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

Sandeep Bhupatiraju Daniel L. Chen Shareen Joshi
Peter Neis

May 20, 2024

Abstract

This paper examines the role of digital platforms, specifically Indian Kanoon, in providing free legal search services and their impact on overcoming economic barriers at the market level. Utilizing a generalized difference-in-differences approach, our analysis reveals that the staggered rollout of Kanoon has significantly improved access to justice and court efficiency without compromising decision quality. Notably, the advent of Kanoon has instigated profound financial alterations in firms, evidenced by marked increases in assets, income, expenditures, and legal charges. These effects were observed not only in litigating firms but across the wider business spectrum. Consequently, this study sheds light on the pivotal role of unrestricted access to legal information in driving economic progress.

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

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

*Bhupatiraju: World Bank, 1818 H Street Washington, DC 20433 (e-mail: sbhupatiraju@worldbank.org); Chen: Toulouse School of Economics, Université Toulouse Capitole, Toulouse, France (e-mail: daniel.chen@iast.fr); Joshi: Walsh School of Foreign Service, Georgetown University, 3700 O Street, Washington DC, 20057 (e-mail: shareen.joshi@georgetown.edu); Neis: Université Clermont Auvergne, CNRS, IRD, CERDI, F-63000, Clermont-Ferrand, France (e-mail: peter.neis@uca.fr); We are grateful to the World Bank and Georgetown University's Summer grant opportunity for financial support. Daniel L. Chen 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. Peter Neis gratefully acknowledges the support received from the *Agence Nationale de la Recherche* of the French government through the program "*Investissements d'avenir*" (ANR-10-LABX-14-01). We thank Sushant Sinha, the founder of Indian Kanoon for sharing data and his experiences of establishing this platform. For research assistance, we thank Ritesh Das, 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

Fair justice systems require meaningful and reliable access to legal information for all stakeholders (United Nations 2015; OECD 2017).[1] Accessing legal information however, remains challenging due to high costs and complexity of legal language and processes (United Nations 2016). As a result, accessing legal information can be prohibitively expensive and difficult to understand for a large portion of the population.

Recent landmark randomized controlled trials, in South Africa and Pakistan most notably, suggest that the provision of legal information can have economic impacts on firm performance and citizen trust in formal institutions (Bertrand and Crépon 2021; Acemoglu et al. 2020). Legal information can potentially mitigate hold-up problems among trading partners, a problem illustrated by Boehm and Oberfield (2020) in their study on court efficiency and vertical integration among Indian firms.

This paper contributes to the literature on the rule(s) of law and macro-development. We examine the general equilibrium impact of a free legal search engine on a range of judicial and economic outcomes in India, the world's largest common law nation and where the costs of accessing legal information are high.[2] Established in 2008 by Sushant Sinha, an overseas graduate student, Indian Kanoon set out with the mission to offer free and readily accessible legal information to all (Iyengar 2010). Today this website includes all laws and all court orders from the Supreme Court, High Courts and Tribunals in court history. It is thus widely regarded as a "first-stop" in a search for legal information in India.[3]

Our study relies on comprehensive data that was scraped from both Kanoon and the eCourts' official records of Indian district and high courts. We extracted all accessible cases from Kanoon, coding them for judgment dates, courts, and citations. Additionally, we retrieved case-level metadata from the eCourts database, including filing and decision dates, to develop measures of court efficiency. A crucial step in our analysis was merging district and high court metadata to trace appeals. To bridge the gap between legal proceedings and economic impacts, we aligned firms mentioned in district court cases with their financial

[1] The 2030 Agenda for Sustainable Development, adopted by the United Nations General Assembly in September 2015, includes Goal 16 on promoting peaceful and inclusive societies, providing access to justice for all and building effective, accountable and inclusive institutions at all levels.

[2] In a survey data reported by the World Justice Project (2019), 32% of Indians report experiencing a legal problem in the past two years, 39% knew where to get legal information, and it took the average respondent 35 months to resolve their legal issue.

[3] Sinha shared in personal conversation in February 2022 that Indian Kanoon experiences 2.9 million search queries and 1.6 million sessions per month, with 6 minutes of time spent on each page.

profiles in the Prowess database, which contains detailed financial indicators sourced from formal firm balance sheets.

We leverage the phased integration of the Kanoon platform across various states in the framework of an event study (Borusyak, Jaravel, and Spiess 2021; Sun and Abraham 2021) to ascertain Kanoon's causal impact on legal cases, court operations, and corporate entities.

Our first analysis explores Kanoon's influence on district courts, revealing that its introduction corresponded with a substantial rise in filings (an additional 6.9 thousand per state and month) and decisions (8.3 thousand per state and month). Despite this uptick, the efficiency metrics, such as the backlog of cases pending over a year and average decision times, remained largely unaffected. Notably, we detected no compromise in the quality of district courts' decisions, evidenced by the unchanged rate of case appeals. Moreover, tracking these cases into the high courts, we observed a 2.6 percentage point increase in case withdrawals and a 13 percentage point decrease in overturn rates. These patterns suggest that Kanoon's deployment enhanced access to judiciary services.

The high courts, moreover, exhibited complementary trends, with a notable reduction in filings and decisions, primarily in original civil litigation cases. Interestingly, this decline did not translate into improved disposition times; in fact, we observed a slight backlog increase. This outcome suggests that Kanoon might have streamlined case processing at the high courts, possibly by curbing erroneous filings, promoting alternative dispute resolutions, and discouraging forum-shopping.

Our panel analysis extends to Kanoon's economic ramifications, particularly its effect on the balance sheets of major firms accounting for 70% of India's industrial output. The data reveals a positive influence on large formal firms' financial health. Firms involved in district court litigation during our study period experienced significant growth in yearly income (48%), assets (62%), expenditures (54%), and legal charges (38%). While these effects were slightly muted across the broader firm spectrum, they remained substantial and were most pronounced in the highly-regulated manufacturing and financial sectors, reflecting reduced informational frictions within these firms.

Collectively, these findings underscore Kanoon's role in diminishing legal information costs, alleviating bottlenecks, and enhancing judicial efficiency, all without eroding decision quality. These improvements had in turn a significant effect on the economy.

Our work contributes to a wider economic literature on the links between the efficiency of legal systems with economic advancement (Djankov et al. 2003; Ponticelli and Alencar 2016; Lichand and Soares 2014; Visaria 2009; Kondylis and Stein 2018; Chemin 2020;

Rao 2021). Notably, our results complement prior findings from randomized control trials accentuating the partial equilibrium causal linkage between legal information and favorable economic outcomes for firms (Bertrand and Crépon 2021). Information friction costs can be quite high for firms (Bloom, Eifert, et al. 2013; Bloom, Manova, et al. 2021; Guillouet et al. 2021). Different from these, by asking what happens when all citizens get access to a free online legal search platform, we substantiate the widely held belief that the absence of common knowledge about laws and regulations levies both direct and secondary costs on economies. When people know that other people have access to the same legal information, societies can have a predictable socio-economic landscape with greater accountability and higher levels of trust in economic transactions (McAdams and Rasmusen 2007). By examining a large scale intervention, i.e. the launch of a free, accessible, and searchable platform offering legal information – we are able to undertake a large-scale examination of these dynamics.

We also contribute to the growing literature on the impact of digital technologies on markets. In the past three decades, digital platforms have transformed information access across various economic sectors, including meteorology, medicine, and academia. Economic theory suggests that reducing the cost of information gathering enhances decision-making efficiency (Stigler 1961; Diamond 1971; Varian 1980). Recent empirical evidence further confirms that decreased search costs lead to increased transaction volumes and overall market efficiency (Goldfarb and Tucker 2019). While past work has often focused on the impact of commercial platforms such as Amazon and eBay, we focus here on the provision of free legal information. This has been identified as an important priority by many governments and the United Nations (United Nations 2015; OECD 2017). Legal information is particularly important in common-law judicial systems that rely on legal precedent. Here, improved access to legal information can contribute to market development by reducing uncertainty and transaction costs, thus enabling more efficient resource allocation. (Gennaioli and Shleifer 2007).

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. 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.[4]

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

Efforts to digitize legal information in India began about twenty years ago 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 courts in the country. The first phase of the program, which lasted from 2007 to 2015, focused on the computerization

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

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.[5] The top panel of Appendix Figure A1 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, a public sector corporation that is responsible for hosting, maintaining and updating the websites of constitutional, central, state, and local government agencies across India. 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".[6] 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).[7]

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, information about laws was made accessible to a broad range of stakeholders. In his own words, Sinha argues

[5]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 payment platforms for making deposits, payment of court fees, fines, etc. Throughout this period, however, cases have been uploaded to the court websites.

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

[7]Iyengar (2010) interviewed 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, Sinha developed patches for the broader community of software developers. His efforts contributed to an improvement in the 'headline citation' functionality of Postgres, which facilitates the retrieval of contextual information associated with search queries.

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 judgments 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 bottom panel of Appendix Figure A1 provides a summary of when the uploading of cases per high court on the website began.

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 judgments 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 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.

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 Online Appendix Section OA1 of this paper. These resources were different from Kanoon in several ways: they were expensive, difficult to access, and had only partial coverage of Indian law. Ethnographic evidence from the grassroots district courts suggests that in the era of our study, access to these databases was negligible (Krishnan et al. 2014). A survey by a legal think-tank found that even in 2020, only half of Indian advocates at high courts had access to these databases (VIDHI 2020).[8]

Since 2015, all available Indian laws as well as their amendments have been curated into a single repository, 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, still remains difficult to access for those who do not have a legal background.[9]

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.[10]

[8] Sinha was well aware of these issues and described these paid 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."* (https://blog.law.cornell.edu/voxpop/tag/indiankanoon/, accessed on June 23, 2022.)

[9] 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 information as a prerequisite for performing a search raises a big barrier to access by common people" (Private Conversation with Sushant Sinha, February 2022).

[10] A cursory examination of Indian 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.

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 and district courts is around two years, we expect the long-term effects to take at least three years to unfold.

3.1 Courts

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 that are in the jurisdiction of the high courts.[11]

Improved accessibility to legal information has the potential to promptly influence both the perceived likelihood of success and the actual probability of winning a case. Individuals utilizing the platform may be empowered to present more coherent and compelling arguments during legal proceedings. This, in turn, can facilitate judges in reaching decisions more efficiently, thereby reducing the time required for case resolution. However, it is important to acknowledge that increased access to information can also introduce additional complexities to the case, potentially prolonging the review process for judges. As these two mechanisms exert opposing effects on the time taken to resolve cases, the overall impact of Kanoon remains a question that must be examined empirically.

While the records of the 2,800 district courts were not uploaded on Kanoon, we can anticipate the impact of Kanoon on these courts for a variety of reasons. First, the rulings of the high court carry authoritative weight and establish precedents for these district courts. Second, considering that Kanoon may have been particularly beneficial to stakeholders in economically disadvantaged regions of India, litigants in these courts have a lot more to gain from Kanoon than their counterparts in the high courts. In the short term, enhanced

[11] 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) 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 and are thus likely to be affected by the launch of Kanoon.

access to legal information can influence decisions regarding initiating a case or filing an appeal. In the long run, if improved information leads to higher-quality decisions, we can expect a reduction in both the number of appeals and the rate of reversals at the high court.

We can also anticipate different effects at the high courts than the district courts. High courts are located in large cities where there are significant populations of lawyers who may have access to the paid databases and are part of networks of lawyers with considerable legal knowledge. Evidence to date suggests that external legal expenses by firms and wealthy individuals have been on the rise in India (Galanter and Robinson 2013; Sankaraguruswamy and Varottil 2023). These populations may thus stand to gain little from a platform like Kanoon.

For both district and high courts we thus examine case-level outcomes as well as court-level measures of efficiency such as the number of filings, the number of decisions, the number of pending cases, the extent of the backlog and the clearance rate.

3.2 Firms

In business, legal knowledge helps with negotiating contracts, resolving disputes, understanding tax obligations, complying with regulations, and generally, reducing risk and uncertainty in economic transactions. Common knowledge is created when parties know that others have access to the same legal information as they do. This potent knowledge symmetry nurtures efficiency, creating a predictable socio-economic landscape with low transactions costs and higher levels of trust among contracting parties (McAdams and Rasmusen 2007).

Such legal knowledge is particularly important in developing countries like India, where firms face a complex regulatory landscape (Bloom, Eifert, et al. 2013; Ahluwalia 2019). We hypothesize that the arrival of Kanoon reduces "communication frictions" in business activities in India (Guillouet et al. 2021). Before Kanoon was available, lawyers were the only source for legal information and were called upon by managers and accountants for all aspects of legal counsel. Kanoon removes this information bottlenecks. We thus 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 activity of the courts themselves.

4 Data and Descriptive Statistics

Our main analysis relies on the samples drawn from the official eCourts data. The timing of the rollout of the Kanoon platform (Appendix Figure A1) together with the filing dates and judgment dates from these records allow us to estimate the impact of the platform on cases as well as court efficiency.

4.1 Rollout dates

Of the 25 Indian high courts established today, we drop four (Manipur, Meghalaya, Telangana, and Tripura High Court) from our analysis, as these were created during or after our study period. For the remaining high courts, we have case level data for 17 high courts. Of these, six are included on Indian Kanoon in January 2009, one in March 2009, six in June 2010, one in July 2010, one in January 2011, one in April 2017 and one in March 2021.

For district courts, we end up with data from 25 states and union territories. Of these 25, two are under the jurisdiction of a high court included on Indian Kanoon in August 2008, eight in January 2009, four in March 2009, six in June 2010, one in July 2010, two in January 2011, one in April 2011, and one in March 2021.

Appendix Table A1 gives an overview over all high courts, the states they have jurisdiction over, their Kanoon rollout date and whether they are included in our analysis or not.

The left panel of Figure 1 illustrates the rollout timing of the 17 high courts in our sample. High courts can have jurisdiction over more than one state or union territory. This is represented in the right panel of Figure 1, which presents the increasing number of states under the jurisdiction of high courts included on Indian Kanoon over time.

4.2 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, including both criminal and civil cases. For each case, we extracted the date of judgment, court name, and the citations of other cases within the judgment.[12]

[12]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.

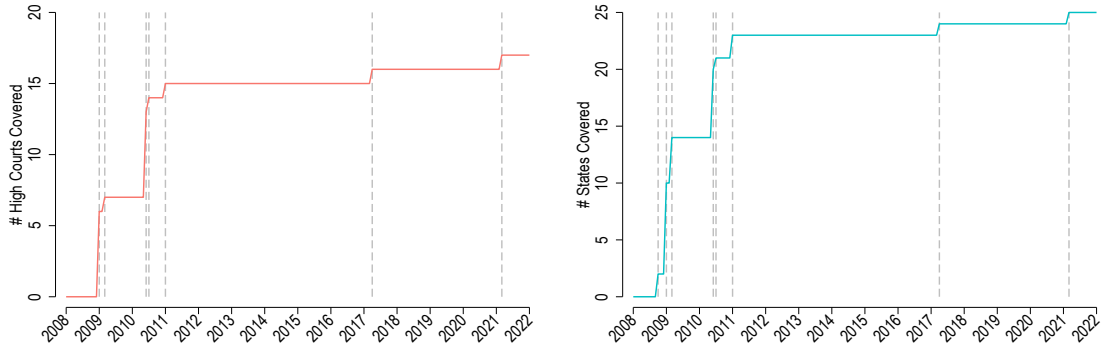


Figure 1: Number of States and High Courts on Kanoon

Note: The left panel displays the cumulative number of high courts included on Indian Kanoon. The right panel displays the cumulative number of states and union territories under the jurisdiction of high courts included on Indian Kanoon. Vertical gray lines represent waves of Kanoon rollout. Appendix Table A1 provides complete information about high courts' inauguration dates, the states and union territories they have jurisdiction over, and their Kanoon rollout dates.

4.3 eCourts Data

For district courts as well as high courts, we scraped 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 eCourts website of an Indian high court, and merged in the Kanoon rollout dates.

For district courts, we focus on all cases that were open at any time between 2005 and 2015. In this period, we observe 43,716,170 filed and 30,303,020 decided cases. For the high courts, we follow a similar approach and observe 6,665,230 filings and 5,274,129 decisions in our sample period.

4.3.1 Data Linkages to Appeal Cases

District court decisions might be appealed in the high court having jurisdiction over the original court. In order to follow the cases from district courts to high courts, we start by associating the records from the lower courts to those of the upper courts (see Online Appendix Section OA2 for a comprehensive explanation of this method). We generate estimates of the total count of cases registered and resolved over each year-month from January 2001 to December 2018 for every state in the district courts data. From the regis-

tered cases, we count the number that are appealed at a high court at any point during our sample period. On average, 0.3% of registered cases are appealed at high courts within a year.[13]

An appeal submitted by a plaintiff progresses through multiple stages of the justice process (see Appendix Figure A2 for an overview of the various stages and potential outcomes of an appeal case.) We employ the courts' own official language to label these stages.[14] Upon its initial filing at the high court, a case is marked as either "Admitted" (all paperwork is in order, and the case is deemed fit for processing at the high court and listed for a hearing/decision) or "Rejected" (the case falls outside the court's jurisdiction—territory or in terms of content—or the paperwork has irreparable defects). If a case is admitted, it can be heard by the high court and then, upon a decision by the bench, it is labeled as "Disposed". The nature of disposal varies: an appeal can be "Overruled" (approved), "Dismissed" by the bench, or "Withdrawn" by the petitioner.[15]

4.4 Prowess

Prowess is a database that curates financial information of around 54,000 listed private and public companies in India. It covers nearly all companies on the National Stock Exchange and Bombay Stock Exchange.[16] These firms account for more than 70% of industrial output and 75% of corporate taxes 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.

To abstract from firm entry and exit, we focus on the panel of 7,522 firms which have

[13] 0.33% cases are appealed within 2 years and 0.35% cases are appealed within 5 years.

[14] These are official terms used in the eCourts system, applied by the e-filing administrator when reviewing the case documents (see https://ecourts.gov.in/ecourts_home/static/manuals/efiling-User-manual.pdf)

[15] In theory, the appellant can decide to withdraw at any time, even during the administrative stage before admission.

[16] 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.

non-zero income data for all years 2005 to 2015. We conduct our analysis on two samples of these firms. First, we use the full panel of firms, regardless of whether the firm has pending cases at the courts. Second, we examine the subsample of these firms that have at least one active case (either pending or new) at any time during the period 2005-2015 in the district courts.

That implied linking the litigants in the eCourts district court database to firms in the Prowess database. Of the 83.8 million cases in the eCourts district court database, we found 718,000 unique cases that matched with at least one firm in the Prowess database. This corresponds to 9,246 unique Prowess firms - around 18% of all the firms in the Prowess data. For the period 2005-2015, we end up with a panel of 1,771 firms which are litigants in at least one active district court case and which have non-zero income data for all financial years of the time period.

Summary statistics of all key variables from these different datasets are presented in Tables A2 and A3. 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

To estimate the impact of Kanoon on court efficiency, appeals, and firm financials, we use an event study framework that exploits the exogeneity of the timing of the Kanoon rollout. A high court, state, case, or firm is considered treated from the date on when the high court's (or the high court having jurisdiction over the state) cases are uploaded on Kanoon. We refer to this event as the "Kanoon rollout date" for that high court. On this date, all cases that were filed *prior* to the rollout date *and* available for review on the high court websites are available in an accessible format. For all the dates that follow, cases are added to Kanoon on the date that they appear on the court websites.

We estimate the following two-way fixed effects (TWFE) model at the state or high court level:

$$Y_{cym} = \alpha + \beta \text{Kanoon}_{cym} + \gamma_c + \delta_y + \eta_m + \varepsilon_{cym} \tag{1}$$

where Y_{cym} is an observed outcome of state or high court c in year y and month m. $Kanoon_{cym}$ is a dummy variable that takes value 1 if the Kanoon rollout of high court c (or the high court having jurisdiction over state c) happened before year-month ym (and 0 otherwise). γ_c, δ_y, and η_m represent year, high court (or state), and month fixed effects. Standard errors

are clustered at the high court-level.

The staggered nature of the treatment produces a complex control group whose population changes over the period of the study. As Kanoon rolls out over the country, courts move from the "untreated" to "treatment" category at different times. Recent literature suggests that in this scenario, coefficients in the standard TWFE model may not represent a simple weighted average of unit-level treatment effects. The problem of "forbidden comparisons", i.e. treated units are inappropriately compared with those untreated and some groups may even receive negative weights, leading to a reversal of the sign of the coefficient relative to estimates seen in an event-study plot (Roth et al. 2023). Our remedy for this issue is to rely on the estimator from Borusyak, Jaravel, and Spiess (2021).[17] The estimator exploits the known and specified structure of the treatment to draw comparisons between treated and never treated groups (where never treated in our setting is equivalent to a Kanoon rollout post 2015), and then aggregates them using user-specified weights to estimate the impact of the legal platform.[18]

We examine several types of outcome variables. First, we consider four district court efficiency parameters aggregated at the state-year-month level: number of filings, number of decisions, backlog (number of unresolved cases older than one year) and disposition time (ratio between the number of unresolved cases and the number of decisions, a measure of how many years it takes on average to resolve a case.)[19] For these analyses, Kanoon rollout is defined with respect to the high court which has jurisdiction over the state.

Second, we study the same outcomes at the high court-year-month level, to estimate the impact of Kanoon's rollout on high court efficiency.

Third, we study the impact on the number and share of district court cases appealed in high courts and the impact on their case status and outcome in the high courts.

In order to study the dynamic effects Kanoon had on the different outcomes and to test for parallel trends, we also estimate the following event-study:

$$Y_{cym} = \alpha + \sum_{j=-4}^{5} \beta_j \text{Kanoon}_{c,y+j,m} + \gamma_c + \delta_y + \eta_m + \varepsilon_{c,y,m} \tag{2}$$

[17] We use the STATA 17 package *did_multiplegt*.

[18] See Online Appendix Tables OA3 and OA4 for robustness checks, where we compare our baseline estimates with simple OLS estimations and two alternative estimators proposed by Callaway and Sant'Anna (2021).

[19] In results available upon request, we also explore the impact of Kanoon on the mean age of decided cases, the mean age of pending cases, and the clearance rate.

where variables are defined as in Equation 1 and $Kanoon_{c,y+j,m}$ is a set of dummy variables equal to 1 if the Kanoon rollout of court c happened j years before (or after) year-month ym (and 0 otherwise).

Estimating this equation via OLS would imply relatively strong assumptions about homogeneity in treatment effects to deliver consistent results. If Kanoon has larger effects for states that are treated early, then as the early-treated become controls for the later-treated, the treatment effects can be biased downwards. To address this identification problem, we draw on new econometric advances. 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).[20] This method is similar to other proposed corrections in recent literature (see for example Callaway and Sant'Anna 2020, De Chaisemartin and d'Haultfoeuille 2020 and Goodman-Bacon 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.

Lastly, we inspect Kanoon's downstream impacts on firms and estimate a yearly model (and, similarly, an event study) on a panel of Indian firms:

$$Y_{fcy} = \alpha + \beta \text{Kanoon}_{fcy} + \delta_y + \eta_f + \varepsilon_{fcy} \tag{3}$$

where Y_{fcy} is a financial indicator of firm f based under the jurisdiction of high court c for financial year y. Kanoon_{fcy} is a dummy variables equal to one if financial year y began after court c's Kanoon rollout. δ_y and η_f are financial year and firm fixed effects, respectively.

The coefficient(s) of interest in all three settings is either β or the set of β_j. In the dynamic estimations using the method from Sun and Abraham (2021), $\text{Kanoon}_{c,y-1,m}$ is excluded and results should be interpreted with respect to this baseline. In order to interpret the coefficients β and $\beta_0, \ldots, \beta_5$ as the average treatment effects on the treated (ATT) of the

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

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 discuss possible threats to the identification of causal effects, 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, year, and month fixed-effects rules out the possibility that Kanoon's observed impact is induced by high court (or state) characteristics, temporal trends such as macroeconomic fluctuations, changes in internet regulations, digital privacy laws, or seasonal factors.

Causal identification of the impact of Kanoon on outcomes, however, hinges on the independence of the platform'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 (Appendix Figure A1). Second, as discussed in Section 2.2, the eCourts web hosting program that permitted users to retrieve information about pending cases was launched after the years that we are focusing on in this analysis. In the period that we are studying, the only information that was available to the public was that content of high court websites, which varied in content and coverage but was always restricted to information about past cases.

Additionally, in Online Appendix Figure OA1 we plot the evolution of the number of filings in district courts over time separately for each of the 25 states and union territories included in our analysis. We note a general pattern of an increase in filings over time, independently of the Kanoon rollout of the high court having jurisdiction over the state or union territory. This strengthens our trust, that although some artifacts of the data are sensitive to an increased digitalization of courts over time, these are similar across states and independent of Kanoon's availability.

Finally, we also face the possibility that the order of expansion across states may be correlated with other confounding factors, creating the challenge of omitted variables and the possibility of pre-trends. We remedy this by constructing event-study plots and examining the evidence for pre-trends, as recommended by Roth et al. (2023). 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 additional exercises suggested by De Chaisemartin and d'Haultfoeuille (2020). These 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 estimates in the Appendix. Both sets of analyses lend confidence to our causal inference.

6 Results

We present and interpret our baseline estimates of the causal effect of Kanoon in three broad areas: access to justice, efficiency of courts, and firm outcomes.

6.1 Impacts on District Courts

We find that Kanoon has substantial positive impacts on access to justice. We consider first the impacts on the district courts (Table 1). Note that there is an observed increase of approximately 6.9 thousand filings per state and month during the post-Kanoon period. This is equivalent to a 52% increase relative to the overall sample mean, and the effect remains stable and statistically significant in all the specifications. We also see a significant increase in the number of decisions by 8.3 thousand cases (90% increase). Panel A of Online Appendix Table OA3 displays estimation results from other estimators and shows that results are robust to the choice of estimator.

While the case volume in the district courts is increasing, it is intriguing to note that there is only a marginal rise in court backlog (i.e., the number of cases pending for over a year) and a slight, statistically insignificant reduction in case disposition times. As district courts are typically the initial judicial recourse for most original litigation, these findings imply that improved access to court data may empower more parties, especially in small claims, to file original cases without significantly affecting the overall pace of case pro-

Table 1: Kanoon's Impact on District Court Efficiency

	Filings			Decisions	Backlog	Disposition Time
	(1)	(2)	(3)	(4)	(5)	(6)
Post-Kanoon introduction	6.856***	6.929***	6.929***	8.256***	105.593*	-2.221
	(1.940)	(1.942)	(1.942)	(2.151)	(56.566)	(12.558)
Year FE	✓	✓	✓	✓	✓	✓
Month FE		✓	✓	✓	✓	✓
Rollout FE	✓	✓				
State FE			✓	✓	✓	✓
Mean Dep. Var.	13.247	13.247	13.247	9.183	386.713	24.414
Observations	3,300	3,300	3,300	3,300	3,300	2,868

Note: This table displays estimation results of Indian Kanoon's impact on court efficiency in district courts. Observations are at the state-year-month level. The data is based on all cases active in the period 2005-2015 and available on eCourts District Court in May 2019. Every estimation is regressing a measure of court efficiency on a dichotomous variable, equal to one if the year-month is post the Indian Kanoon rollout for the respective high court, and several fixed effects. *Filings* counts the number of filings per state-year-month (in thousand). *Decisions* counts the number of decided cases per state-year-month (in thousand). *Backlog* counts the number of undecided cases in a state at the end of a year-month, and which were filed at least one year prior to this year-month (in thousand). *Disposition Time* is the ratio between the number of unresolved cases at the end of a year-month and the number of decisions in that year-month in a state, and is a measure of how many years it takes on average to resolve a case. 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.

cessing.

The dynamic effects of Kanoon on district courts are presented in Figure 2. We observe that the number of filings and decisions at the district courts increases steadily in the four years after the rollout of Kanoon. The overall positive effects noted in Table 1 are largely driven by the positive and significant effects in the second, third, and fourth year of the launch of the platform. At its peak, four years after the launch, we observe an additional 10 thousand filings and 12 thousand decisions per state, year, and month. We see no statistically significant impacts in any year on the disposition time. There is a small and statistically significant increase in backlog in the second year of the rollout.

In summary, we see strong impacts of the Kanoon platform on access to district courts without adverse impacts on their efficiency. The analogous results for high courts will be examined later in this section. For now, we turn to the impacts of the platform on a specific category of cases that originate in the district courts: appeals. A closer look at the trajectory of appeals in the Indian justice system before and after the rollout provides insight into the question of what drives the higher case volume, and whether there are changes in the case "quality".

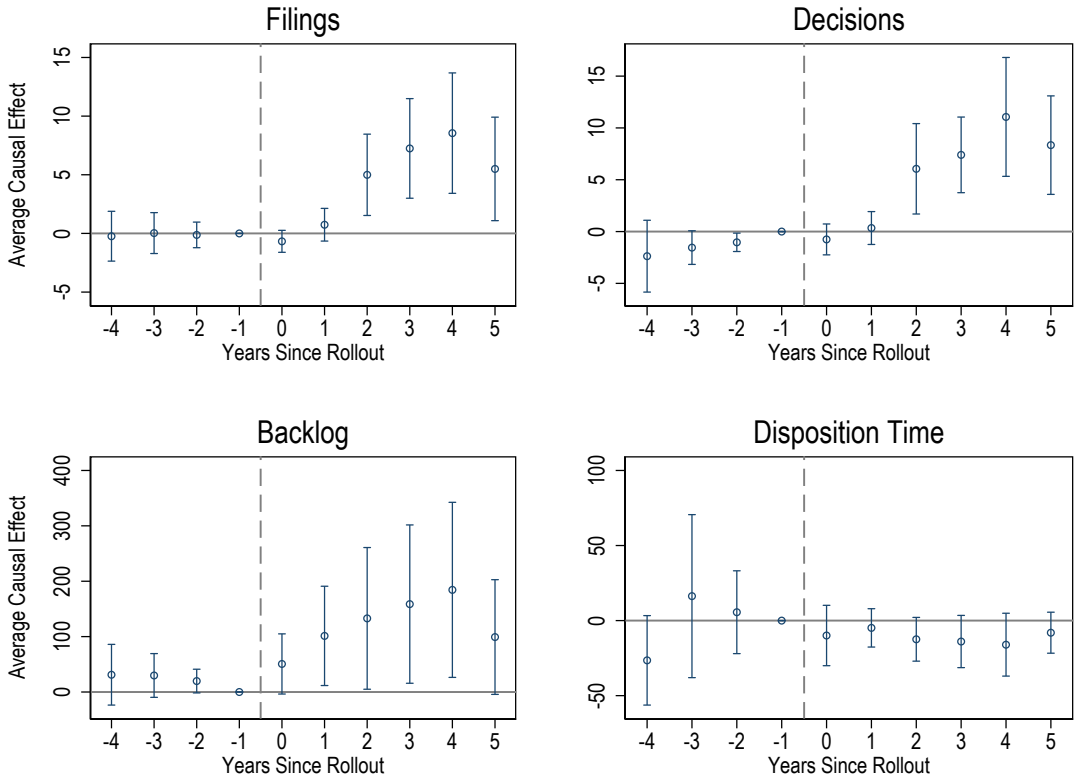


Figure 2: Dynamic Impact of Kanoon Rollout on District Court Efficiency

Note: This figure presents dynamic estimation results of the impact of Indian Kanoon on court efficiency in district courts. Observations are at the state-year-month level. The data is based on all cases active in the period 2005-2015 and available on eCourts District Court in May 2019. Leads and lags are defined as the yearly difference between the year and month of an observation and the year and month of the rollout of the state's high court on Indian Kanoon. *Lead 1* is the omitted category. *Filings* counts the number of filings per state-year-month (in thousand). *Decisions* counts the number of decided cases per state-year-month (in thousand). *Backlog* counts the number of undecided cases in a state at the end of a year-month, and which were filed at least one year prior to this year-month (in thousand). *Disposition Time* is the ratio between the number of unresolved cases at the end of a year-month and the number of decisions in that year-month in a state, and is a measure of how many years it takes on average to resolve a case. Any court that has not been rolled out by 2017 is used as control cohort. The estimation controls for state, year, and month fixed effects and is robust to heterogeneous treatment effects. Standard errors are clustered at the state level. Confidence intervals are at the 95% level.

Table 2: Kanoon's Impact on Appeals of District Court Cases in High Courts

	(1) Appeals	(2) % Appealed
Post-Kanoon introduction	.044**	.0011
	(.019)	(.0023)
Year FE	✓	✓
Month FE	✓	✓
State FE	✓	✓
Mean Dep. Var.	.0972	.00706
Observations	1,980	1,980

Note: This table presents estimation results examining the impact of Indian Kanoon on appeals of district courts cases in high courts. The observations are at the court-year-month level. The data includes all cases active during the period 2005-2015, as obtained from eCourts District Court in May 2019. Each estimation regresses a measure of court efficiency on a dichotomous variable, which equals one if the year-month is after the Indian Kanoon rollout of the high court with jurisdiction over the state, along with several fixed effects. *Appeals* represents the number of district court cases filed in a state-year-month and later appealed in the high court (in thousand). *% Appealed* represents the share of district court cases filed in a state-year-month that are appealed later on in the high court. Standard errors are clustered at the high court level. Stars (**, **, ***) indicate that the p-value is below 0.1, 0.05, and 0.01, respectively.

6.2 Appeals from District Courts at High Courts

We follow appeal cases from the district courts through the high courts in a linked dataset described in section 4.2.1. The first step is to examine the number of appeals that emanate from the district courts. Table 2 presents the impact of Kanoon on the number of cases appealed (Column 1) as well as percentage (Column 2). On average, we note a statistically significant uptick of 44 appeals (or 45%) per state, year, and month after the Kanoon rollout (Column 1). We do not observe a significant increase in the proportion of district court cases appealed to high courts (Column 2). The dynamic estimates, as presented in Online Appendix Figure OA2, follow a similar pattern as shown in the district court measures discussed in the previous section, i.e., we see a gradual increase in the number of appeals over the initial years, with particular prominence achieved during years 3 to 5 after the Kanoon rollout. There is, however, no sustained statistically significant increase in the percentage of cases that are appealed in this period. Except for a small increase in the first year after the rollout, Kanoon is thus not changing the likelihood of any litigant filing an appeal in this era.

Table 3: Impact on Outcomes of District Court Case Appeals in High Courts (in %)

	% of Appeals	% of Admitted	% of Disposed	
	(1) Admitted	(2) Disposed	(3) Withdrawn	(4) Overruled
Post-Kanoon introduction	.0013	-.00037	.026***	-.13***
	(.028)	(.00033)	(.0096)	(.028)
Year FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
State FE	✓	✓	✓	✓
Mean Dep. Var.	.83	.999	.0795	.25
Observations	1,519	1,507	1,507	1,507

Note: This table displays estimation results of the impact of Indian Kanoon on outcomes of district courts cases appealed in the high courts. Observations are at the court-year-month level. The data is based on all cases active in the period 2005-2015 available on eCourts District Court in May 2019 and appealed later on in the high court. Every estimation is regressing a conditional share of cases with a given appeal status on a dichotomous variable, equal to one if the district court filing date is post the Indian Kanoon rollout of the high court having jurisdiction over the state, and several fixed effects. *Admitted* measures the share of all district court appeal cases for which all paperwork is in order, and the case is deemed fit for processing at the high court and listed for a hearing/decision. *Disposed* measures the share of *admitted* appeal cases which is heard by the high court and decided by the bench. *Withdrawn* measures the share of *disposed* appeal cases withdrawn by the petitioner. *Overruled* measures the share of *disposed* appeal cases (at least partly) approved by the high court. 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.

Did the increase in litigation volume following the Kanoon rollout coincide with a decline in the "quality" of cases, meaning a higher likelihood of dismissal or rejection when appealed in the high courts? To examine this, we analyze how these cases fare upon entry into the high court, where they may either be admitted, rejected, or transferred. Our findings indicate that the vast majority of appeal cases (approximately 83%) are admitted, meaning they successfully pass the initial stage in the high courts and proceed for further consideration. In simpler terms, despite Kanoon's notable impact on increasing case volumes (as shown in column (1) of Table 3), there is no statistically significant impact on the proportion of appeals that are admitted.

This suggests that the increase in litigation quantity following the rollout of Kanoon does not lead to a deterioration in the quality of cases being appealed in the high courts. The majority of cases continue to be deemed worthy of further consideration, despite the

higher number of cases being filed overall.

Next, we examine the impact of Kanoon on the outcomes of the appeals at the high court. As explained earlier, admitted cases can be disposed in three ways: "dismissed" (the district-court decision stands), modified or "overruled" (the district court decision is modified) or "withdrawn" (the litigant decides to not pursue the case and the original district court decision stands). To study these outcomes, we construct conditional percentages along the appeal process in each node of this tree.[21][22]

Among the admitted appeals, we find no changes in the likelihood of disposal (as indicated in column (2) of Table 3) in the post-Kanoon period. However, we do find significant effects on the *type* of disposals. Specifically, the rollout of Kanoon appears to have increased case withdrawals by 2.6 percentage points, while simultaneously decreasing the proportion of cases (at least partly) overruled by 13 percentage points. These findings suggest that litigants still get their opportunity for a fair hearing in court; however, some opt to withdraw their cases during the course of proceedings. Furthermore, the decrease in the share of cases overruled implies that the district court decisions were not of lower quality. Instead, this outcome suggests that Kanoon may have made appellants more aware of their chances of success, leading some to withdraw their cases in anticipation of an unfavorable ruling. In essence, the implementation of Kanoon seems to have facilitated better-informed decisions among litigants, impacting the types of disposals while maintaining access to justice for all parties involved.

The event study plots for these outcomes are presented in Online Appendix Figure OA3. Note that there is a decrease in the likelihood of being admitted in the first three years after the rollout, but this decline appears to be transitory. For admitted cases, disposed cases, reversals and withdrawals, the dynamic results in Online Appendix Figure OA3 are also consistent with the overall impact in Table 3.

Overall, 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. Conditional on being admitted, however, litigants may decide against the continuation of the legal process more than they did before they had access to the Kanoon platform.

[21] Conditional in the sense that the denominator for each case status is the last node in Figure A2. For the variable % Admitted, 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 two variables % Withdrawn and % Overruled, it is the number of disposed cases.

[22] See Online Appendix Figure OA4 for the results on the total quantity of appealed cases by outcome.

Table 4: Kanoon's Impact on High Court Efficiency

	(1) Filings	(2) Decisions	(3) Backlog	(4) Disposition Time
Post-Kanoon introduction	-1.165**	-1.028***	20.864**	-0.013
	(0.549)	(0.294)	(8.149)	(0.905)
Year FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
State FE	✓	✓	✓	✓
Mean Dep. Var.	3.144	2.461	74.225	0.799
Observations	2,120	2,143	2,187	2,142

Note: This table displays estimation results of Indian Kanoon's impact on court efficiency in high courts. Observations are at the high court-year-month level. The data is based on all cases active in the period 2005-2015 and available on eCourts High Court in November 2020. Every estimation is regressing a measure of court efficiency on a dichotomous variable, equal to one if the year-month is post the Indian Kanoon rollout for the respective high court, and several fixed effects. *Filings* counts the number of filings per state-year-month (in thousand). *Decisions* counts the number of decided cases per state-year-month (in thousand). *Backlog* counts the number of undecided cases in a state at the end of a year-month, and which were filed at least one year prior to this year-month (in thousand). *Disposition Time* is the ratio between the number of unresolved cases at the end of a year-month and the number of decisions in that year-month in a state, and is a measure of how many years it takes on average to resolve a case. 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.

6.3 Impacts on High Courts

Next, we examine the impact of Kanoon on high courts. Given that we already noted an increase in case-load at the district courts, it is reasonable to speculate that the high court case-load would also be increasing in the aftermath of Kanoon. Appeals, however, account for only 41.6% of the case-load in our data (see footnote 11). In other areas of court jurisdiction (such as "original jurisdiction"), it is unlikely that Kanoon had significant impacts on litigants. Such cases are typically contested by urban lawyers who maintain close affiliations to the courts, have greater access to (paid) digital databases and maintain thick professional networks for the exchange of legal information Galanter and Robinson 2013). Legal professionals in this domain likely already had robust mechanisms to access legal information at the time of the Kanoon rollout.

Table 4 presents estimates of the number of cases filed, decisions made, backlog and disposition time at the high courts. In contrast to the increased case load at the district

courts, here we observe a decrease in the number of cases filed and decisions made in all specifications. On average, the Kanoon rollout led to 1.2 thousand fewer cases filed per state, year and month. Similarly, it reduced decisions by one thousand cases. Unlike for district courts, we observe an increase in backlog in the high courts: the Kanoon rollout led to a statistically significant increase in the stock of pending cases older than one year (backlog) by 28%. Panel B of Online Appendix Table OA3 shows that results are robust to the choice of different estimation techniques.

The results of the estimation of the dynamic impacts of Kanoon on the high courts are presented in Figure 3. We note an overall downward trend in the estimate of Kanoon on the number of filings. Backlog increases over the years after the Kanoon rollout, peaking within three to four years after the rollout, where there are additions of approximately 20,000 cases per court per month. Here too, we see no significant change in the disposition time.

What drives the decline in cases at the high courts? Appendix Table A4 presents the impacts of Kanoon on the different case-types at these courts: criminal, civil, writs, appeals, and "other".[23] We note that there is a negative and statistically significant effect on the number of civil cases. This is, for the most part, driving the overall decline.

But then the next question is why Kanoon would induce such a decline in the number of civil cases filed. Here, we speculate that the Kanoon may help citizens take greater advantage of the policy reforms that were being undertaken to reduce the backlog in the justice system throughout this time. The judiciary was actively promoting the establishment of tribunals (for example, the debt recovery tribunal and the green tribunal), alternate dispute resolution systems (for example, *Lok Adalats*) and the launch of digital services that were discussed earlier in this paper. The arrival of free legal information for citizens may have reduced communications and information frictions in the rollout of these platforms. Kanoon may have been able to reduce jurisdictional confusion (Krishnan et al. 2014).[24] It may

[23] Criminal, civil, writ and appeal are not mutually exclusive case types. E.g., a "criminal appeal" is counted as both criminal and appeal. The category "other" includes cases that can not be categorized in any of the other categories (example: judicial section cases that pertain to personell decisions within the judiciary, or judicial review activities).

[24] The issue of the overlapping jurisdictions of the high courts, tribunals and subordinate courts has been a frequent issue of contention in the judiciary due to a lack of clarity in the Indian Constitution. Article 226 for example, gives High Courts the ability to issue instructions, orders, and writs to any person or authority, including the government. Article 227 gives High Courts the power of superintendence over all courts and tribunals in the territory over which they have jurisdiction. The precise meaning of the term “superintendence" was intensely debated for years (Supreme Court Online 2021). Kanoon may have provided some clarity to confused citizens and deterred them from filing cases in the High Courts.

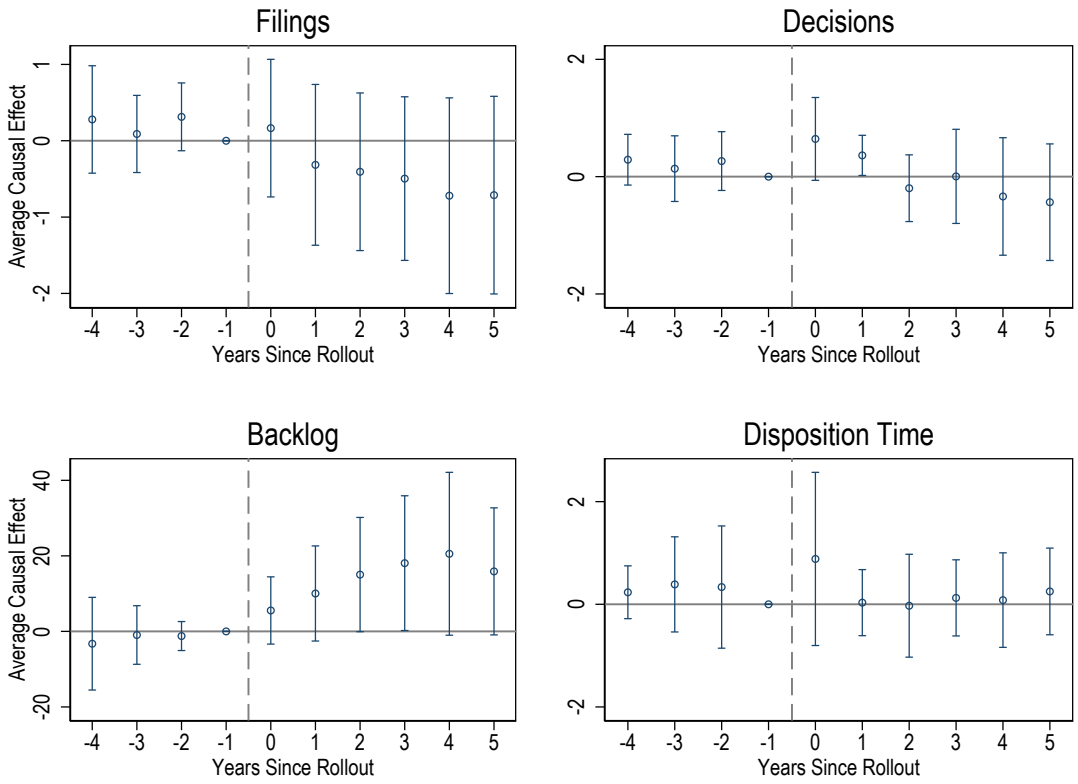


Figure 3: Impact of Kanoon Rollout on Aggregate Efficiency Measures of High Courts

Note: This figure presents dynamic estimation results of the impact of Indian Kanoon on overall court efficiency in high courts. Observations are at the court-year-month level. The data is based on all cases active in the period 2005-2015 and available on eCourts High Court in November 2020.The leads and lags are defined as the yearly difference between the year and month of an observation and the year and month of the rollout of the state's high court on Indian Kanoon. *Lead 1*, 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. *Filings* counts the number of filings per state-year-month (in thousand). *Decisions* counts the number of decided cases per state-year-month (in thousand). *Backlog* counts the number of undecided cases in a state at the end of a year-month, and which were filed at least one year prior to this year-month (in thousand). *Disposition Time* is the ratio between the number of unresolved cases at the end of a year-month and the number of decisions in that year-month in a state, and is a measure of how many years it takes on average to resolve a case. The estimation uses the *eventstudyinteract* command in STATA to estimate consistent and heterogeneous treatment robust average dynamic treatment effects. Any court that has not been rolled out by 2017 is used as control cohort. The estimation controls for state, year, and month fixed effects. Standard errors are clustered at the state level. Confidence intervals are at the 95% level.

also deter citizens from "forum shopping", i.e. intentionally filing cases in high courts with the expectation of gaining advantages, such as faster resolution, favorable judgments, or strategic reasons, despite the case being more appropriately handled in a lower court. Our current event study framework does not provide us with the flexibility to test this hypothesis, through, for example, constructing the interaction of Kanoon and these interventions, but we believe this is an important and fascinating area for additional research.

We also conducted robustness tests of all the event study estimations above by estimating them using the method from De Chaisemartin and d'Haultfoeuille (2020). Results are presented in Online Appendix Figures OA8, OA9 and OA10. The results suggest the absence of the pre-trends for all the outcome variables and display similar effects post-Kanoon rollout.

6.4 Impacts on Firms

In this last part of our analysis, we study Kanoon's impact on firms. As conjectured earlier, free legal information within a firm can facilitate business activity by enabling smoother contract negotiation, dispute resolution, tax compliance, and regulatory understanding. This clarity breeds trust and predictability in commercial relationships, increasing "common knowledge" among market participants. To capture these broad effects, we consider the simplest possible measures of the financial well-being of firms that may be affected by improved access to legal information across the markets in which they operate: income, assets, and expenditures. To capture more direct effects, we also present results on the presence of external legal charges, i.e. legal expenses that exclude in-house legal counsel services and pertain to services outside the firm.

We restrict our sample to firms for which we have financial information for all years 2005-2015 and, therefore, abstract from entry and exit decisions of firms. We also emphasize here that the firms in the Prowess database are *not* representative of all Indian firms – they mostly represent large firms in the formal sector that have financial information available for all the years we are considering here.[25]

We analyze the impact of Kanoon on firms in two ways. First, we estimate the impact of Kanoon on all firms in our sample, regardless of whether they had any cases within

[25] In our sample of 5,240 firms, analysis of national industry classification codes suggests that 2,190 (41.79%) are in the manufacturing sector, 1,126 (21.49%) are in the financial sector and the remaining 24% of firms are spread across a variety of sectors that include information technology (5.59%), construction (4.4%), administrative services (3.36%), transportation (2.08%), etc.

Table 5: Effects on Firm Financials

	(1) Income	(2) Assets	(3) Expenditures	(4) Legal Charges
Panel A: Partial Equilibrium				
Post-Kanoon introduction	5.261***	13.026***	5.393***	0.005***
	(1.309)	(2.344)	(1.267)	(0.001)
Mean Dep. Var.	10.89	20.86	10.33	0.01
Observations	19,481	19,481	19,481	18,689
Panel B: General Equilibrium				
Post-Kanoon introduction	2.349***	5.730***	2.452***	0.004**
	(0.790)	(1.756)	(0.772)	(0.002)
Mean Dep. Var.	6.05	11.87	5.72	0.01
Observations	82,555	82,555	82,555	78,639
Year FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓

Note: This table display estimation results of the impact of Indian Kanoon on several firm financials. Observations are at the firm-financial year level, and outcomes are measured in a billion INR. The data includes firms with non-zero income data available for all years 2005-2015 in the Prowess database. Every estimation is regressing a financial variable on a dichotomous variable, equal to one if the financial year started post the Indian Kanoon rollout of the high court having jurisdiction over the state the firm is registered in, and year and firm fixed effects. A financial year is defined as April 1st to March 31st of the following year. Panel A restricts the sample to firms which had at least one active case in the period 2005-2015. Panel B includes all firms. Estimates are robust to heterogeneous treatment effects. 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.

our considered time frame. We refer to this as a "general equilibrium" effect.[26] Then, we restrict our analysis to firms having at least one case at the courts during the 2005-2015 time period. We refer to this as the "partial equilibrium" analysis. For this, we linked firm-level data to the legal data by matching the eCourts district court case data to the firms in the Prowess database.

In each regression, we note that the unit of observation is a financial year of a firm.

[26]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 that are not directly involved in litigation and therefore impacted by Kanoon.

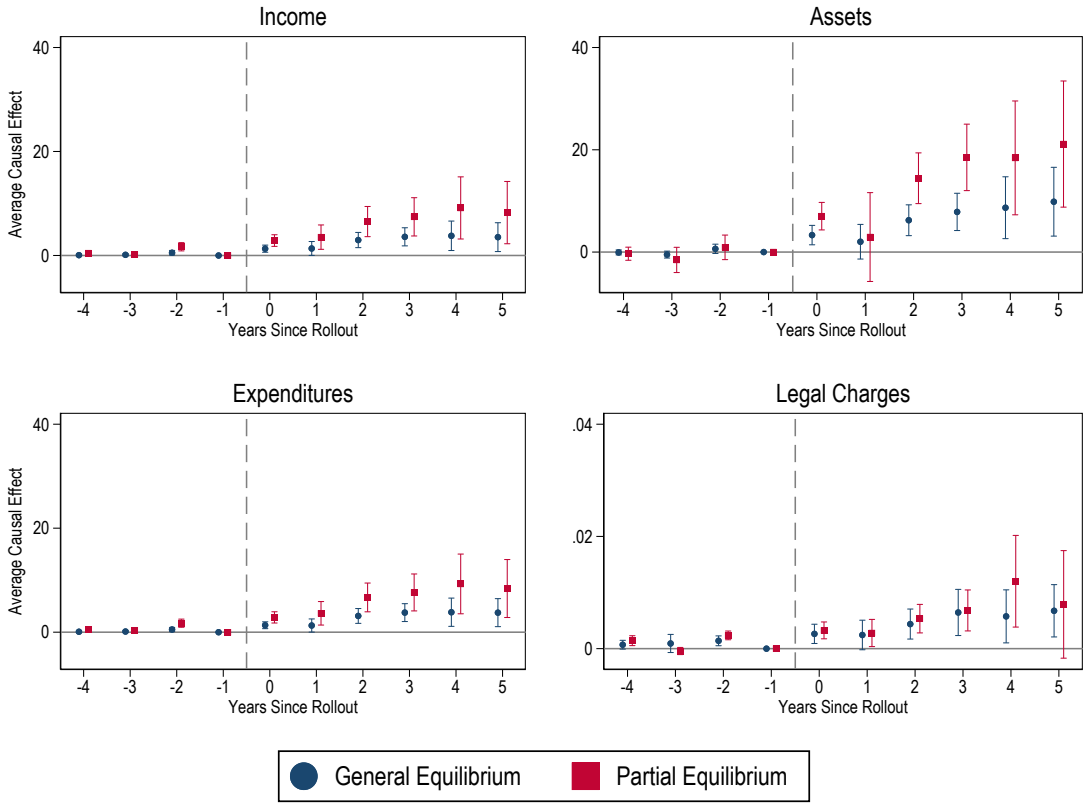


Figure 4: Effects of Kanoon rollout on Firm Financials

Note: This figure presents dynamic estimation results of the impact of Indian Kanoon on firm outcomes for all firms (General Equilibrium) and firms having at least one district court case during the considered time frame (Partial Equilibrium). Observations are at the firm-fiancial year level and outcomes are measured in billion Indian rupees. The data includes firms with non-zero income data available for all years 2005-2015 in the Prowess database. Leads and lags are defined as the yearly difference between the start of the financial year and the date of rollout on Indian Kanoon of the firm's registration state's high court. *Lead 1*, the dummy variable which is equal to one if the Kanoon rollout falls into the financial year, is the omitted category. Any court that has not been rolled out by 2017 is used as control cohort. The estimation controls for firm and year fixed effects. Standard errors are clustered at the state level. Confidence intervals are at the 95% level.

Panel A of Table 5 presents the estimates for the "partial equilibrium" impacts, this is, the impact on firms that had at least one identified active case in a district court at any time during the period 2005-2015. We observe that income increased by 5,261 (48%), assets by 13,026 (62%), and expenditures by 5,393 (54%) million Indian rupees (INR).[27] We also observe a 38% increase in legal charges for these firms.

Panel B of Table 5 displays the overall effect Kanoon had on the "general equilibrium" outcomes of firms. We observe smaller impacts than the partial equilibrium estimates. On average, income increased by 2,349 million INR (39%), assets by 5,730 mill. INR (48%), expenditures by 2,452 mill. INR (43%), and legal charges by 4 mill. INR (40%).

Figure 4 displays the dynamic effects of Kanoon on these firm financial variables. Note that in both general and partial equilibrium settings, the Kanoon rollout had large and significant effects on firms' financial variables, increasing over the first three years. The effect on assets is steadily increasing over the first years after Kanoon rollout before leveling off at a longer term effect around 10,000 (general equilibrium) and 20,000 million INR (partial equilibrium).

Table 6 presents supplementary estimates derived from subsamples of firms across various sectors. Notably, the manufacturing sector (column 1) experiences the most robust impacts of the rollout of Kanoon. We see that income, assets, and expenditures all witness a substantial increase of around 60% in the aftermath of the rollout. The financial sector (column 2) also shows significant effects, albeit with lower level effects compared to manufacturing, yet proportionately significant. Outside these two sectors, we find that the rollout of Kanoon does not exhibit any statistically significant impact. It is worth noting that manufacturing and financial services remain among the most heavily regulated sectors in India, although there has been some relaxation of regulations since the pro-market reforms of 1991 (Ahluwalia 2019). The similarity in the impacts on these two sectors aligns with recent literature that underscores the close linkages between them: evidence indicates that capital market liberalization has facilitated the improvements in manufacturing efficiency during the post-reform years (Arnold et al. 2016; Bau and Matray 2023).

In Appendix Table A5 we investigate in deeper detail how the Kanoon rollout impacted firms' legal charges. Especially, we study whether the observed impact on legal charges is driven by changes in the intensive margin (the amount of legal charges if you have any) or in the extensive margin (the probability of having any legal charges). For firms with at least one active case in our database during the study period, we find an increase of 35%

[27] 5,393 million INR corresponds to approximately 65 million dollars.

Table 6: Heterogeneous General Equilibrium Effects on Firm Financials by Sector

	(1) Income	(2) Assets	(3) Expenditures
Panel A: Manufacturing Sector			
Post-Kanoon introduction	5.025***	5.035***	4.973***
	(1.780)	(1.401)	(1.748)
Mean Dep. Var.	9.12	8.16	8.78
Observations	32,615	32,615	32,615
Panel B: Financial Sector			
Post-Kanoon introduction	1.446***	12.713**	1.214***
	(0.481)	(5.226)	(0.450)
Mean Dep. Var.	2.21	19.68	1.89
Observations	19,844	19,844	19,844
Panel C: Other Sectors			
Post-Kanoon introduction	0.577	1.552	1.023*
	(0.550)	(1.232)	(0.548)
Mean Dep. Var.	5.18	9.39	4.86
Observations	30,096	30,096	30,096
Year FE	✓	✓	✓
Firm FE	✓	✓	✓

Note: This table display estimation results of the impact of Indian Kanoon on several firm financials by firms' sector. Observations are at the firm-financial year level, and outcomes are measured in billion Indian rupees. The data includes firms with non-zero income data available for all years 2005-2015 in the Prowess database. Every estimation is regressing a financial variable on a dichotomous variable, equal to one if the financial year started post the Indian Kanoon roll-out of the high court having jurisdiction over the state the firm is registered in, and year and firm fixed effects. A financial year is defined as April 1st to March 31st of the following year. Sectors are defined by divisions of the firms' National Industrial Classification code. "Financial" is short for "financial and insurance activities". "Other" encompasses every division except "manufacturing" and "financial". Estimates are robust to heterogeneous treatment effects. 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.

increase in the intensive margin of legal charges and an 8.7 percentage point increase in the probability of having any legal charges. Turning to all firms, we find a 25% increase on the intensive margin and a zero or at most small effect on the extensive margin. Appendix Figure A3 shows that these effects are increasing over time.

Online Appendix Figure OA5 presents additional results on Kanoon's impact on court-level efficiency measures for cases involving firms from the Prowess data. We note here a discernable decline in the disposition time of cases in the third and fourth year of the rollout.

We also present placebo tests for the partial and general equilibrium estimations and the court-level efficiency estimates (see Online Appendix Table OA4 and Figures OA11, OA12 and OA13). These results strengthen our confidence in the absence of pre-trends.

To interpret the impacts on firm financial indicators, we draw on the growing literature that demonstrates the substantial impact of providing relevant legal or management information on the productivity of firms in developing countries. In a recent randomized controlled trial that offered firm management consulting services in India, Bloom, Manova, et al. (2021) note a remarkable 200% increase in firm assets (via the establishment of new plants). In South Africa, the provision of legal information to firms increased employment by 12% over just a six-month period (Bertrand and Crépon 2021). We complement these findings by studying when all citizens get access to a free online legal search platform. We substantiate the widely held belief that the absence of common knowledge about laws and regulations levies both direct and secondary costs on economies (McAdams and Rasmusen 2007). Our intervention, consisting of a free, accessible, and searchable platform offering legal information, offers an unprecedentedly large-scale examination of these dynamics, bypassing the typical constraints associated with most RCTs.

7 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 judgments 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 individual cases, courts, and firms. Our working sample exists of all district court and high court cases active during the period 2005-2015.

Our findings are threefold, encompassing access to justice, judicial efficiency, and the financial status of firms. Firstly, Indian Kanoon's rollout corresponded with an uptick in case filings and resolutions, without exacerbating backlogs or prolonging case disposition times. Importantly, we observed no decline in decision quality, as indicated by appeals to high courts and their subsequent outcomes.

At the high court level, complementary trends emerged, with noticeable reductions in case filings and decisions, primarily in original civil litigation. Here, Kanoon appears to have curbed erroneous filings, encouraged alternative dispute resolutions, and reduced forum-shopping, thereby enhancing judicial efficiency.

Finally, our analysis indicates beneficial effects on the financial health of large formal firms. We noted a robust positive correlation between the platform's introduction and improvements in firms' income, assets, and expenditures, consistent across the full spectrum of firms and those involved in district court litigation during our study period. These improvements came along with similar relative increases in legal charges. Notably, these improvements were most pronounced in the manufacturing and financial sectors.

These results collectively suggest that Kanoon's provision of accessible legal information has had broad-ranging effects in India, including reducing bottlenecks, elevating efficiency, and enhancing legal research quality in courts. This case study has relevance far beyond India - it suggests that free legal search can reduce market-level constraints to economic development.

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.

Ahluwalia, Montek S (2019). "India's economic reforms: Achievements and next steps". In: *Asian Economic Policy Review* 14.1, pp. 46–62.

Arnold, Jens Matthias et al. (2016). "Services reform and manufacturing performance: Evidence from India". In: *The Economic Journal* 126.590, pp. 1–39.

Bau, Natalie and Adrien Matray (2023). "Misallocation and capital market integration: Evidence from India". In: *Econometrica* 91.1, pp. 67–106.

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

Bloom, Nicholas, Kalina Manova, et al. (2021). "Trade and management". In: *Review of Economics and Statistics* 103.3, pp. 443–460.

Boehm, Johannes and Ezra Oberfield (2020). "Misallocation in the Market for Inputs: Enforcement and the Organization of Production". In: *The Quarterly Journal of Economics* 135.4, pp. 2007–2058.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2021). "Revisiting event study designs: Robust and efficient estimation". In: *arXiv preprint arXiv:2108.12419*.

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

— (2021). "Difference-in-differences with multiple time periods". In: *Journal of econometrics* 225.2, pp. 200–230.

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.
Galanter, Marc and Nick Robinson (2013). “India’s Grand Advocates: a legal elite flourishing in the era of globalization”. In: *International Journal of the Legal Profession* 20.3, pp. 241–265.
Gennaioli, Nicola and Andrei Shleifer (2007). “The evolution of common law”. In: *Journal of Political Economy* 115.1, pp. 43–68.
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*.
Guillouet, Louise et al. (2021). *Language barriers in multinationals and knowledge transfers*. Tech. rep. National Bureau of Economic Research.
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.
McAdams, Richard H and Eric B Rasmusen (2007). “Norms and the Law”. In: *Handbook of law and economics* 2, pp. 1573–1618.
OECD (2017). “OECD Policy Roundtable on Equal Access to Justice”. In: *Roundtable Notes, OECD Paris 22-23 May*.
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.
Roth, Jonathan et al. (2023). “What’s trending in difference-in-differences? A synthesis of the recent econometrics literature”. In: *Journal of Econometrics*.

Sankaraguruswamy, Srinivasan and Umakanth Varottil (2023). "Which companies pay more (or less) in legal fees? An empirical study of India". In: *International Review of Law and Economics* 75, p. 106138.

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.

Supreme Court Online, Blog Post (2021). *Whether the arbitral process could be interfered under Article 226/227 of the Constitution, and under what circumstance? Supreme Court explains*. URL: https://www.scconline.com/blog/post/2021/01/14/whether-the-arbitral-process-could-be-interfered-under-article-226-227-of-the-constitution-and-under-what-circumstance-supreme-court-explains/.

United Nations (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. General Assembly. A/RES/70/1. URL: https://undocs.org/en/A/RES/70/1.

— (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.

VIDHI, Centre for Legal Policy (2020). *A Survey of Advocates Practicing Before the High Courts*. URL: https://vidhilegalpolicy.in/research/a-survey-of-advocates-practicing-before-the-high-courts/.

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.

Appendix

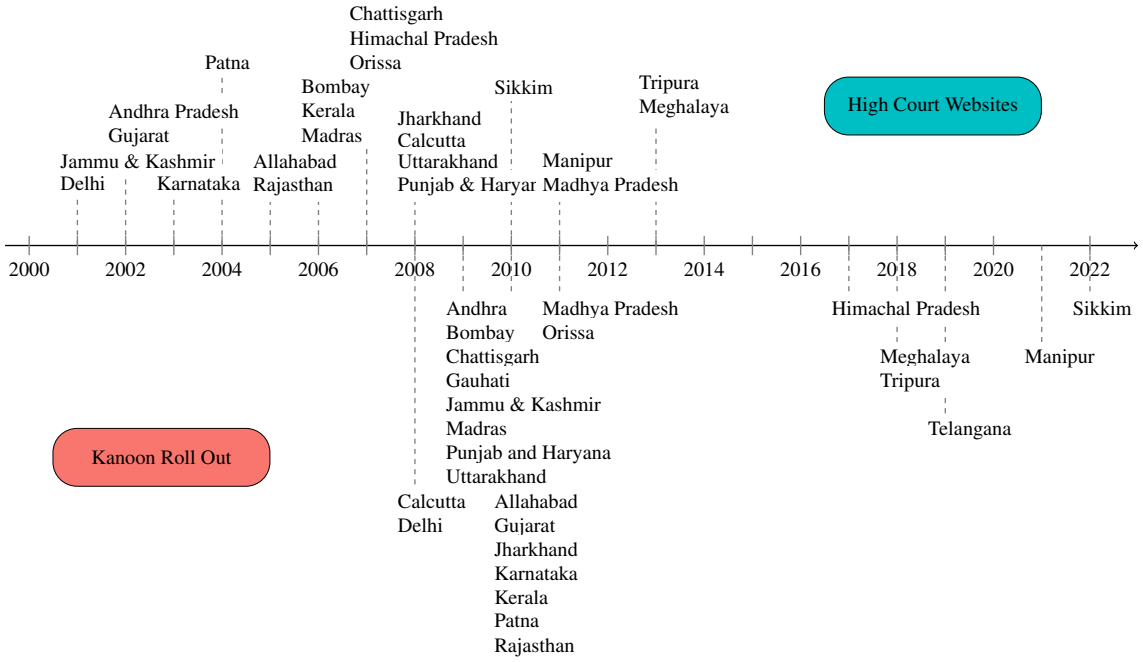


Figure A1: Rollout of High Court Websites versus Indian Kanoon

Note: This figure illustrates that the rollout of Indian Kanoon across high courts (bottom panel) did not follow closely or in a clear pattern the availability of data via high courts' websites (top panel). Names above the vertical timeline represent the year in which high court websites were first publicly accessible. The names below represent the year the Indian Kanoon database included the high courts for the first time. For exact dates of these events and additional information, see Appendix Table A1.

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

						In Analysis?	
High Court	**Established**	**State**	**State Created**	**Population**	**Kanoon**	**HC**	**DC**
Allahabad HC	1866-03-17	Uttar Pradesh	1950-01-26	19,9812,341	2010-07-08	y	y
Andhra Pradesh HC[a,b]	2019-01-01	Andhra Pradesh	1956-11-01	4,9506,799	2009-01-18	y	y
		Telangana (until Jan 2019)					n
Bombay HC	1862-08-14	Goa	1987-05-30	145,8545	2009-01-18	y	y
		Maharashtra	1960-05-01	112,374,333			y
		Dadra & N. H. & D. & D. (UT)	2020-01-26	586,956			n
Calcutta HC	1862-07-02	West Bengal	1950-01-26	91,276,115	2008-10-26	n	y
		Andaman & Nicobar Islands (UT)	1956-11-01	380,581			n
Chhattisgarh HC	2000-11-01	Chhattisgarh	2000-11-01	25,545,198	2009-01-18	y	y
Delhi HC	1966-10-31	Delhi (UT)	1956-11-01	16,787,941	2008-10-26	n	y
Gauhati HC[c]	1948-03-01	Arunachal Pradesh	1987-02-20	1,383,727	2009-01-18	y	n
		Assam	1950-01-26	31,205,576			y
		Mizoram	1987-02-20	1,097,206			y
		Nagaland	1963-12-01	1,978,502			n
		Meghalaya (until Mar 2013)					n
		Manipur (until Mar 2013)					n
		Tripura (until Mar 2013)					n
Gujarat HC	1960-05-01	Gujarat	1960-05-01	60,439,692	2010-06-13	y	y
Himachal Pradesh HC	1971-01-25	Himachal Pradesh	1971-01-25	6,864,602	2017-04-15	y	y
Jammu & K. and L. HC	1928-03-26	Jammu and Kashmir (UT)	2019-10-31	12,258,433	2009-01-18	y	y
		Ladakh (UT)	2019-10-31	290,492			n
Jharkhand HC	2000-11-15	Jharkhand	2000-11-15	32,988,134	2010-06-13	y	y

Continuation of Table A1

High Court	Established	State	State Created	Population	Kanoon Rollout	HC	DC
Karnataka HC	1905-02-26	Karnataka	1956-11-01	61,095,297	2010-06-13	y	y
Kerala HC	1956-11-01	Kerala	1956-11-01	33,406,061	2010-06-26	y	y
		Lakshadweep (UT)	1956-11-01	64,473			n
Madhya Pradesh HC	1936-01-02	Madhya Pradesh	1950-01-26	72,626,809	2011-01-29	n	y
Madras HC	1862-08-15	Tamil Nadu	1956-11-01	72,147,030	2009-01-18	y	y
		Puducherry (UT)	1962-08-16	1,247,953			n
Manipur HC[c]	2013-03-25	Manipur	1972-01-21	2,855,794	2020-02-27	n	n
Meghalaya HC[c]	2013-03-23	Meghalaya	1972-01-21	2,966,889	2018-10-21	n	n
Orissa HC	1948-04-03	Odisha	1950-01-26	41,974,218	2011-01-26	y	y
Patna HC	1916-09-02	Bihar	1950-01-26	10,409,9452	2010-06-13	y	y
Punjab and Haryana HC	1947-08-15	Chandigarh (UT)	1966-11-01	1,055,450	2009-03-20	n	y
		Haryana	1966-11-01	25,351,462			y
		Punjab	1966-11-01	27,743,338			y
Rajasthan HC	1949-06-21	Rajasthan	1950-01-26	68,548,437	2010-06-13	y	y
Sikkim HC	1975-05-16	Sikkim	1975-05-16	61,0577	2021-03-01	y	y
Telangana HC[a,b]	2019-01-01	Telangana	2014-06-02	35,193,978	2019-08-30	n	n
Tripura HC[c]	2013-03-23	Tripura	1972-01-21	367,3917	2018-10-30	n	n
Uttarakhand HC	2000-11-09	Uttarakhand	2000-11-09	10,086,292	2009-03-20	y	y

[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 Summary Statistics

Table A2: Summary Statistics for Courts Data

	N	Mean	SD	Min	Max
eCourt District Courts Data (State-Year-Month Level)					
Filings (Thsd.)	3,300	13.25	19.02	0.00	132.60
Decided cases (Thsd.)	3,300	9.18	19.90	0.00	353.02
Backlog (Thsd.)	3,300	386.71	619.53	0.01	3,661.21
Disposition time (Years)	2,924	24.41	102.45	0.00	1,650.17
eCourt High Courts Data (State-Year-Month Level)					
Filings (Thsd.)	2,120	3.14	3.37	0.00	20.83
Decided cases (Thsd.)	2,143	2.46	2.76	0.00	35.32
Backlog (Thsd.)	2,187	74.22	72.60	0.00	329.39
Disposition time (Years)	2,142	0.80	4.00	0.01	99.41
Filings Criminal	2,120	0.60	0.97	0.00	5.81
Filings Civil	2,120	0.74	1.65	0.00	12.73
Filings Writ	2,120	0.33	0.86	0.00	10.89
Filings Appeal	2,120	0.88	1.89	0.00	16.32
Filings Other	2,120	0.59	1.24	0.00	10.30
Appeals of District Court Cases (State-Year-Month Level)					
Appeals (Thsd.)	1,980	0.10	0.17	0	1.1
% Appealed	1,980	0.01	0.01	0	.077
High Court Case Status of District Court Appeals (State-Year-Month Level)					
Admitted	1,645	59.03	106.66	0	667
Disposed	1,980	48.95	99.63	0	667
Withdrawn	1,980	2.11	5.56	0	56
Overruled	1,980	13.63	28.38	0	167
% Admitted	1,522	0.83	0.18	0	1
% Disposed (of Admitted)	1,509	1.00	0.01	.94	1
% Withdrawn (of Disposed)	1,509	0.08	0.13	0	1
% Overruled (of Disposed)	1,509	0.25	0.21	0	1

Note: All data sources cover the years 2005 - 2015.

Table A3: Summary Statistics for Firm Data

	N	Mean	SD	Min	Max
Firm Financial Data (Yearly): All Firms					
Income (billion INR)	82,742	6.05	64.93	-0.000	5,439.60
Assets (billion INR)	82,742	11.87	116.39	-0.002	6,487.84
Expenditures (billion INR)	82,742	5.72	62.64	-1.825	5,380.96
External Legal Charges (billion INR)	78,826	0.01	0.07	0.000	6.56
External Legal Charges (if > 0) (billion INR)	43,542	0.02	0.10	0.000	6.56
Any External Legal Charges	78,826	0.55	0.50	0.000	1.00
Firm Financial Data (Yearly): Firms with at least one case in 2005-2015					
Income (billion INR)	19,481	10.89	110.06	0.000	5,439.60
Assets (billion INR)	19,481	20.86	179.20	-0.002	6,487.84
Expenditures (billion INR)	19,481	10.33	107.60	-0.107	5,380.96
External Legal Charges (billion INR)	18,689	0.01	0.09	0.000	6.56
External Legal Charges (if > 0) (billion INR)	11,659	0.02	0.11	0.000	6.56
Any External Legal Charges	18,689	0.62	0.48	0.000	1.00
eCourt High Courts Cases linked to Firms (State-Month Level)					
Filings (thousand)	2,772	0.14	0.31	0.000	4.22
Decisions (thousand)	2,772	0.10	0.31	0.000	7.48
Backlog (thousand)	2,772	2.40	5.51	0.002	33.85
Disposition Time (years)	1,697	0.42	0.72	0.002	7.96

Note: For financial data, every observation represents one financial year of a firm (April 1st to March 31st of the following year). Financial data is obtained from Prowess and subset to the years 2005–2015 and to firms which have non-zero income data for all 11 years.

A2 Additional Appeal Information

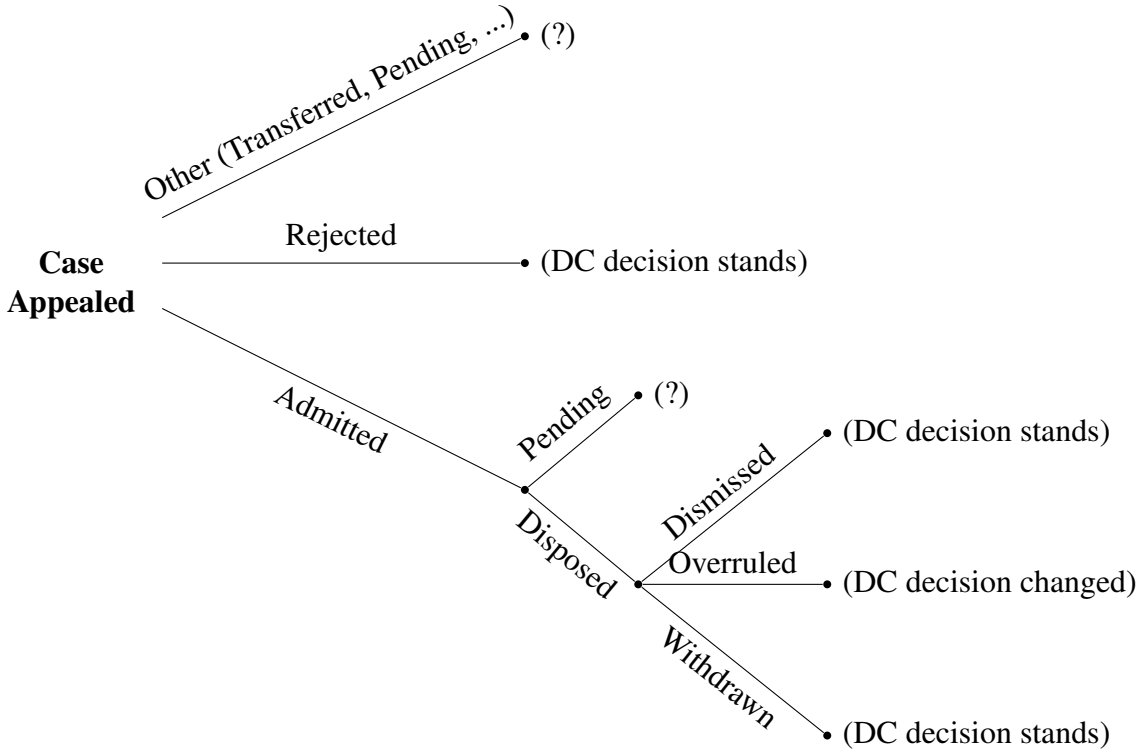


Figure A2: Decision Tree of District Court Appeals in High Courts

Note: *Admitted:* the appeal case's paperwork is in order, and the case is deemed fit for processing at the high court and listed for a hearing/decision. *Rejected:* the appeal cases falls outside the court's jurisdiction or its paperwork has irreparable defects. *Disposed:* the appeal cases was heard by the high court and decided by the bench. *Withdrawn:* the appeal cases was withdrawn by the petitioner. *Overruled:* the appeal cases was (at least partly) approved by the high court. *Dismissed* the appeal cases was dismissed by the high court.

A3 Additional Results on High Courts

Table A4: Kanoon's Effect on Filings in High Courts per Case Type

	Number of Filings of Case Type (in thsd)				
	(1) Criminal	(2) Civil	(3) Writ	(4) Appeal	(5) Other
Post-Kanoon introduction	0.065	-0.748**	-0.154*	-0.359	0.030
	(0.193)	(0.306)	(0.087)	(0.242)	(0.051)
Court FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Mean Dep. Var.	0.60	0.74	0.33	0.88	0.59
Observations	2,120	2,120	2,120	2,120	2,120

Note: This table displays estimation results of the impact of Indian Kanoon on the number of filings per case type in high courts (in thousand). Observations are at the court-year-month level. The data is based on all cases active in the period 2005-2015 and available eCourts High Court in November 2020 (Panel B). Every estimation is regressing the number of filings on a dichotomous variable, equal to one if the year-month is post the Indian Kanoon rollout for the respective high court, and several fixed effects. Each column is counting the number of filings of the following case types: Criminal, Civil, Writ, Appeal, and Other if the case can not be categorized in any of the other categories. Note that criminal, civil, writ and appeal are not mutually exclusive case types. E.g., a "criminal appeal" is counted as both criminal and appeal. 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.

A4 Additional Results on Firms

Table A5: Effects on External Legal Charges

	(1) External Legal Charges (if > 0)	(2) Any External Legal Charges
Panel A: Partial Equilibrium		
Post-Kanoon introduction	0.007^{***}	0.087^{***}
	(0.002)	(0.015)
Mean Dep. Var.	0.02	0.62
Observations	10,809	18,689
Panel B: General Equilibrium		
Post-Kanoon introduction	0.005^{***}	0.015^{*}
	(0.002)	(0.009)
Mean Dep. Var.	0.02	0.55
Observations	39,486	78,639
Year FE	✓	✓
Firm FE	✓	✓

Note: This table display estimation results of the impact of Indian Kanoon on several firm financials. Observations are at the firm-financial year level. The data includes firms with non-zero income data available for all years 2005-2015 in the Prowess database. Every estimation is regressing a financial variable on a dichotomous variable, equal to one if the financial year started post the Indian Kanoon rollout of the high court having jurisdiction over the state the firm is registered in, and year and firm fixed effects. A financial year is defined as April 1st to March 31st of the following year. *External Legal Charges (if > 0)* keeps all observations with external legal charges greater than zero and is measured in billion Indian rupees. *Any External Legal Charges* is a dichotomous variable equal to one if external legal charges of a firm in a financial year are greater than zero (and zero otherwise). Panel A restricts the sample to firms which had at least one active case anytime in the period 2005-2015. Panel B includes all firms in the sample. The regressions use the *did_imputation* command in STATA to estimate treatment effects robust to staggered rollout. 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.

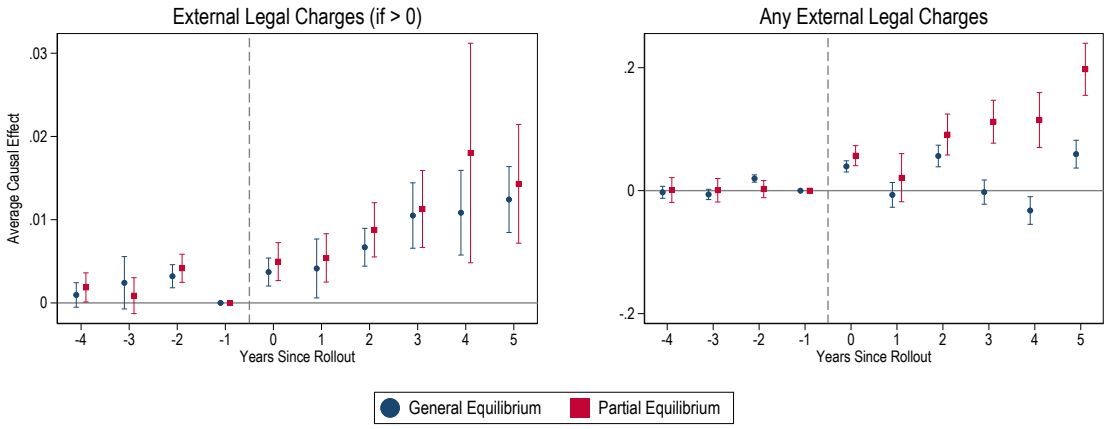


Figure A3: Dynamic Effects of Kanoon rollout on External Legal Charges

Note: This figure presents dynamic estimation results of the impact of Indian Kanoon on various firm outcomes for all firms (General Equilibrium) and firms having at least one district court case during the considered time frame (Partial Equilibrium). Observations are at the firm-financial year level and outcomes are measured in billion Indian rupees. The data includes firms with non-zero income data available for all years 2005-2015 in the Prowess database. The leads and lags are defined as the yearly difference between the start of the financial year and the date of rollout on Indian Kanoon of the firm's registration state's high court. *Lead 1*, 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 *eventstudyinteract* command in STATA to estimate consistent and heterogeneous treatment robust average dynamic treatment effects. Any court that has not been rolled out by 2017 is used as control cohort. The estimation controls for firm and year fixed effects. Standard errors are clustered at the state level. Confidence intervals are at the 95% level.

Online Appendix

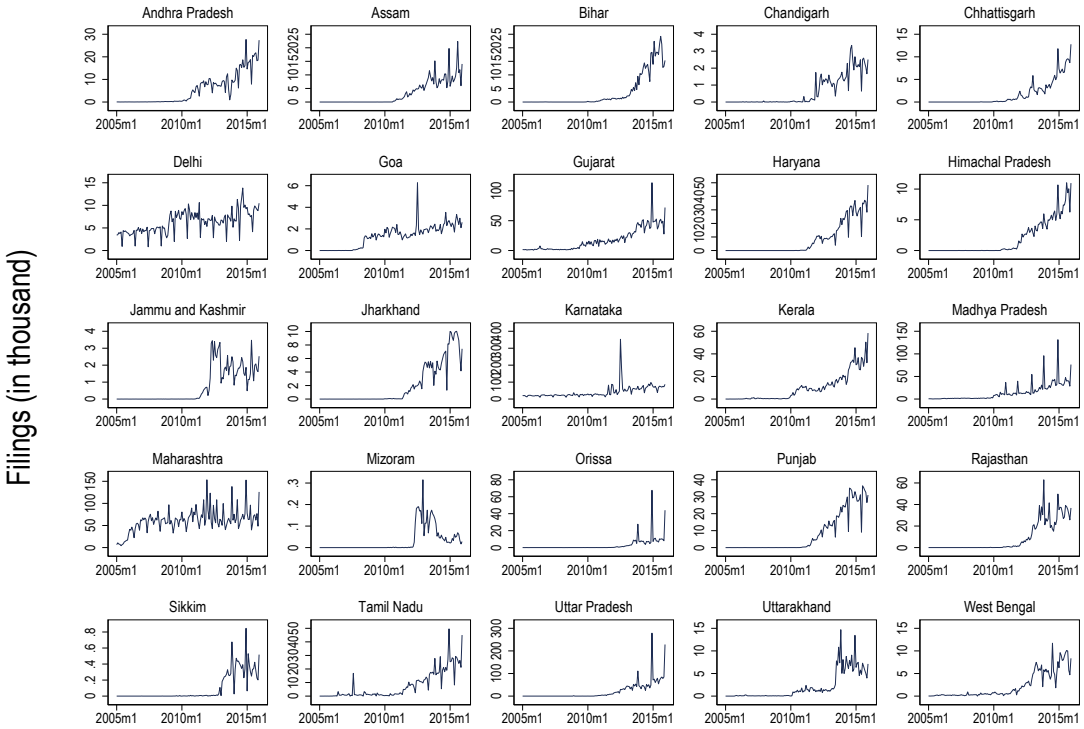


Figure OA1: Number of Filings in Indian District Courts per State

This figure presents the time evolution of the number of cases filed in Indian district courts per month and state. Filings are counted in thousand cases.

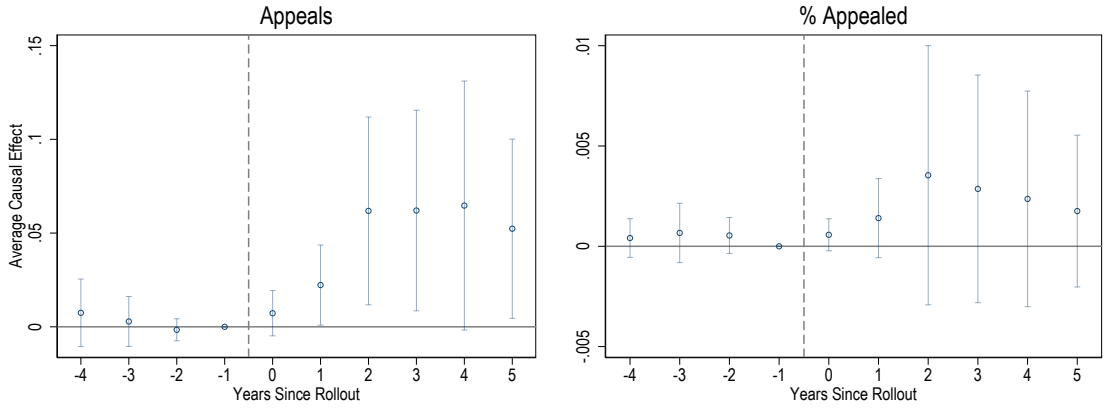


Figure OA2: Impact of Kanoon on Appeals of District Court Cases in High Courts

Note: This table presents dynamic estimation results examining the impact of Indian Kanoon on appeals of district courts cases in high courts. Observations are at the court-year-month level. The data is based on all cases active in the period 2005-2015 available on eCourts District Court in May 2019 and appealed later on in the high court. The leads and lags are defined as the yearly difference between the filing year and month of a case in the district court and the year and month of the rollout of the state's high court on Indian Kanoon. *Lead 1*, 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. *Appeals* counts the number of district court cases filed in a state-year-month and appealed later on in the high court (in thousand). *% Appealed* is the share of district court cases filed in a state-year-month and which are appealed later on in the high court. The estimation uses the *eventstudyinteract* command in STATA to estimate consistent and heterogeneous treatment robust average dynamic treatment effects. Any court that has not been rolled out by 2017 is used as control cohort. The estimation controls for state, year, and month fixed effects. Standard errors are clustered at the state level. Confidence intervals are at the 95% level.

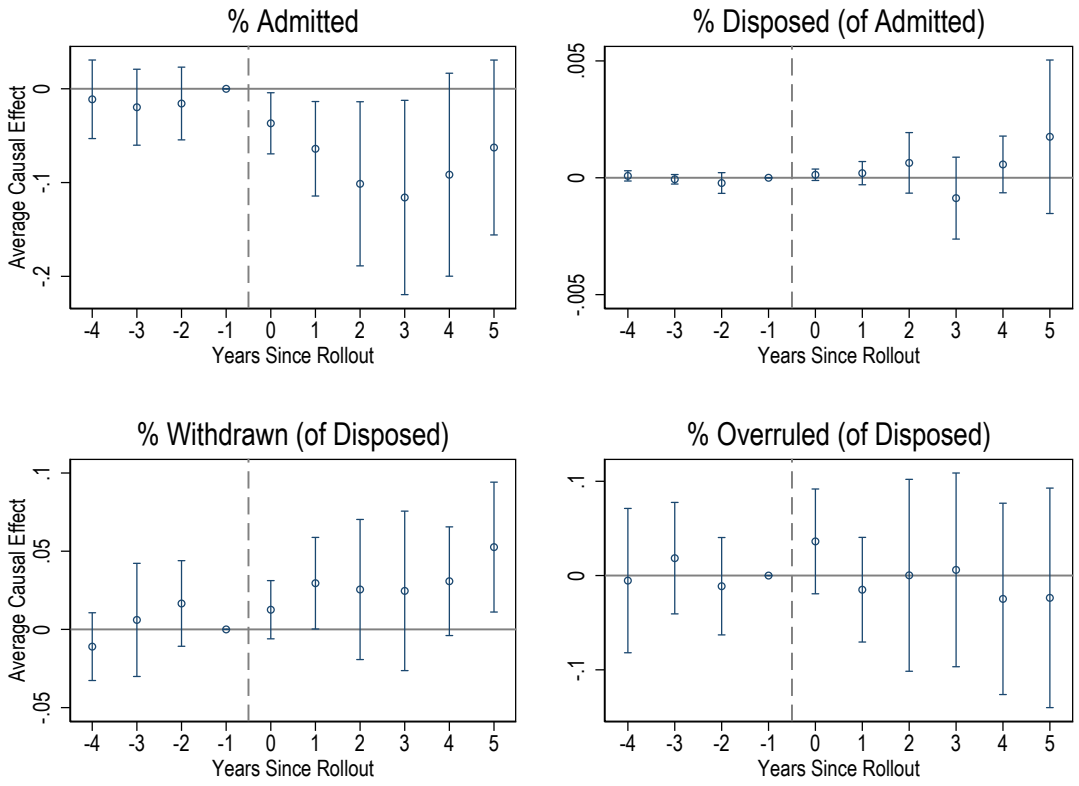


Figure OA3: Impact of Kanoon on Appeal Outcomes of District Court Cases in High Courts (in %)

Note: This figure presents dynamic estimation results of the impact of Indian Kanoon on outcomes of district courts cases appealed in the high courts. Observations are at the court-year-month level. The data is based on all cases active in the period 2005-2015 available on eCourts District Court in May 2019 and appealed later on in the high court. The leads and lags are defined as the yearly difference between the year and month of an observation and the year and month of the rollout of the state's high court on Indian Kanoon. *Lead 1*, 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. *% Admitted* measures the share of all district court appeal cases for which all paperwork is in order, and the case is deemed fit for processing at the high court and listed for a hearing/decision. *% Disposed* measures the share of *admitted* appeal cases which is heard by the high court and decided by the bench. *% Withdrawn* measures the share of *disposed* appeal cases withdrawn by the petitioner. *% Overruled* measures the share of *disposed* appeal cases (at least partly) approved by the high court. The estimation uses the *eventstudyinteract* command in STATA to estimate consistent and heterogeneous treatment robust average dynamic treatment effects. Any court that has not been rolled out by 2017 is used as control cohort. The estimation controls for state, year, and month fixed effects. Standard errors are clustered at the state level. Confidence intervals are at the 95% level.

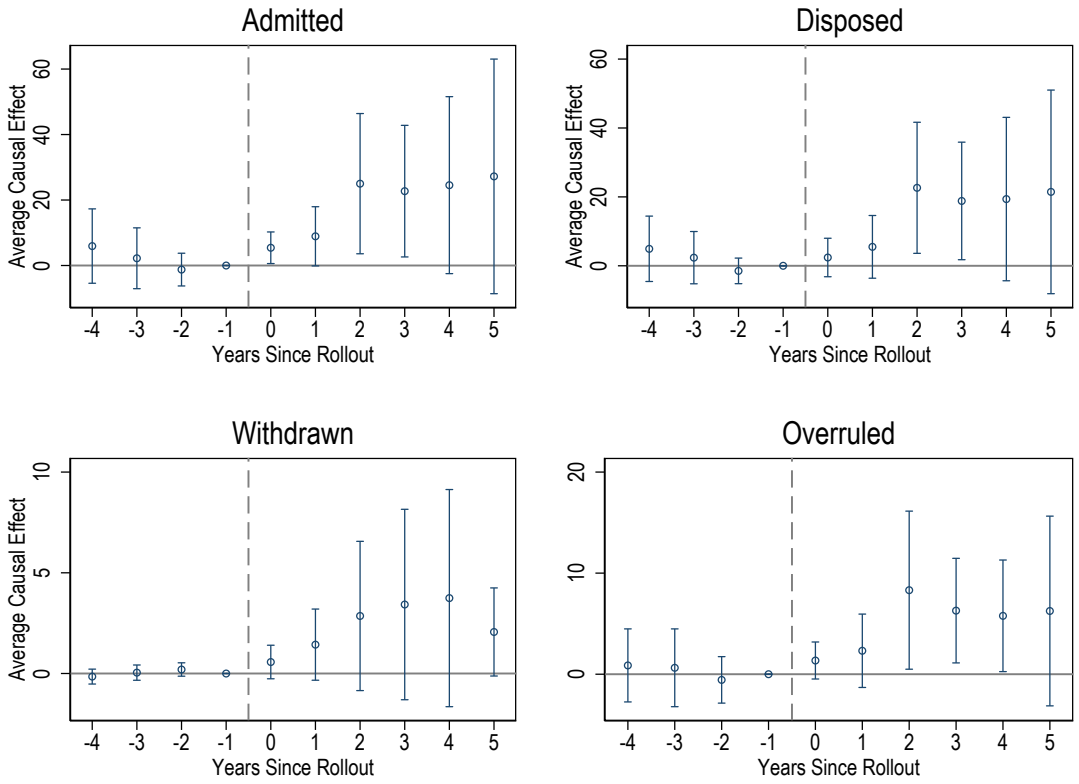


Figure OA4: Impact of Kanoon on Appeal Outcomes of District Court Cases in High Courts (in Levels)

Note: This figure presents dynamic estimation results of the impact of Indian Kanoon on outcomes of district courts cases appealed in the high courts. Observations are at the court-year-month level. The data is based on all cases active in the period 2005-2015 available on eCourts District Court in May 2019 and appealed later on in the high court. The leads and lags are defined as the yearly difference between the year and month of an observation and the year and month of the rollout of the state's high court on Indian Kanoon. *Lead 1*, 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. *Admitted* counts the number of district court appeal cases for which all paperwork is in order, and which are deemed fit for processing at the high court and listed for a hearing/decision. *Disposed* counts the number of appeal cases which are heard by the high court and decided by the bench. *Withdrawn* counts the number of appeal cases withdrawn by the petitioner. *Overruled* counts the number of appeal cases which are (at least partly) approved by the high court. The estimation uses the *eventstudyinteract* command in STATA to estimate consistent and heterogeneous treatment robust average dynamic treatment effects. Any court that has not been rolled out by 2017 is used as control cohort. The estimation controls for state, year, and month fixed effects. Standard errors are clustered at the state level. Confidence intervals are at the 95% level.

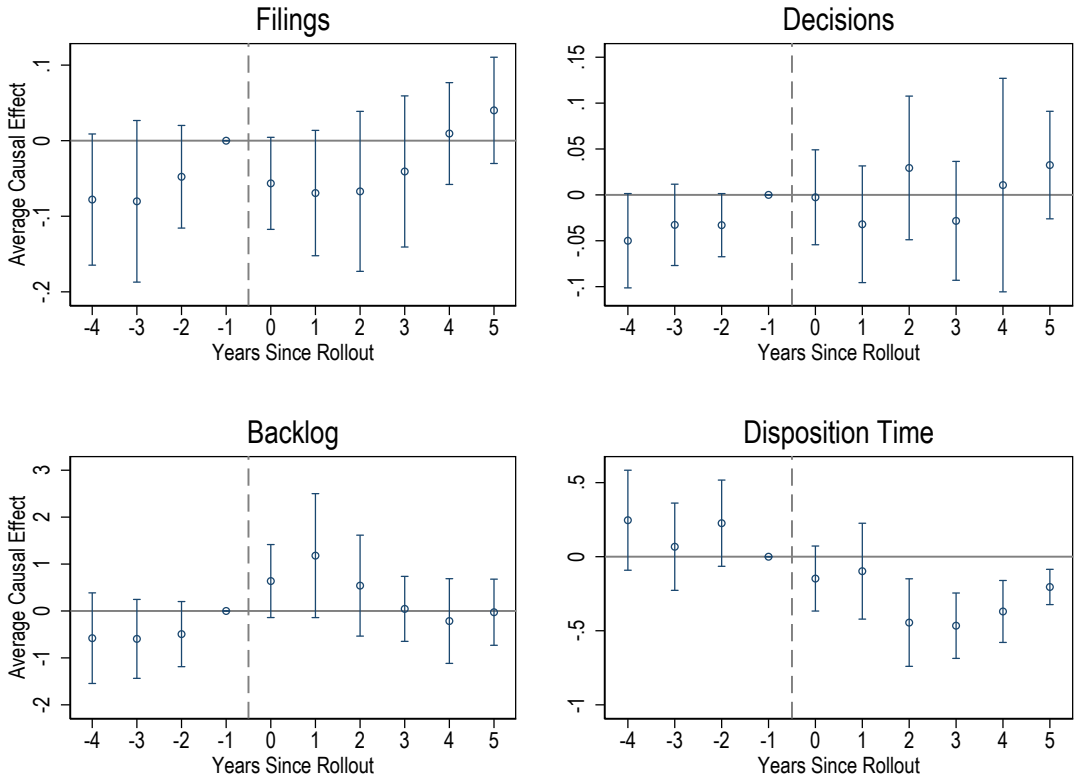


Figure OA5: Impact of Kanoon on District Court Efficiency Measures for Cases linked to Firms

Note: This figure presents dynamic estimation results of the impact of Indian Kanoon on district court efficicency measures of the subset of cases linked to firms in the Prowess datatabase. Observations are at the state-year-month level. The data is based on all cases active in the period 2005-2015 and available on eCourts District Court in May 2019 and which are linked to (at least) one firm in the Prowess database. Leads and lags are defined as the yearly difference between the year and month of an observation and the year and month of the rollout of the state's high court on Indian Kanoon. *Lead 1*, 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. *Filings* counts the number of filings per state-year-month (in thousand). *Decisions* counts the number of decided cases per state-year-month (in thousand). *Backlog* counts the number of undecided cases in a state at the end of a year-month, and which were filed at least one year prior to this year-month (in thousand). *Disposition Time* is the ratio between the number of unresolved cases at the end of a year-month and the number of decisions in that year-month in a state, and is a measure of how many years it takes on average to resolve a case. The estimation uses the *eventstudyinteract* command in STATA to estimate consistent and heterogeneous treatment robust average dynamic treatment effects. Any court that has not been rolled out by 2017 is used as control cohort. The estimation controls for state, year, and month fixed effects. Standard errors are clustered at the state level. Confidence intervals are at the 95% level.

OA1 Kanoon's Competition: An Overview of Other Electronic Legal Databases

As noted in the paper, Kanoon was not the only electronic source of legal information 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.

OA2 Appeals Data Creation

For the appeal regressions in Tables 2 and 3 and Online Appendix Figures OA2, OA3, OA4, OA8, OA9, and OA10 we created a district-year-month level sample with variables drawn from the database of cases at the district courts. These include: *Appeals*, *% Appealed*, and high court variables *Admitted*,*Disposed*, *Withdrawn*, and *Overruled*. In the following, we describe the creation of this sample.

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*, *Disposed*, *Withdrawn*, and *Overruled* 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.

OA3 Additional Analysis: Did Kanoon Change Citation Practices?

The introduction of Kanoon, a user-friendly, comprehensive legal database in India, has revolutionized the study of past legal decisions, a cornerstone of the common law system. Here, we look into variables capturing the relationship between high court rulings and previous and subsequent cases. To accomplish this, we scrutinize the following four principal citation metrics. *Backward citations*, representing the quantity of previously ruled cases referenced in a judgment. *Forward citations*, representing the frequency a judgment is cited in subsequent cases. *Backward self-citations*, representing the quantity of previously ruled cases from the same high court referenced in a judgment. *Forward self-citations*, representing the frequency a judgment is cited in subsequent cases within the same court.

Analyzing the impact of Kanoon on the extensive margin of citations, we employ dichotomous variables, assigning a value of one if the respective citation measure exceeds zero. Our dataset covers the period from 2005 to 2015, with 2,162,698 documents accessible through the Kanoon website. Interestingly, we find that only 6% of these documents cite at least one other case, and the percentage is even lower for the other citation measures (Table OA1). When examining the impact of Kanoon's implementation on the share of cases with at least one citation, we perceive a roughly 30% augmentation compared to the overall mean, although this result is not statistically significant. (Panel A of Table OA2). This result is anticipated, given that many documents on Kanoon are intermediate orders rather than final verdicts, often lacking references to other cases.

Moving to the intensive margin, where we focus solely on cases with at least one citation, we find that these cases, on average, cite 3.3 past cases (1.6 when limited to cases from the same court) and are cited by other cases an average of 2.7 times (2.59 when restricted to cases from the same court). Notably, Kanoon's launch has a significant effect, culminating in a 13% increase in citations between cases from the same court (Panel B of Table OA2). This corroborates intuition, as rulings from the same court substantially influence future cases. The availability of a free, readily accessible database, especially in the context of a common law system, is greatly beneficial in fortifying courtroom arguments.

Panels A and B of Table OA2 sought to understand Kanoon's impact on citation practices within High Courts, relying on cases uploaded to Kanoon and assuming pre and post Kanoon case similarities. To assuage potential concerns regarding differential case upload influencing citation patterns, we narrow our focus to a critical subset of cases.

Hence, we zero in on cases featured in the All India Reporter (AIR), a venerated legal publication in India. Court verdicts in AIR are regularly cited in official proceedings. By using a 'comprehensive search' query to access appeal cases from district courts on the publisher's site, we compile an exhaustive dataset spanning from 1980 to 2021. These cases are then matched with entries in the Indian Kanoon database using decision dates and litigant names. Our analysis is limited to cases that a committee of juridical experts deems as significant precedents. The AIR database comprises cases approved by this committee for citation by judges in official court proceedings across India, making it a pivotal sample influencing future judicial decisions. By examining the impact of Kanoon's free

Table OA1: Summary Statistics for Citations

	N	Mean	SD	Min	Max
Any Backward Citation	3,234,083	0.08	0.27	0	1
Any Forward Citation	3,234,083	0.03	0.17	0	1
Any Backward Self-Citation	3,234,083	0.04	0.20	0	1
Any Forward Self-Citation	3,234,083	0.02	0.15	0	1
Backward Citations of all cases by all cases	265,223	3.00	4.45	1	368
Forward Citations of all cases by all cases	101,515	3.10	33.58	1	9,131
Backward Self-Citations of all cases by all cases	130,214	1.53	1.45	1	124
Forward Self-Citations of all cases by all cases	77,272	2.99	37.19	1	9,121
Backward Citations of AIR cases by all cases	265,223	0.08	0.30	0	9
Forward Citations of AIR cases by all cases	9,369	3.43	12.21	1	820
Backward Self-Citations of AIR cases by all cases	130,214	0.09	0.32	0	9
Forward Self-Citations of AIR cases by all cases	7,278	3.33	13.33	1	811

Note: Observations are at the case level. The data is based on all judgments and orders from the period 2005-2015 available in March 2021 on Indian Kanoon. *Backward citations* measures the number of (past) (AIR) cases a judgment is citing. *Forward citations* measures the number of times an (AIR) judgment is cited in the future. *Self backward citations* is the number of times a judgment is citing (AIR) cases from the same court in which it was heard. *Self forward citations* is the number of times an (AIR) judgment is cited in the future by cases in the same court.

legal search on this specific subset, we maintain the composition of cases before and after Kanoon's rollout, enabling us to investigate the causal effect of reducing information frictions on the rule of law.

Notably, this selected subset of cases demonstrates a heightened citation frequency, both across all cases and within the same court (Table OA1). The analysis of these cases indicates a notable 35% surge in citations and a 30% rise in self-citations of AIR cases facilitated by Kanoon (Panel C of Table OA2). This suggests that Kanoon has efficaciously enabled stakeholders to identify pertinent citations for their cases, though the additional step of securing the official citation from another source may still be required.

Table OA2: Impact of Kanoon on Citations of Published Rulings

	Citations		Self-Citations	
	(1) Backward	(2) Forward	(3) Backward	(4) Forward
Panel A: Citations to All Cases - Extensive Margin				
Publication Post-Kanoon introduction	0.0094	0.0012	0.0049	0.0035
	(0.00980)	(0.00575)	(0.00430)	(0.00350)
Mean Dep. Var.	0.08	0.04	0.03	0.02
N	3,234,083	3,234,083	3,234,083	3,234,083
Panel B: Citations to All Cases - Intensive Margin				
Publication Post-Kanoon introduction	0.226	-0.445*	0.128**	-0.670*
	(0.223)	(0.252)	(0.048)	(0.381)
Mean Dep. Var.	3.00	3.10	1.53	2.99
N	265,223	101,515	130,214	77,272
Panel C: Citations to AIR Cases - Intensive Margin				
Publication Post-Kanoon introduction	0.005	0.548	0.009	0.495
	(0.014)	(0.494)	(0.017)	(0.715)
Mean Dep. Var.	0.08	3.43	0.09	3.33
N	265,223	9,368	130,214	7,278
Court FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓

Note: This table displays estimation results of the impact of Indian Kanoon on citation practice in high courts. Observations are at the case level. The data is based on all judgments and orders from the period 2005-2015 available in March 2021 on Indian Kanoon. Every estimation is regressing a citation count on a dichotomous variable, equal to one if the document's decision date is post the Indian Kanoon rollout of the high court, and several fixed effects. Panel A includes all available judgments and orders, the dependent variables are dummies equal to 1 if the respective citation count is greater than 0. Panel B restricts the observations to cases where the respective citation count is greater than 0. Panel C restricts the citations to cases published in the All India Reporter (AIR) and other allied journals from the high courts and the supreme court of India, spanning the years from 1980 to 2021, and keeps only cases where the respective citation count is greater than 0. *Backward citations* measure the number of (past) (AIR) cases a judgment is citing. *Forward citations* measure the number of times an (AIR) judgment is cited in the future. *Self backward citations* is the number of times a judgment is citing (AIR) cases from the same court in which it was heard. *Self forward citations* is the number of times an (AIR) judgment is cited in the future by cases in the same court. 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.

OA4 Robustness to Different Estimation Techniques

Table OA3: Kanoon's Impact on Court Efficiency - Robustness

	(1) Filings	(2) Decided	(3) Backlog	(4) disposition
Panel A: District Courts				
Borusyak, Jaravel, & Spiess (2023)	6.929***	8.256***	105.593*	-2.221
	(1.942)	(2.151)	(56.566)	(12.558)
TWFE	4.741**	6.711***	114.289*	-15.603
	(1.732)	(2.080)	(55.552)	(20.889)
Callaway & Sant'Anna (2021) - Never Treated	5.036**	7.434***	81.487*	-13.021
	(2.034)	(2.519)	(46.989)	(12.641)
Callaway & Sant'Anna (2021) - Not Yet Treated	4.577**	7.194***	69.515	-6.944
	(2.034)	(2.621)	(45.723)	(9.367)
Sun & Abrahams (2020)	2.547	4.759	49.192	3.459
	(2.114)	(1.703)	(54.319)	(15.265)
Mean Dep. Var.	13.25	9.18	386.71	24.41
Observations	3,300	3,300	3,300	2,924
Panel B: High Courts				
Borusyak, Jaravel, & Spiess (2023)	-1.165**	-1.028***	20.864**	-0.013
	(0.549)	(0.294)	(8.149)	(0.905)
TWFE	-0.610	-0.037	16.410**	0.154
	(0.647)	(0.520)	(7.480)	(0.259)
Callaway & Sant'Anna (2021) - Never Treated	-0.927*	-0.638***	16.093**	0.835
	(0.488)	(0.240)	(6.412)	(0.607)
Callaway & Sant'Anna (2021) - Not Yet Treated	-0.914	-0.580**	15.888**	0.680
	(0.570)	(0.291)	(6.284)	(0.554)
Sun & Abrahams (2020)	-0.656	-0.062	13.048	0.208
	(0.685)	(0.539)	(7.859)	(0.258)
Mean Dep. Var.	3.14	2.46	74.22	0.80
Observations	2,120	2,143	2,187	2,142
Year FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓

Note: This table displays estimation results from different estimators of the impact of Indian Kanoon on overall court efficiency in district courts (Panel A) and high courts (Panel B). Observations are at the court-year-month level. The data is based on all cases active in the period 2005-2015 and available on eCourts District Court in May 2019 (Panel A) and eCourts High Court in November 2020 (Panel B). Every estimation is regressing a measure of court efficiency on a dichotomous variable, equal to one if the year-month is post the Indian Kanoon rollout for the respective high court, and several fixed effects. *Filings* counts the number of filings per state-year-month (in thousand). *Decisions* counts the number of decided cases per state-year-month (in thousand). *Backlog* counts the number of undecided cases in a state at the end of a year-month, and which were filed at least one year prior to this year-month (in thousand). *Disposition Time* is the ratio between the number of unresolved cases at the end of a year-month and the number of decisions in that year-month in a state, and is a measure of how many years it takes on average to resolve a case. The estimations use, respectively, *did_imputation*, *reghdfe*, *csdid*, and *csdid* with option *notyet* commands in STATA to estimate treatment effects. Standard errors are clustered at the high court level. Stars *, ** and *** indicate that the

Table OA4: Effects on Firm Financials: Robustness

	(1) Income	(2) Assets	(3) Expenditures	(4) Legal Charges
Panel A: Partial Equilibrium				
TWFE	-0.193	1.089	-0.126	-0.001
	(0.555)	(0.861)	(0.513)	(0.001)
Borusyak, Jaravel, & Spiess	5.261***	13.026***	5.393***	0.005***
	(1.309)	(2.344)	(1.267)	(0.001)
Callaway & Sant'Anna - Never Treated	3.100***	11.217***	3.201***	0.001
	(1.187)	(2.140)	(1.058)	(0.007)
Callaway & Sant'Anna - Not Yet Treated	3.235***	10.441***	3.439***	0.001
	(1.147)	(2.072)	(1.009)	(0.006)
Mean Dep. Var.	10.89	20.86	10.33	0.01
Observations	19,481	19,481	19,481	18,689
Panel B: General Equilibrium				
TWFE	0.022	0.490	0.017	0.000
	(0.178)	(0.360)	(0.189)	(0.001)
Borusyak, Jaravel, & Spiess	2.349***	5.730***	2.452***	0.004**
	(0.790)	(1.756)	(0.772)	(0.002)
Callaway & Sant'Anna - Never Treated	1.669***	4.929***	1.695***	0.003
	(0.526)	(1.068)	(0.477)	(0.002)
Callaway & Sant'Anna - Not Yet Treated	1.626***	4.525***	1.684***	0.003
	(0.497)	(1.027)	(0.445)	(0.002)
Mean Dep. Var.	6.05	11.87	5.72	0.01
Observations	82,742	82,742	82,742	78,826
Year FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓

Note: This table displays estimation results of the impact of Indian Kanoon on several firm financials. Observations are at the firm-financial year level. The data includes firms with non-zero income data available for all years 2005-2015 in the Prowess database. Every estimation is regressing a financial variable on a dichotomous variable, equal to one if the financial year started post the Indian Kanoon rollout of the high court having jurisdiction over the state the firm is registered in, and year and firm fixed effects. A financial year is defined as April 1st to March 31st of the following year. Panel A restricts the sample to firms which had at least one active case anytime in the period 2005-2015. Panel B includes all firms in the sample. The estimations use, respectively, *reghdfe*, *did_imputation*, *csdid*, and *csdid* with option *notyet* commands in STATA to estimate treatment effects. 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.

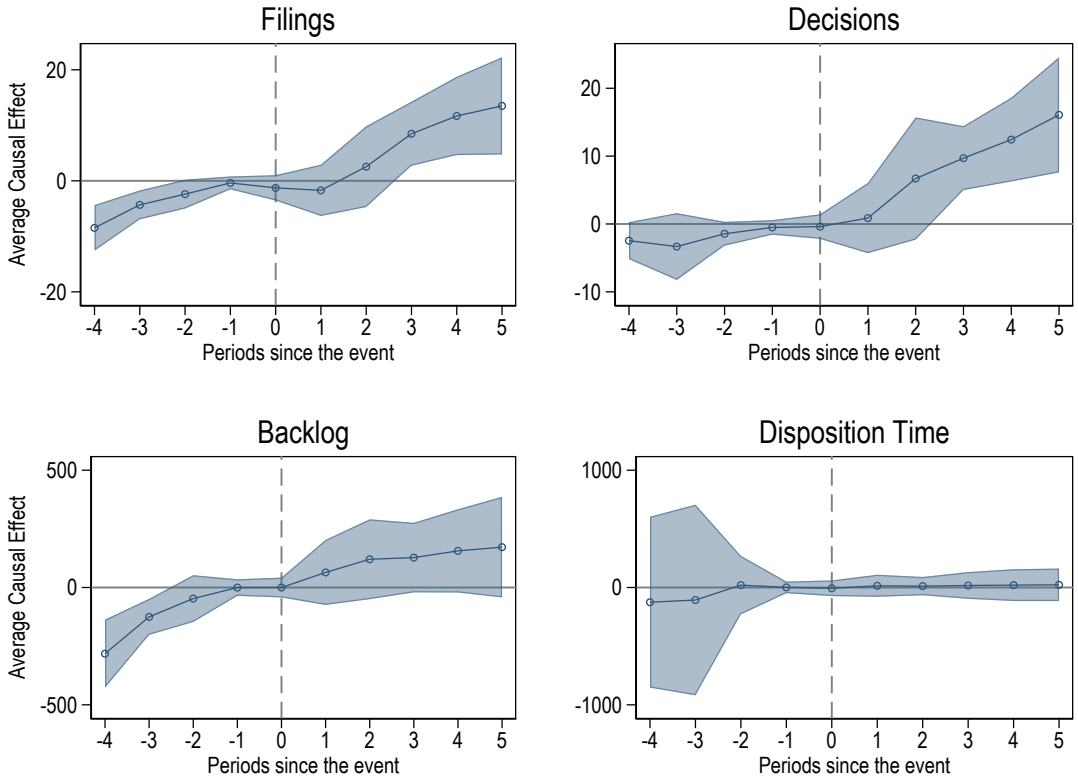


Figure OA6: Robustness Check for Aggregate Efficiency Measures of District Courts

Notes: This figure shows the results of the Placebo test for checking pre-trends as suggested by de Chaisemartin and D'Houltfoeuille, 2020. The Placebo tests are constructed assuming that, for tests whose treatment actually happens at t, treatment occurred 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, and month fixed effects. Standard errors are clustered at the state level.

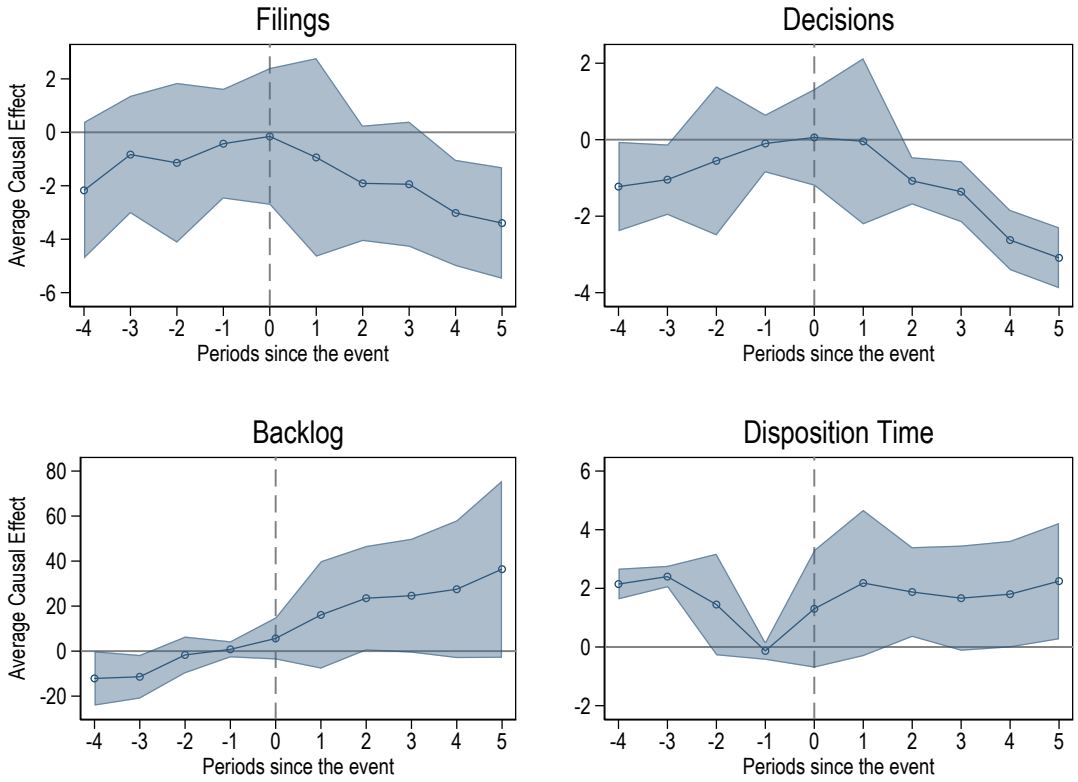


Figure OA7: Robustness Check for Aggregate Efficiency Measures of High Courts

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 occurred 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, and month fixed effects. Standard errors are clustered at the state level.

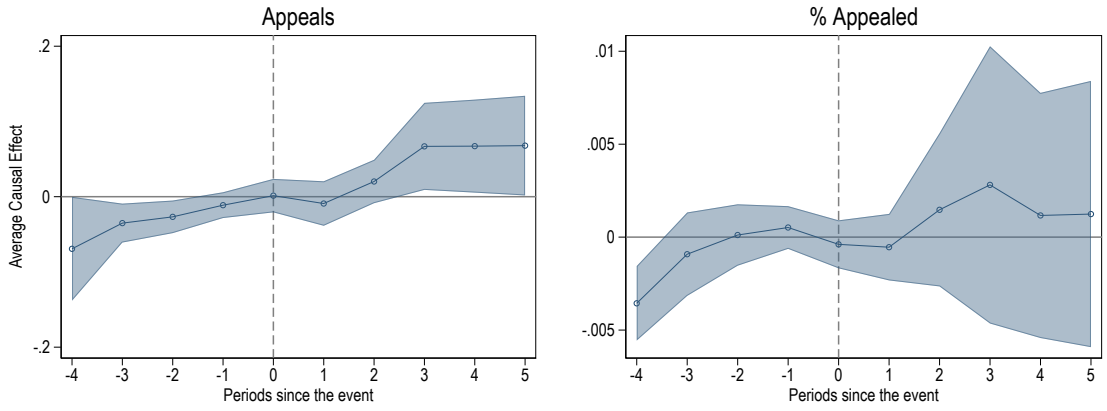


Figure OA8: Robustness Check for Appeals of district court cases.

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 occurred 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.

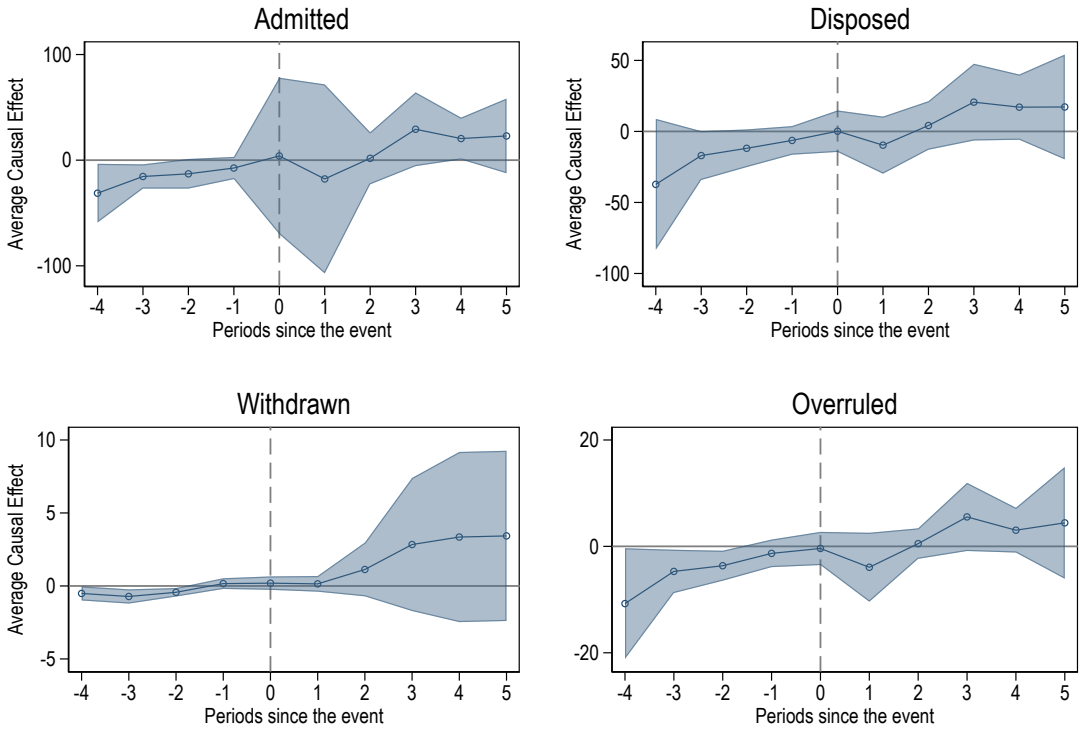


Figure OA9: Robustness Check for Kanoon's Impact on Appeal Outcomes

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 occurred 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.

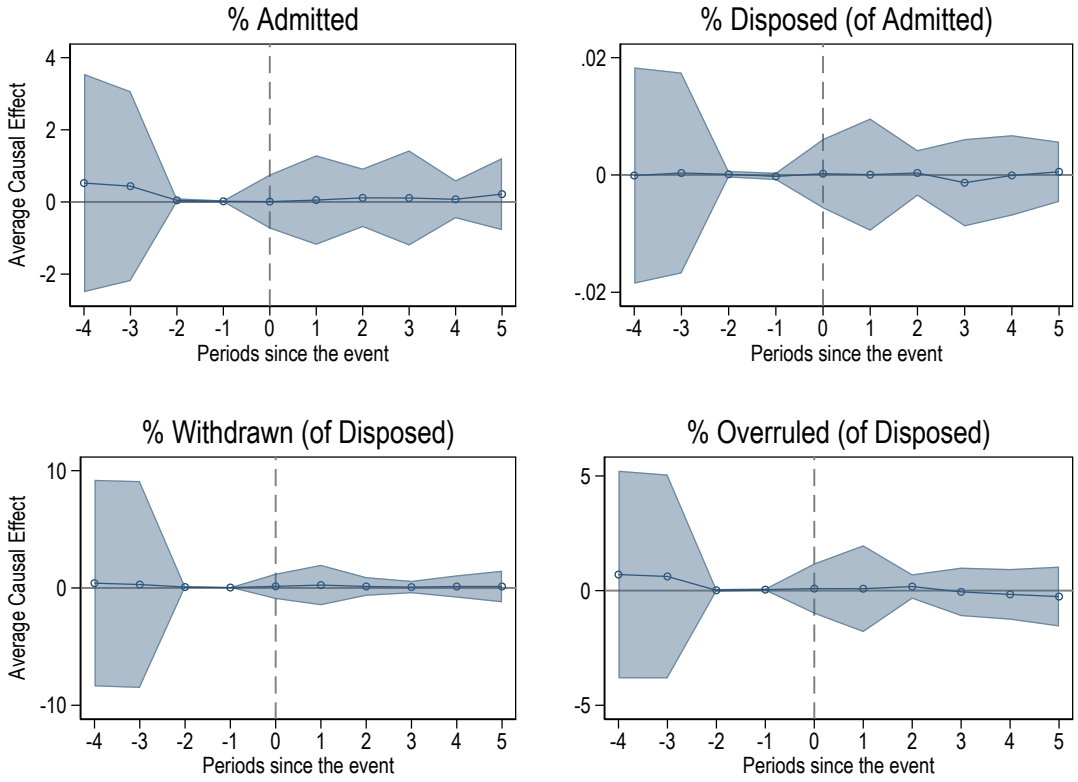


Figure OA10: Robustness Check for Kanoon's Impact on Appeal Outcomes, with variables defined in conditional percentages based on the sample in the previous stage of the pipeline of justice.

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 occurred at the time t-k, for k in (1,2,3,4).

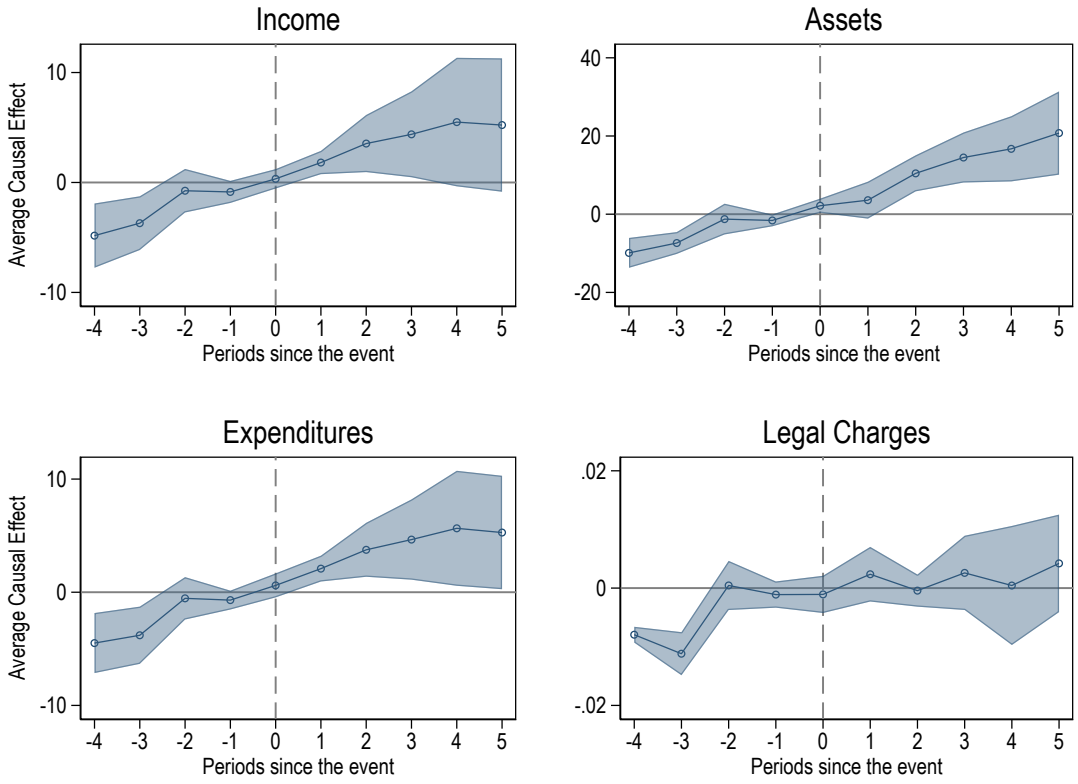


Figure OA11: Robustness Check for Kanoon's Impact on Partial Equilibrium effects on Firm Financials

Notes: This figure shows the results of the Placebo test for testing pre-trends 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.

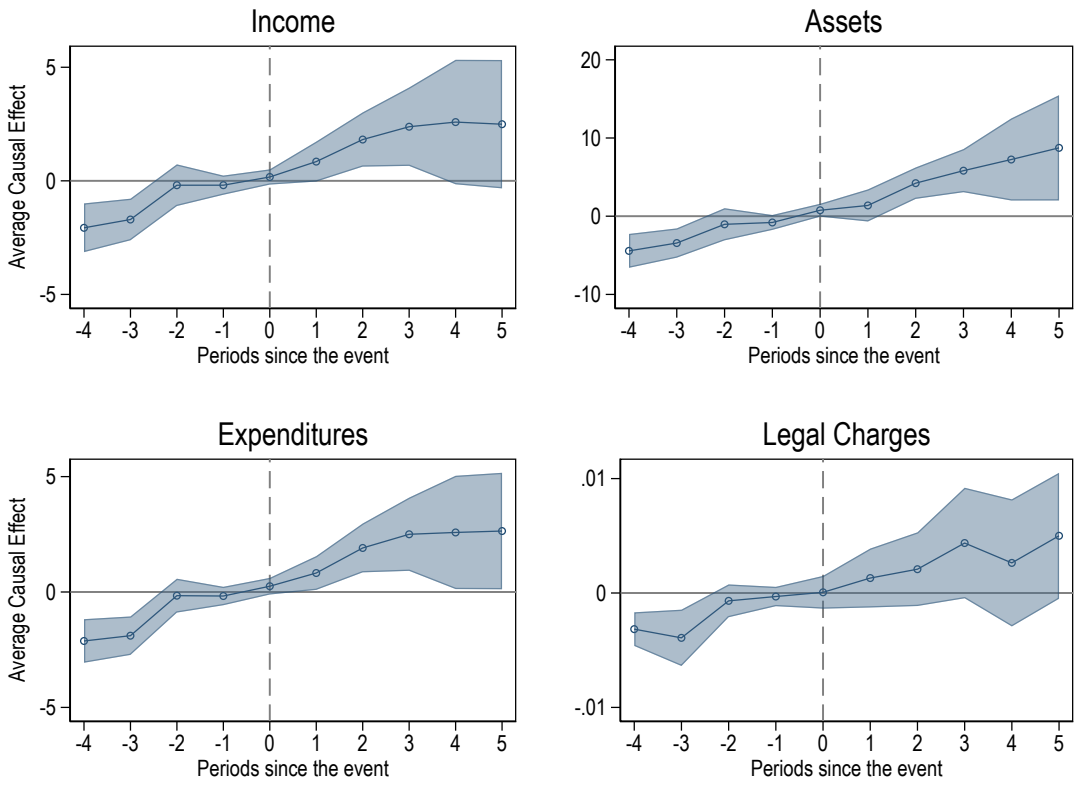


Figure OA12: Robustness Check for Kanoon's Impact on General Equilibrium effects on Firm Financials

Notes: This figure shows the results of the Placebo test for testing pre-trends 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.

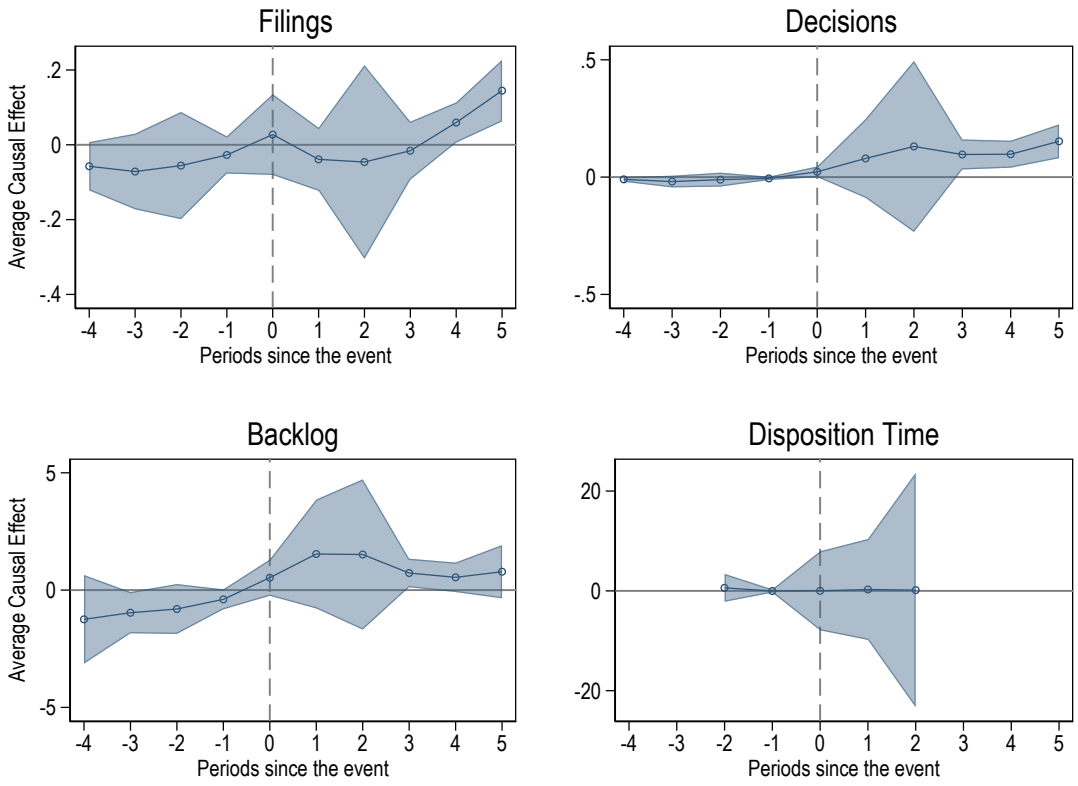


Figure OA13: Robustness Check for Kanoon's Impact on District Court Efficiency Measures for Cases linked to Prowess

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 occurred 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, and month fixed effects. Standard errors are clustered at the state level.