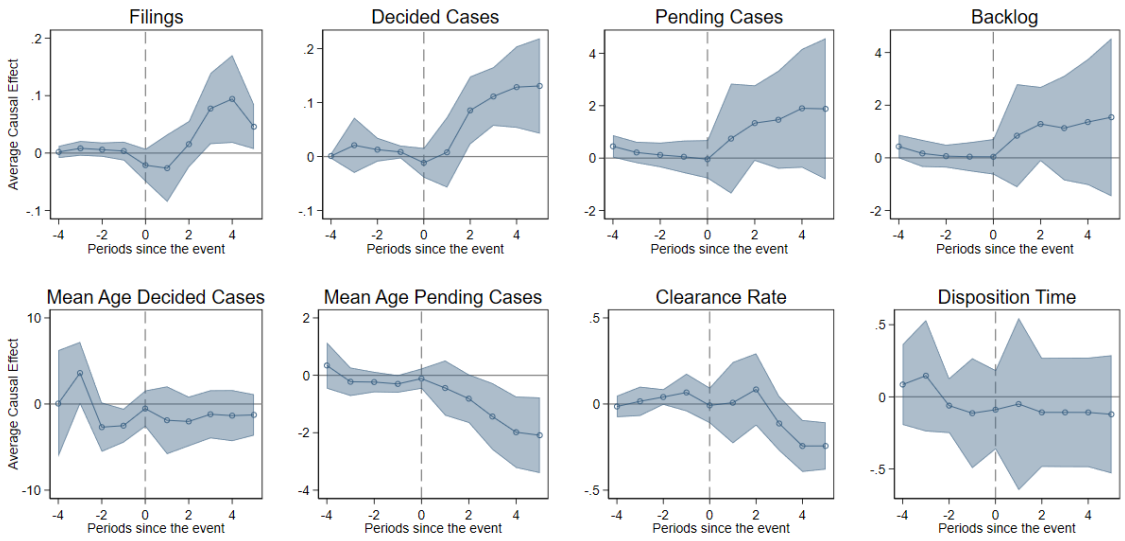


Figure A6: Placebo test to check pre-trends of the Aggregate Efficiency Measures of District Courts

*Notes:* This figure shows the results of the Placebo test for checking pre-trends as suggested by Chaisemartin and D'Houltfoeuille, 2020. The Placebo tests are constructed assuming that, for tests whose treatment actually happens at t, treatment occured at the time t-k, for k in (1,2,3,4). Outcomes are normalized in the following way: Filings, Decided Cases, Pending Cases and Backlog are measured in 100,000 cases. Disposition time divided by 100,000. Mean Age Decided Cases and Mean Age Pending Cases in 100 days. All variables are calculated at the state-year–month level. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

# A4 Additional Results for Appeals

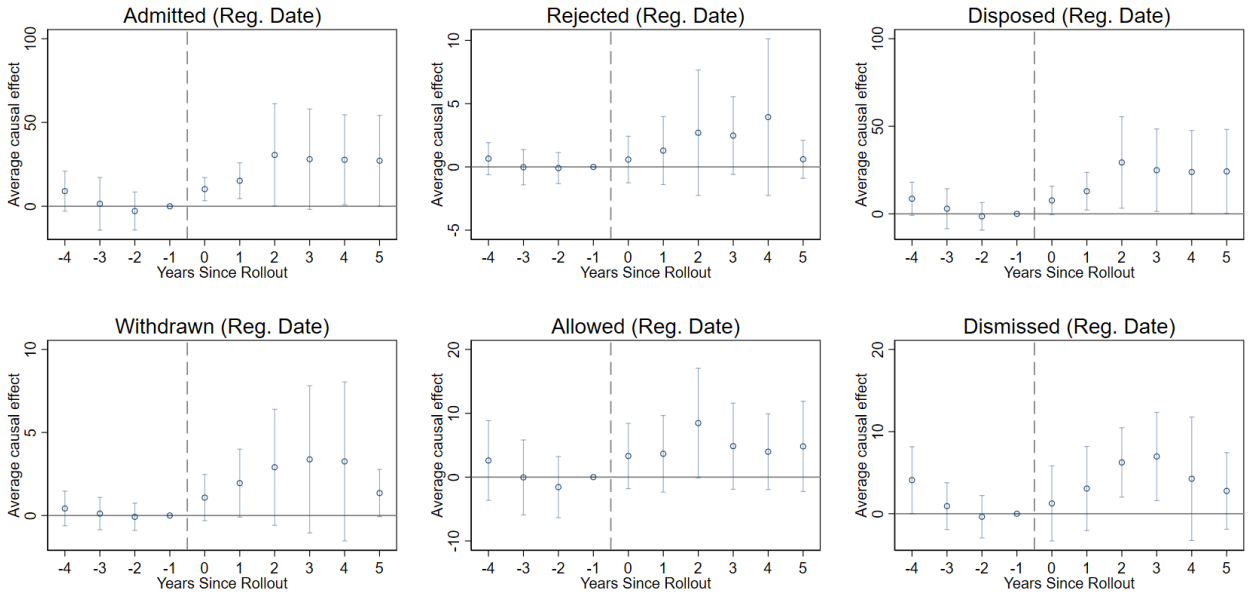


Figure A7: Event study analysis of the impact of Kanoon rollout on appeal outcomes based on the district court registration date

*Notes:* All variables are calculated at the state-year–month level. The event studies use the algorithm of Sun and Abraham (2021), where Himachal Pradesh (rollout 04/2017) is defined as control cohort. The estimation controls for state, year, state × year and month fixed effects. Standard errors are clustered at the state level. All variables are calculated at the state-year–month level. The estimation controls for state, year, state × year and month fixed effects. Standard errors are clustered at the state level.

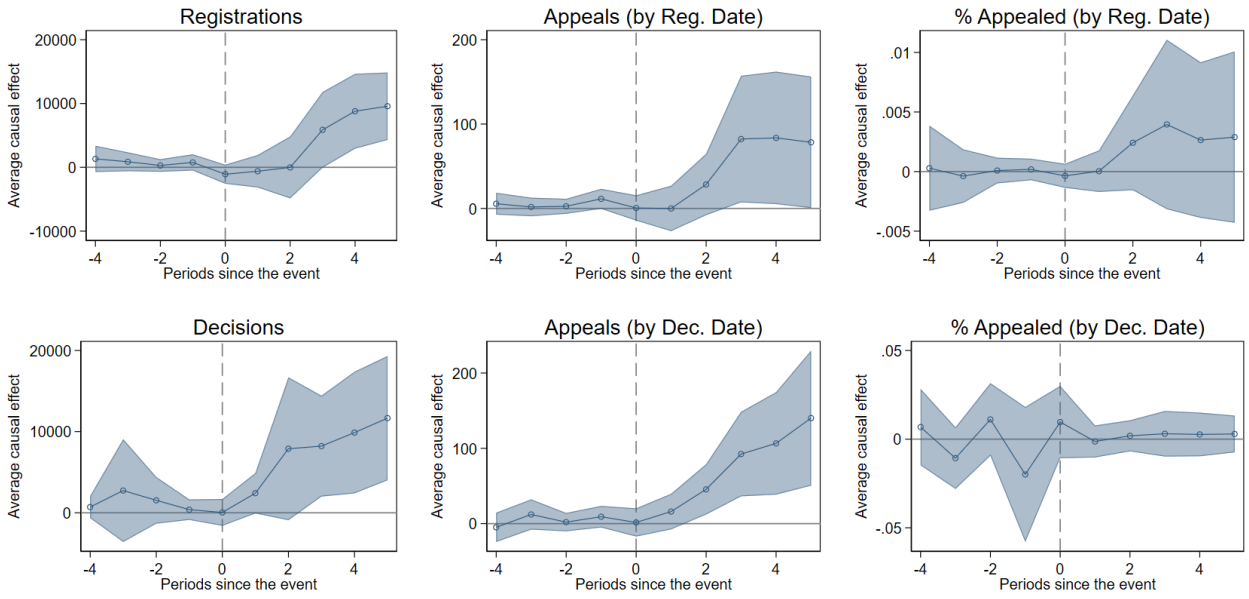


Figure A8: Placebo test to check pre-trends in the effects of Kanoon rollout on Appeals of district court cases.

*Notes:* This figure shows the results of the Placebo test for checking pre-trends as suggested by Chaisemartin and D'Houltfoeuille, 2020. The Placebo tests are constructed assuming that, for tests whose treatment actually happens at t, treatment occured at the time t-k, for k in (1,2,3,4). All variables are calculated at the state-year–month level. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

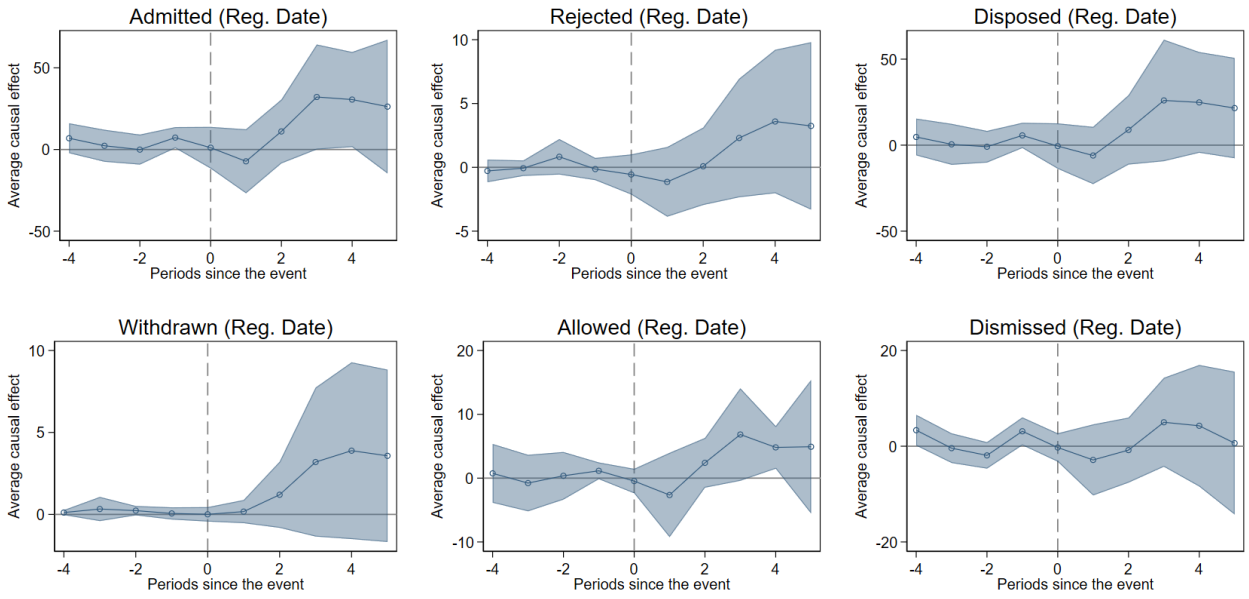


Figure A9: Placebo test to check pre-trends in the effects of Kanoon rollout on appeal outcomes based on district court registration date.

*Notes:* This figure shows the results of the Placebo test for checking pre-trends as suggested by Chaisemartin and D'Houltfoeuille, 2020. The Placebo tests are constructed assuming that, for tests whose treatment actually happens at t, treatment occured at the time t-k, for k in (1,2,3,4). All variables are calculated at the state-year–month level. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

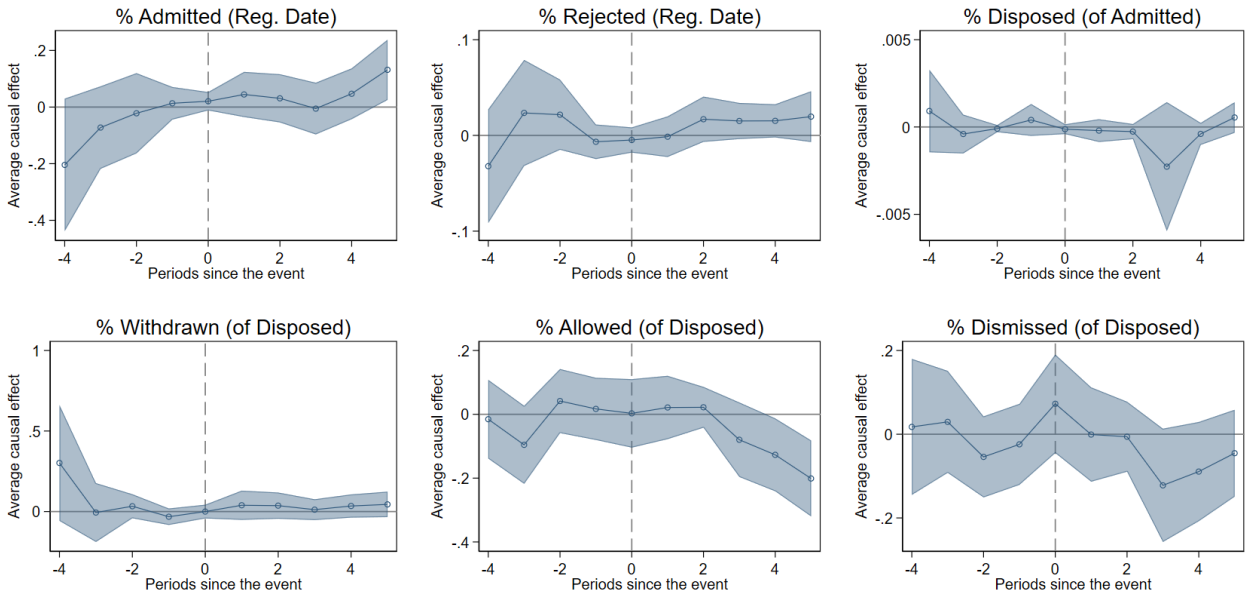


Figure A10: Placebo test to check pre-trends in the effects of Kanoon rollout on appeal outcomes based on the district court registration date, with variables defined in conditional percentages based on the sample in the previous stage of the pipeline of justice.

*Notes:* This figure shows the results of the Placebo test for checking pre-trends as suggested by Chaisemartin and D'Houltfoeuille, 2020. The Placebo tests are constructed assuming that, for tests whose treatment actually happens at t, treatment occured at the time t-k, for k in (1,2,3,4).

### A4.1 Appeals Data Creation

For the appeal regressions in Tables 7 and 8 and Appendix Tables A7, A8, A9, and A10 we created a district-year-month level sample with variables drawn from the database of cases at the district courts. These include: *Registrations*, *Decisions*, *Appeals by Reg. Date*, *Appeals by Dec. Date*, *% Appealed by Reg. Date*, *% Appealed by Dec. Date* and high court variables *Admitted*, *Rejected*, *Withdrawn*, *Allowed*, *Dismissed* and *Disposed*. In the following we describe the creation of this sample. Figure A11 gives an overview of the merges of the initial and intermediate datasets.

1. **Identify DC cases appealed in HCs**
   High court cases have some information about the subordinate court case giving rise to the appeal. This information is in the form of the decision date of the lower court case, the registration number, and the registration year (or a subset of these variables). However, this information is not sufficient to uniquely identify the lower court case. In order to find the appropriate lower court case, we restrict the potential set of matches in the district data to those cases in the appropriate state (determined by which states are administratively under the given High court) matching the registration number, registration year, and the decision date. Finally, from this restricted set of cases, we choose the case where the litigant's names in the district case closely match (i.e., above a threshold) the litigant's names in the High court case.

2. **Count number of cases decided / registered in DCs per state-month**
   From the original eCourt DC data, we count how many cases were registered and decided per state-year-month (*Registrations* and *Decisions*).

3. **Calculate appeal counts per state-month (DC / HC and Registrations / Decisions)**
   To the matched appeal case dataset, we merge in district court registration and decision dates using the unique district court case identifier (CINO) as merge key. From there, we can then calculate the number of appeals per state-year-month based on the DC registration date and based on the DC decision date (*Appeals by Reg. Date* and *Appeals by Dec. Date*).

4. **Percent of cases appealed**
   We merge together the state-year-month counts of registrations and decisions in DCs with the state-year-month counts of appeal cases by registration and decision date. This allows us to calculate the two variables *% Appealed by Reg. Date* and *% Appealed by Dec. Date*.

5. **Calculate number and percentage of HC appeals**
   To the matched appeal case dataset, we merge in HC registration and decision dates and additional information of these HC cases, especially "disposal name". We use the string variable disposal name to create the variables *Admitted*, *Rejected*, *Withdrawn*, *Allowed*, *Dismissed* and *Disposed* and create aggregate counts and percentages of them at the state-year-month level based on the DC registration date and decision date.

6. **Merge all state-year-month variables together**
   Finally, we merge all these state-year-month counts and percentages of DC registrations and decisions, the counts and percentages of DC cases appealed in HC, and the counts and percentages of the outcomes in the HCs together in one final dataset.

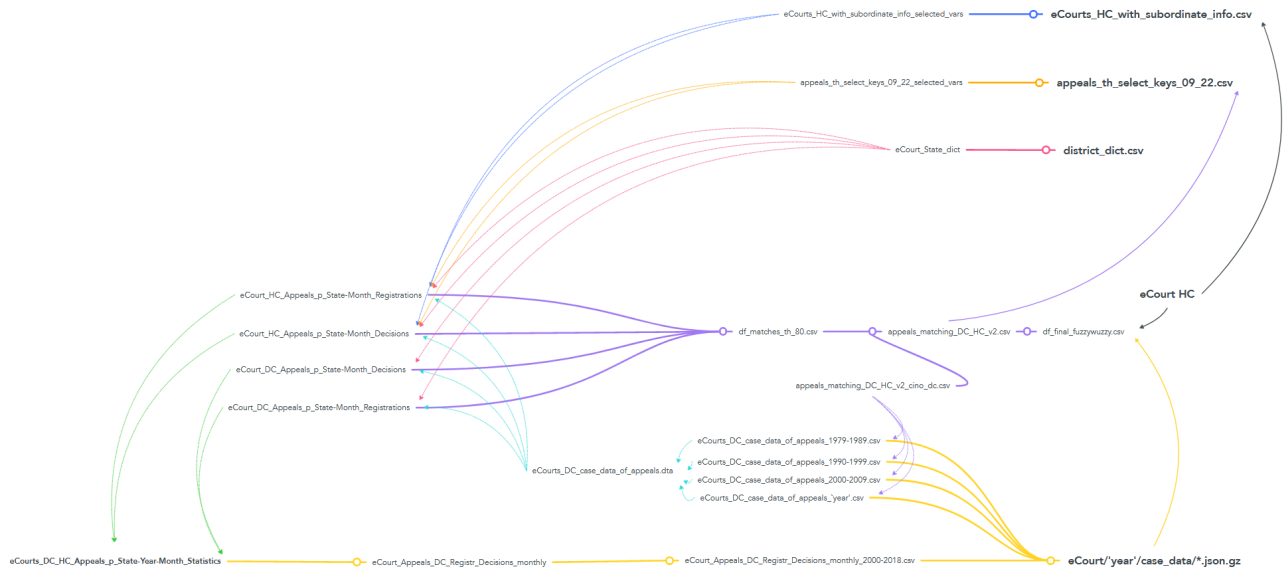


Figure A11: Construction of final Appeal Regressions dataset

*Notes:* This figure gives an overview how we created, starting from our three initial datasets (eCourt HC, eCourt DC and a district dictionary) the final district-year-month dataset used for the appeal regressions.

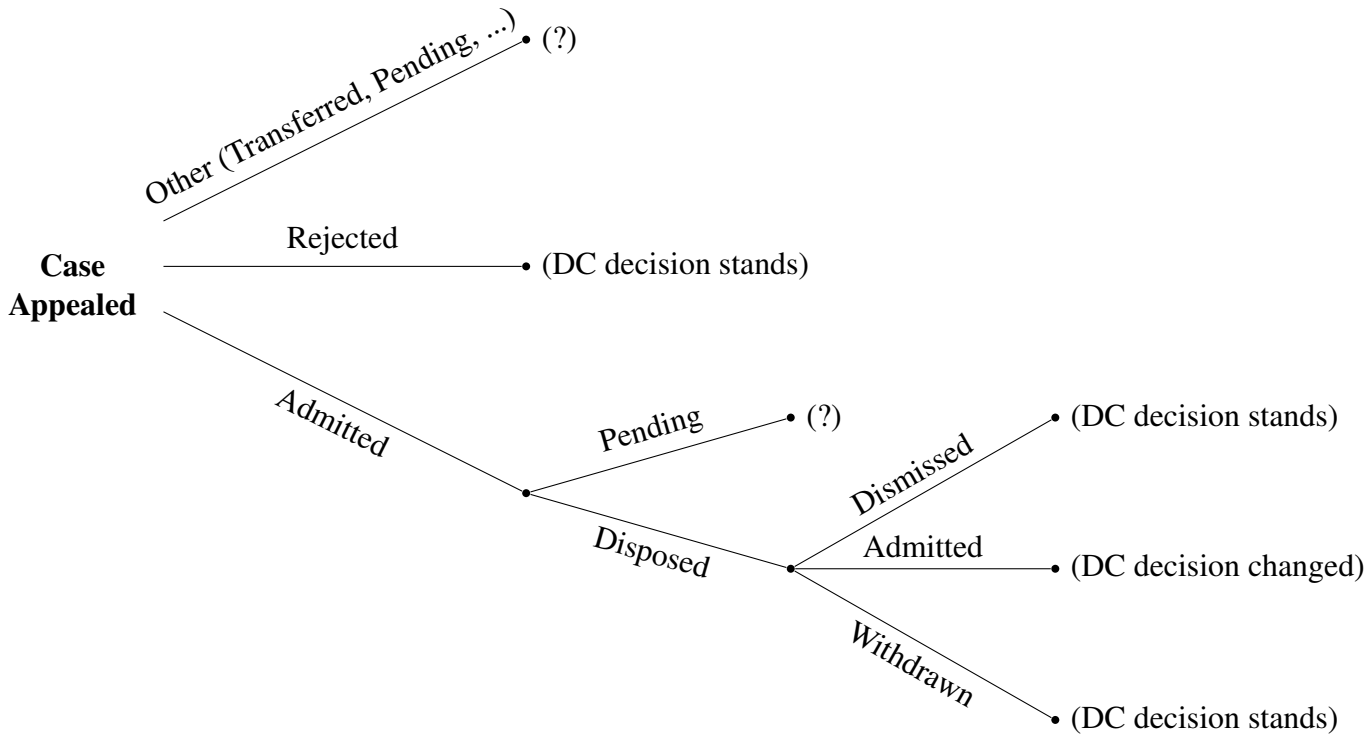


Figure A12: Decision Tree for DC cases appealed in the HC

# A5 Additional Results on Firms

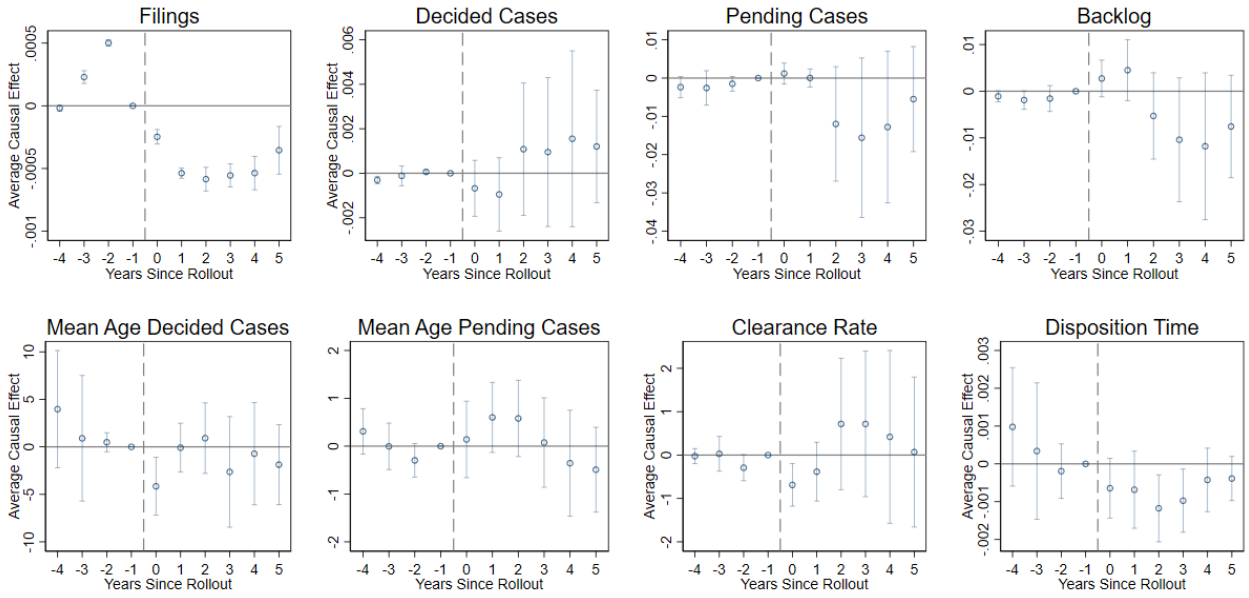


Figure A13: Impact of Kanoon Rollout on Aggregate Efficiency Measures for Cases linked to Prowess.

*Notes:* All variables are calculated at the state-year–month level. The event studies use the algorithm of Sun and Abraham (2021), where Himachal Pradesh (rollout 04/2017) is defined as control cohort. The estimation controls for state, year, state × year and month fixed effects. Standard errors are clustered at the state level.

*Notes:*

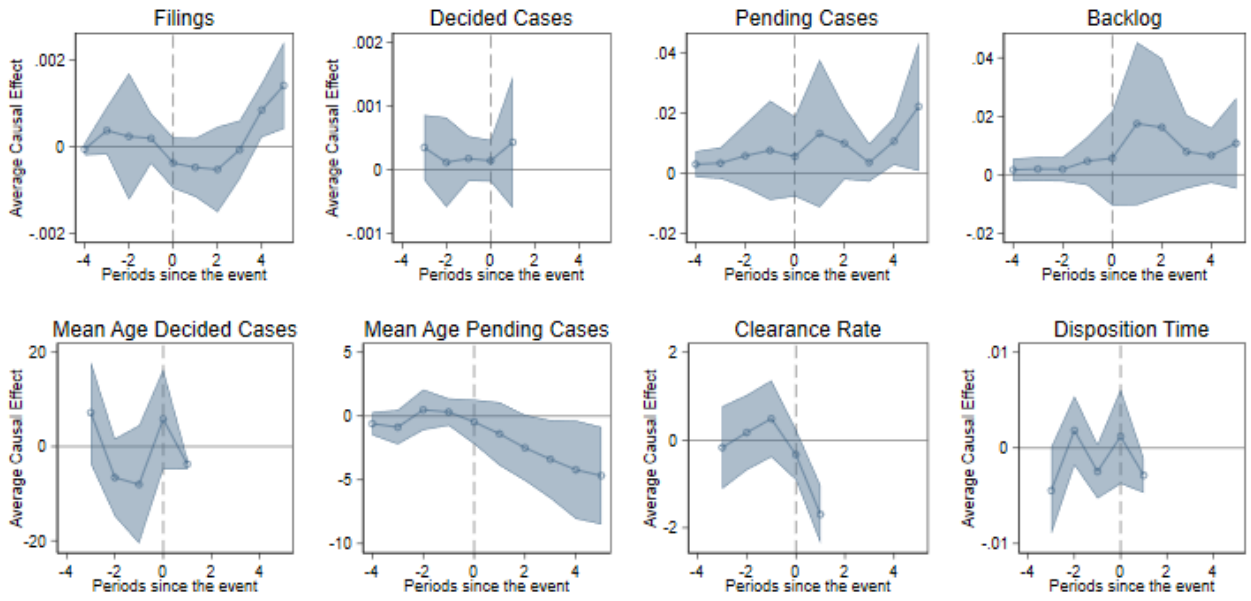


Figure A14: Placebo test to check pre-trends of the Aggregate Efficiency Measures for Cases linked to Prowess

*Notes:* This figure shows the results of the Placebo test for checking pre-trends as suggested by Chaisemartin and D'Houltfoeuille, 2020. The Placebo tests are constructed assuming that, for tests whose treatment actually happens at t, treatment occured at the time t-k, for k in (1,2,3,4). All variables are calculated at the state-year–month level. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

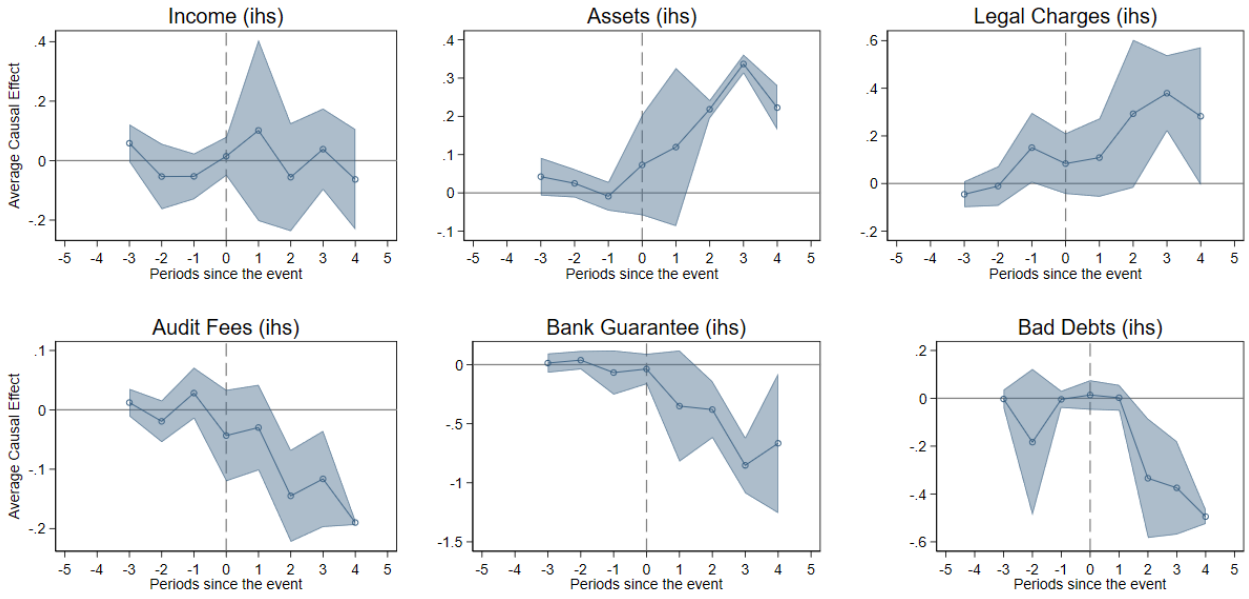


Figure A15: Placebo test to check pre-trends in the General Equilibrium effects of Kanoon rollout on Firm Financials

*Notes:* This figure shows the results of the Placebo test for testing pretrends as suggested by Chaisemartin and D'Haultfoeuille, 2020. The placebo estimates are constructed assuming that, for units whose treatment actually happens at $t$, treatment occurred at time $t-k$ for $k \in (1,2,3,4)$. All variables are calculated at the state-year–month level. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

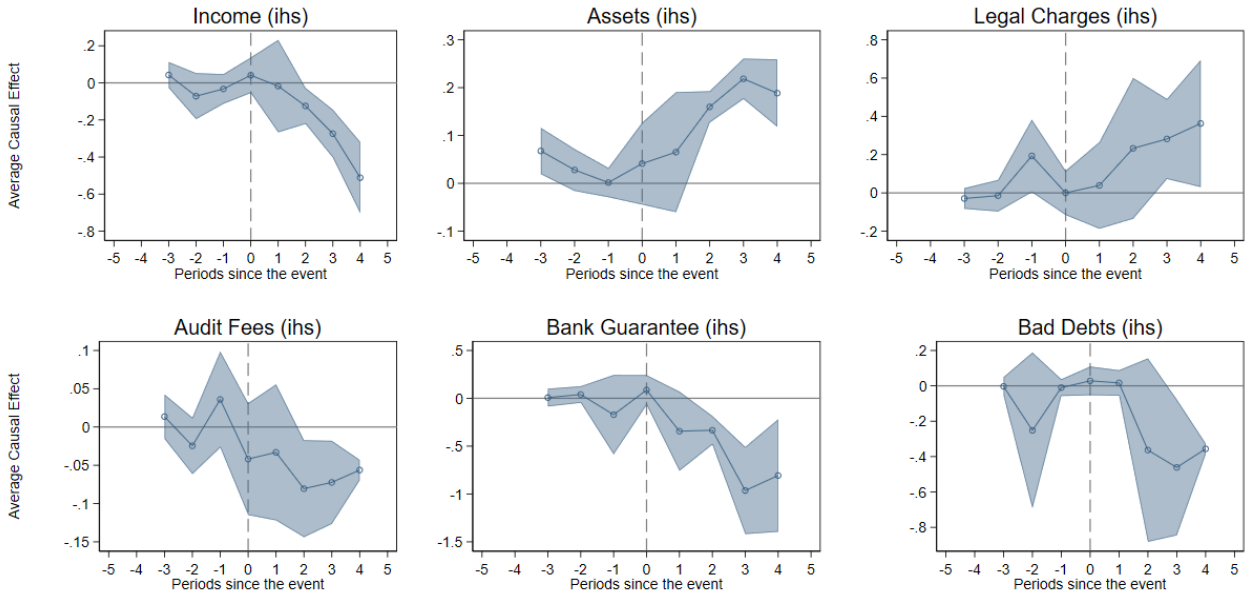


Figure A16: Placebo test to check pre-trends in the Partial Equilibrium effects of Kanoon rollout on Firm Financials

*Notes:* This figure shows the results of the Placebo test for testing pretrends as suggested by Chaisemartin and D'Haultfoeuille, 2020. The placebo estimates are constructed assuming that, for units whose treatment actually happens at $t$, treatment occurred at time $t-k$ for $k \in (1,2,3,4)$. All variables are calculated at the state-year–month level. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the state level.

## A6 Additional Analysis: Did Kanoon Change Citation Practices?

Since Kanoon makes it easier for all stakeholders to examine past cases of relevance – a key feature of a common law system – we also examine a set of outcome variables that measure the links between a high court case and other cases that may have occurred prior to this case, or going forward. *Backward citations* measures the number of (past) cases a judgment is citing. *Forward citations* counts the number of times a judgment is cited in the future. *Self backward citations* is the number of times a judgment is citing cases from the same court in which it was heard. *Self forward citations* is the number of times a judgment is cited in the future by cases in the same court. *Degree centrality* is a measure of the number of cases a particular case is linked to (either citing or cited by). *Eigenvector centrality* is a measure of how influential a case is: i.e., either citing more influential cases or being cited by more influential cases (here influence can be interpreted to be the number of cases it cites or gets cited by). *Square concentration* is measured as the square of the share of the number of citations of a particular case in a particular year. Summing over the squared concentrations of all the cases in a given year gives us the Herfindahl-Hirschman index (HHI) for that year.

We conduct our analysis for not only cases on Kanoon but also a subset of cases listed in the All India Reporter (AIR). This is one of the oldest and most respected legal publications in India. Court decisions that are published in the AIR are routinely cited in official proceedings. We access appeal cases (cases originating in the district courts) on the publisher's website by using the 'comprehensive search' feature by querying for cases by appeal year. This data set includes cases published in AIR and other allied journals from the high courts and the Supreme Court of India spanning the years from 1980 to 2021. The cases were then matched to cases in the Indian Kanoon database by using the decision date and litigant names.

### Kanoon cases

Table A3 presents results on backward citations that build up to our preferred specification (Column 4). Table A4 presents the regression results for this preferred specification for all outcomes of interest. The results in Table A3 suggest that Kanoon has a positive impact on the number of backward citations by year 3 after the rollout. The result is significant at the 1% confidence level in full specification with all the fixed effects (Column 4). This suggests that having greater information at hand for lawyers and judges led to more legal support included as citations by the third year of the rollout. We note however, that the magnitude of the effect is quite small, it is only observed in two specifications and the effect is quite delayed: there is less than 1 extra citation three years after the rollout.

Additional measures of the information content of cases are explored in Table A4. The results column 5 in Table A4 suggest that Kanoon has a positive impact on the eigenvector centrality in the year of its rollout and that this effect is significant at the 5% level in the year of rollout. There is, however, no other impact on any of the other citation variables which we considered.

In light of these results, we believe it is extremely unlikely that any of the observed impacts of Kanoon were driven by sustained or pronounced changes in the citation styles or citation patterns of the litigants or lawyers. In other words, the system of referencing prior precedents in India's common law system does not appear to be likely to drive most of the results seen in this paper. We believe it is more likely that it was the decline in the processing times of cases, and the greater efficiency of the system that was benefited most by Kanoon.

**Analysis of AIR Cases**

The last subsection studied the impact of Kanoon rollout on citation patterns of high courts. The analysis is based on the cases uploaded to Kanoon, so relying on the assumption that the cases uploaded to Kanoon are similar between those which were decided before the Kanoon rollout and those decided afterwards. To relax this assumption and exclude that the observed null effects on citation patterns are driven by differential case upload, we now focus on a very specific but important subset of cases.

We limit our analysis to cases deemed to constitute important precedents by a committee of juridical experts. The AIR database consists of cases that are approved by this committee for citing by judges in official proceedings at the courts of India. It serves as the important sample that sets policy (by influencing future judges). We conduct our analysis of the impact of the free Kanoon legal search on this sample to keep similar the composition of cases before and after Kanoon is rolled out. This allows us to study the causal impact of reducing information frictions on rule of law.

Figure A17 presents results for these cases. We note that Kanoon seems to facilitate the self-citations of AIR cases especially within the first two years of its rollout. Apart from this, there is not much effect of Kanoon on the citations of or by AIR cases. We interpret this as additional evidence that the arrival of Indian Kanoon did not induce significant changes in the processes of information curation within India's common law system. Rather, it was the productivity and efficiency gains within courts and firms that is likely to generate the results seen in this paper.

Table A2: Additional Summary Statistics

| | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| *Indian Kanoon Data* | | | | | |
| Backward Citations | 196,030 | 2.81 | 4.19 | 1 | 179 |
| Forward Citations | 38,942 | 3.27 | 16.28 | 1 | 1,621 |
| Backward Self-citations | 46,617 | 1.56 | 1.33 | 1 | 94 |
| Forward Self-citations | 27,502 | 2.76 | 14.43 | 1 | 1,616 |
| Squared Concentration | 196,030 | 21.66 | 256.92 | .00096 | 10,000 |
| Degree Centrality | 2,306,231 | 0.02 | 0.05 | .00062 | 1.4 |
| Eigenvector Centrality | 2,306,231 | 0.06 | 0.12 | 0 | .78 |
| *All India Reporter (AIR) Data (Case Level)* | | | | | |
| AIR cases citing AIR cases | 23,658 | 0.11 | 0.39 | 0 | 7 |
| AIR cases citing all cases | 23,658 | 4.47 | 5.80 | 1 | 116 |
| AIR cases citing non-AIR cases | 23,658 | 4.36 | 5.70 | 0 | 113 |
| All cases citing AIR cases | 948,144 | 0.03 | 0.20 | 0 | 9 |
| Non-AIR cases citing AIR cases | 924,486 | 0.03 | 0.20 | 0 | 9 |
| Self-citations of AIR cases (same court) | 23,658 | 0.55 | 1.14 | 0 | 25 |
| Outside-citations of AIR cases (different courts) | 23,658 | 1.18 | 1.72 | 0 | 44 |
| Self-citations - Outside-citations | 23,658 | -0.63 | 1.90 | -44 | 19 |

Table A3: Impact of Kanoon Rollout on Backward Citations.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Year 3 before rollout | 0.906 | 0.714 | 0.971 | 0.805 |
| | (0.65) | (0.38) | (0.70) | (0.67) |
| Year 2 before rollout | 0.553 | -0.208 | 0.243 | 0.152 |
| | (0.34) | (0.25) | (0.23) | (0.24) |
| Year of rollout | 0.056 | 0.187 | 0.050 | 0.128 |
| | (0.16) | (0.20) | (0.19) | (0.18) |
| Year 1 after rollout | 0.068 | 0.515 | -0.053 | 0.138 |
| | (0.30) | (0.40) | (0.21) | (0.19) |
| Year 2 after rollout | 0.357 | 0.817 | 0.152 | 0.456 |
| | (0.30) | (0.41) | (0.30) | (0.23) |
| Year 3 after rollout | 0.384 | $0.936^{*}$ | 0.326 | $0.731^{**}$ |
| | (0.33) | (0.42) | (0.26) | (0.23) |
| constant | $2.414^{***}$ | $2.106^{***}$ | $2.583^{***}$ | $2.414^{***}$ |
| | (0.29) | (0.25) | (0.11) | (0.14) |
| Court FE | Y | Y | Y | Y |
| Year FE | | Y | Y | Y |
| Court X Year FE | | | Y | Y |
| Month FE | | | | Y |
| R-sqr | 0.061 | 0.068 | 0.081 | 0.082 |
| N | 196030 | 196030 | 196029 | 196029 |
| mean Y | 2.8136 | 2.8136 | 2.8136 | 2.8136 |

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

*Backward citations* are defined as the number of (past) cases a judgement is citing. Estimation uses the algorithm of Sun and Abraham (2021), where Himachal Pradesh (rollout 04/2017) is defined as control cohort. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the high court level.

Table A4: Impact of Kanoon on Citations and Centrality.

| | Backward Citations | Forward Citations | Backward Self-citations | Forward Self-citations | Eigenvector Centrality | Degree Centrality | Square Concentration |
|---|---|---|---|---|---|---|---|
| Year 3 before rollout | 0.805 | -0.595 | -0.007 | 0.047 | 0.034 | 0.030** | 27.278 |
| | (0.67) | (0.79) | (0.03) | (1.05) | (0.02) | (0.01) | (20.23) |
| Year 2 before rollout | 0.152 | -0.314 | -0.042 | 0.506 | 0.007 | 0.011** | 22.372 |
| | (0.24) | (0.65) | (0.02) | (0.98) | (0.01) | (0.00) | (19.88) |
| Year of rollout | 0.128 | -0.687 | 0.036 | -0.601 | 0.008* | 0.002 | 1.787 |
| | (0.18) | (0.46) | (0.02) | (0.36) | (0.00) | (0.00) | (4.83) |
| Year 1 after rollout | 0.138 | 1.034 | -0.028 | 1.024 | 0.007 | 0.001 | -7.506 |
| | (0.19) | (0.95) | (0.06) | (0.79) | (0.01) | (0.00) | (6.34) |
| Year 2 after rollout | 0.456 | 0.518 | -0.002 | 0.304 | 0.000 | -0.001 | -17.080 |
| | (0.23) | (1.09) | (0.05) | (1.07) | (0.01) | (0.00) | (11.21) |
| Year 3 after rollout | 0.731** | 0.855 | 0.028 | 0.727 | -0.003 | -0.004 | -20.570 |
| | (0.23) | (1.37) | (0.06) | (1.39) | (0.01) | (0.00) | (12.91) |
| constant | 2.414*** | 3.188*** | 0.868*** | 2.239** | 0.057*** | 0.021*** | 26.466** |
| | (0.14) | (0.58) | (0.03) | (0.61) | (0.01) | (0.00) | (7.82) |
| R-sqr | 0.082 | 0.005 | 0.106 | 0.082 | 0.100 | 0.239 | 0.053 |
| N | 196,029 | 38,939 | 46,613 | 27,498 | 2,306,231 | 2,306,231 | 196,029 |
| mean Y | 2.81 | 3.27 | 1.56 | 2.76 | 0.06 | 0.02 | 21.66 |
| Court FE | Y | Y | Y | Y | Y | Y | Y |
| Year FE | Y | Y | Y | Y | Y | Y | Y |
| Court X Year FE | Y | Y | Y | Y | Y | Y | Y |
| Month FE | Y | Y | Y | Y | Y | Y | Y |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Backward citations* measures the number of (past) cases a judgement is citing. *Forward citations* is a similar measure that measures the number of times a judgement is cited in the future. *Self backward citations* is the number of times a judgement is citing cases from the same court in which it was heard. *Self forward citations* is the number of times a judgement is cited in the future by cases in the same court. *Degree centrality* is a measure of the number of cases a particular case is linked to (either citing or cited by). *Eigenvector centrality* is a measure of how influential a case is i.e. either citing more influential cases or being cited by more influential cases (here influence can be interpreted to be the number of cases it cites or gets cited by). *Square concentration* is measured as the square of the share of the number of citations of a particular case in a particular year. Estimation uses the algorithm of Sun and Abraham (2021), where Himachal Pradesh (rollout 04/2017) is defined as control cohort. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the high court level.

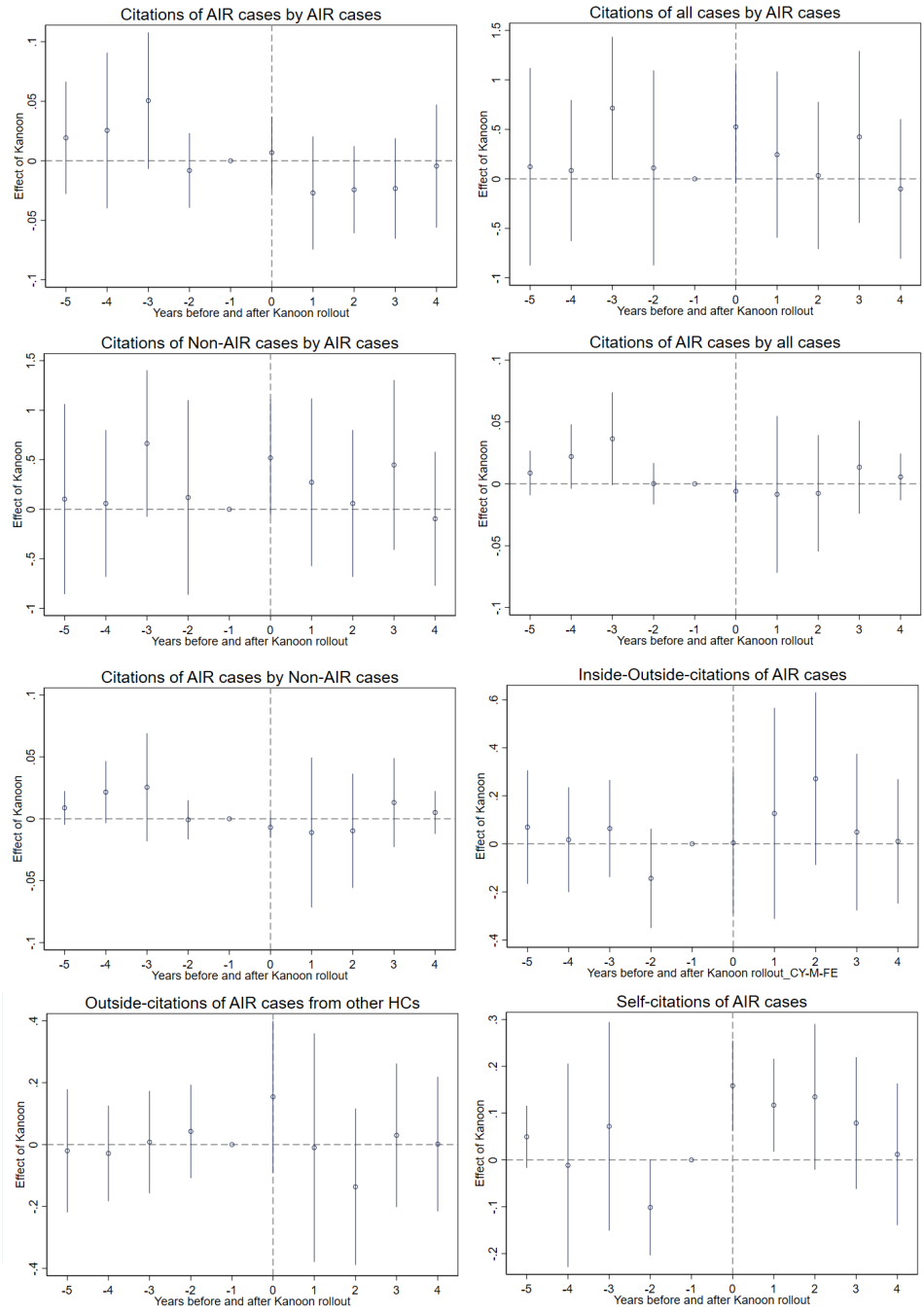


Figure A17: Citation analysis of the AIR cases.

This estimation is conducted on the subset of appeal cases (cases originating in the district courts) on the publisher's website by using the 'comprehensive search' feature by querying for cases by appeal year. This data set includes cases published in AIR and other allied journals from the high courts and the supreme court of India spanning the years from 1980 to 2021. The cases were then matched to cases in the Indian Kanoon database by using the decision date and litigant names. The estimation controls for state, year, state $\times$ year and month fixed effects. Standard errors are clustered at the high court level.