Courtroom Surveillance: Evidence from a State Intervention in the Courts

(Job Market Paper)

Click here for the latest version

Vorada Limjaroenrat*

November 2, 2024

Abstract

Judicial independence stands as the cornerstone of any democratic nation. How does state intervention, specifically through the monitoring of judges, affect the judiciary? To study this question, I collect and transcribe novel case-level administrative complaint and judgment texts from criminal courts in Thailand during the autocratic government period of 2015 to 2020. I analyze the effects of the 2017 constitutional reform mandated by a military coup. This reform requires draft verdicts for severe and politically sensitive cases to be reviewed by a superior court before judgment delivery. Using a difference-in-differences design, I find large and significant impacts on judge behavior: judges became more stringent in cases subject to superior court review. The effect does not fully emerge until 2019 when extra monitoring and enforcement is introduced, and is primarily driven by previously lenient judges. This indicates that the reform increased the uniformity of verdicts as intended. However, heterogeneity analyses revealed that while severe drugs and social stability offenses received longer prison sentences, cases involving politicians and public officers' malfeasance received significantly more lenient treatment. Additionally, these effects are more pronounced in regions that are more politically aligned with autocratic rule.

*Email: vorada.limjaroenrat@economics.gu.se. I am grateful to my supervisors, Randi Hjalmarsson and Andreea Mitrut, for feedback and support throughout the project. I thank Andrea Tesei; Bansi Malde; Warn Lekfuangfu; Jan Stuhler; Luigi Minale; Ricardo Mora; Ronja Helénsdotter; seminars participants at the University of Gothenburg, Stockholm University, the London School of Economics, InsTED Workshop at King's College London, Universidad Carlos III de Madrid, ASWEDE, EALE Annual Conference for comments and suggestions. Financial support from the Royal Society of Arts and Science in Gothenburg (KVVS), Herbert & Karin Jacobssons Stiftelse is gratefully acknowledged. This research has been approved by the Swedish Ethical Review Authority.

1 Introduction

Political parties and judicial review are two distinct institutions that mutually protect and balance each other to ensure the stability of democratic states. In recent decades, however, government interference in judicial processes has increased in various countries. This interference occurs through both formal mechanisms—such as constitutional and legislative changes (like Hungary, Israel, Thailand, or South Africa)—and informal political maneuvers (like Pakistan, or Poland's far-right Law and Justice party). This raises important questions: To what extent should the state intervene with and monitor the courts? What are the underlying motives behind state intervention in the judiciary?

Monitoring is a common tool used to mitigate agency problems in any organization. Economic theory suggests that monitoring can reduce corruption or malfeasance by increasing the cost and likelihood of detection (Becker and Stigler 1974), thereby improving organizational performance (Ferraz and Finan 2008; Ferraz and Finan 2011; Di Tella and Schargrodsky 2003; Borcan, Lindahl, and Mitrut 2017, Vannutelli 2023). However, the context of judicial monitoring differs from other contexts. Judges are highly capable individuals, and monitoring them can compromise judicial independence, potentially weakening the system of checks and balances in democratic countries (Ginzburg and Moustafa 2008).[1]

Despite academic discussion and profound anecdotal evidence of political attacks on the judiciary, the exact motives and consequences of these interventions remain unclear. Very little research has documented the extent to which state intervention in the court affect case outcomes and the judiciary. This paper addresses the gap by providing empirical evidence on the consequences of monitoring judges on case outcomes. Identifying the causal effect of monitoring on case outcomes is challenging given the many factors that affect both monitoring and case outcomes; cases are selected on observed (and unobserved to the researcher) dimensions that are also related to sentences. Thus, previous research has often been limited to qualitative studies or estimating correlations.

To identify the effect of monitoring on case outcomes, I examine a unique large-scale reform implemented by the military-coup government in Thailand. The 2017 judicial reform in Thailand mandated that draft verdicts for selected severe and politically sensitive cases be reported to regional court superiors before judgment delivery. Additionally, final verdicts for these cases must be reported after the judgment if the defendant confesses early, i.e., before the case assignment. These requirements–specifically which cases must be reported–vary

[1] "The court would determined the future of our democracy. It was as important to the new constitution as the parliament and presidency. Judges should be *creative and independent* in ensuring that, in contrast to apartheid, no person was above the law, regardless of their race, power or wealth." (Nelson Mandela, South Africa Constitutional Court opening 1995)

across regions in Thailand.

Although the stated aim of the reform was to increase transparency and uniformity of verdicts across judges, its consequences are unclear and have never been evaluated. With rising public concerns among lawyers and NGOs that the reform could erode judicial independence, the reform was more heavily enforced in 2019 to ensure its effectiveness and continues to be implemented to this day.

To study the effects of the reform, I construct a novel case-level dataset based on Thai Provincial Courts' complaint and judgment texts, which span the period of military control during 2015-2020. These data have never been linked across different courts or used by researchers before. I linked this dataset with (i) individual-level court registry data using anonymized case IDs to obtain case characteristics, defendant characteristics, and prosecuter and judge identifiers, (ii) charge-level punishment range data, which I scraped from the universe of criminal laws in Thailand, and (iii) constituency-level electoral data. By transcribing and coding the exact characteristics of each case, I can identify the precise offense category and the maximum prison sentence range associated with each charge. Using a difference-in-differences design, I exploit suboffense-by-region variation in reform exposure and timing to compare changes in sentencing outcomes between treated cases that contain suboffense types subject to superior court review to control cases that are not subject to additional superior court review.

I find that the reform has a large and robust effect on judges' behavior. Cases subjected to superior court revision before judgment received significantly longer sentences by 38.57 months (107.1 percent increase relative to the mean) following the reform. This stands in contrast with the earlier belief that the reform had no systematic impact on case outcomes in the years that followed. Also, the dynamics of the estimates reveal that although the reform was implemented in 2017Q1, it was only after the introduction of extra monitoring and enforcement in the second phase of the reform (2019Q1) that estimates increased significantly. The results are large and signicant only for cases that are subjected to draft verdict revision before judgment delivery, but insignicant for cases for which final verdicts are to be reported after judgment delivery. This highlights that the success of the reforms depends crucially on strong monitoring and enforcement.

In additional analyses, I demonstrate that the results cannot be explained by changes in the number of crimes, prosecutor behavior, or the case assignment process. Instead, results are mainly driven by previously lenient judges becoming more stringent after the reform. This suggests that the reform increased both punishment stringency and sentencing uniformity across judges.

To better understand the political incentives behind the reform and explore whether it

served as a monitoring tool to reduce judicial bias, a political strategy to gain support, or a means to consolidate power and facilitate state corruption, I document significant variation across offense categories and local political alignments. The reform's impact is primarily driven by severe drug-related offenses, which constitute the majority of serious criminal cases, as well as cases related to political unrest. In contrast, cases involving politicians and public officials' malfeasance received significantly more lenient treatment. Moreover, I find that the effect is more pronounced in areas with stronger political support for the autocratic regime. Overall, my findings provide empirical evidence that the state leverages the courts to expand social and political control and enforce greater uniformity among judges. However, the results also suggest that increased judicial oversight could undermine the rule of law.

This paper contributes to several strands of the literature. First, it speaks to the literature on an autocrat's strategy to maintain political power. Several papers investigate how autocrats manage information flow, through propaganda (Adena et al. 2015; Yanagizawa-Drott 2014), censorship (King, Pan, and Roberts 2014; Rozenas and Stukal 2019; Simonov and Rao 2022), repression (Gehlbach et al. 2024), or collect informtion through local elections (Martinez-Bravo et al. 2022). Other studies have shown that autocrats reward connected politicians (Jia, Kudamatsu, and Seim 2015; Xu 2018) or bureaucrats (Wen 2024) with promotions. Despite the crucial role of judicial capture for autocrats, no paper has empirically documented the consequences of autocratic intervention in the court process. My work builds on this literature by providing case-level evidence of how judicial reforms under an autocratic government affects the rule of law.

Second, this article provides further insight into institutions and development, and particularly the importance of checks and balances on executive power (La Porta et al. 2004). A smaller literature has focused on the role of judicial institutions in economic development. Axbard (2024) studies how convicting corrupt officials affect economic development. Mehmood (2022) explores removals of presidential discretion of judicial appointments (*ex ante*) on case outcomes. I contribute to the literature by studying interference in *ex post* judicial decisions.

Third, this article builds on existing research regarding monitoring as a tool to solve principal-agent problems in the public sector and their implications for state performance. Existing literatures have explored monitoring local government budgets through official auditors (Ferraz and Finan 2008; Ferraz and Finan 2011; Di Tella and Schargrodsky 2003) and through external random auditors (Vannutelli 2023), monitoring exams through CCTV cameras (Borcan, Lindahl, and Mitrut 2017), and central government monitoring of pollution (Axbard and Deng 2024).. I contribute by exploring a new context, monitoring the judges, where the directional effect of the policy is ex-ante ambiguous. Thus, my findings serve not

just as an exercise in quantifying the effects of monitoring the judges but as important inputs for policy debate around judicial independence.

Fourth, this paper contributes to the literature on the organization of the state, and tradeoffs involved in delegating tasks to subordinates (Besley and Coate 2003). There exists empirical works that study the effect of removing discretion in the selection process of state officers, e.g. auditors (Vannutelli 2023) and judges (Mehmood 2022). I contribute to this strand of literature by studying how increases in vertical control, empowering senior judges to screen the verdict of junior judges, could impact the institution's performance.

In the context of criminal justice organization, Downey and Grunwald (2023) study the removal of rule-based prosecutorial discretion in the US, requiring federal prosecutors to charge and seek conviction on the most severe offense in each of their cases. Ernest, Yi Lu, and Wang (2022) examine judicial independence reforms in China that removed local governments' control over local civil courts' financial and personnel decisions. They found that judicial independence can reduce local protectionism, foster cross-regional economic integration, and increase GDP in China. My study builds upon this literature by examining a direct intervention in the judicial decision-making process, which has not yet been explored in the literature.

Finally, the paper also builds upon the literature on bias in the legal system. This includes political bias (Anwar, Bayer, and Hjalmarsson 2019; Poblete-Cazenave 2023), racial bias (Anwar, Bayer, and Hjalmarsson 2019; Alesina and La Ferrara 2014), gender and religion (Ash et al. 2021), and other in-group bias (B. Mullen, Brown, and Smith 1992; Shayo and Zussman 2011; Mehmood 2022).

The remainder of this article is organized as follows. Section 2 provides the institutional background of the Thai judicial system and the 2017 judicial reform. Section 3 describes the data and summary statistics. Section 4 discusses the identification strategy and provides evidence of its validity. Section 5 presents primary estimation results for prison sentence lengths, and Section 6 explores mechanisms by estimating differential effects by judge stringency.

2 Institutional Background

2.1 Thai Court System

The Thai court system is structured into three levels: Courts of First Instance, Courts of Appeal, and the Supreme Court. The majority of cases are resolved at the Court of First Instance, which include the Criminal Court of Thailand in Bangkok Metropolis, 96 Provincial

Courts, 27 District Courts, and 29 Juvenile and Family Courts.

Within the capital city of Bangkok, criminal cases are tried at the Criminal Court of Thailand, Southern Bangkok Courts, Thonburi Courts, Minburi Provincial Court, and seven District Courts. The number of judges presiding over a case varies depending on the court's jurisdiction, ranging from one to five. Outside Bangkok, in other parts of the country, the venue for criminal cases depends on their severity. Cases for which the maximum potential sentence exceeding 3 years in prison, 60,000 baht in fines, or both are tried at the provincial courts. Less severe cases that do not meet these criteria are heard in district courts.

The composition of judicial panels differs between provincial courts and district courts. Cases in district courts are heard by a single professional judge, whereas cases in provincial courts are heard by a panel of two professional judges. In this paper, I focus on criminal cases tried at seven provincial courts outside Bangkok.

In Thailand, career judges are recruited through a merit-based system and appointed as civil servants upon passing a qualification exam. The hierarchy among judges is determined by both the date of passing the exam and the ranking of the exam score. The administration of each local court is led by a Chief Judge, who is appointed from the pool of career judges based on seniority.

All Courts of First Instance outside Bangkok are overseen by one of the nine regional courts, each of which is presided over by a *Chief Regional Judge*. The selection of Chief Regional Judges (see Figure 1) is carried out by the Office of the Judicial Commission (JC).[2]

The Chief Regional Judge can preside over cases within his/her region, including those involving offenses against public security, serious criminal offenses, high-value claims, and contempt of court. Additionally, if necessary, the Chief Regional Judge can temporarily transfer a judge within his/her region to another court for up to three months, with the judge's consent. Any such transfer must be promptly reported to the President of the Supreme Court.

2.2 Thai Criminal Justice Process

The processing of criminal cases in Thailand is summarized in Figure 2. Once a case is filed in court, either by a citizen or through the prosecutor,[3] it first reaches the arraignment hearing stage. At this stage, the defendant can either confess or plead not guilty. Unlike in the US court system, there is no plea bargaining in Thai courts.

[2]The Office of the Judicial Comission (JC) consists of the President of the Supreme Court, six judges from the Supreme court, four Court of Appeal judges, two lower court judges, and two qualified Judicial Committee members.

[3]Thai courts generally accept cases filed by public prosecutors without holding investigative or preliminary hearings, but cases filed by citizen will go through preliminary hearings.

If the defendant confesses and the case involves a minimum sentence of less than 5 years of incarceration, a judge on duty can adjudicate the case right away. These cases, typically small and non-serious, do not require further investigation by the court and proceed without an attorney. A significant number of cases fall into this category, and this streamlined process significantly alleviates the court's workload. As a result, the majority of cases are resolved at the arraignment hearing stage.[4]

However, if the defendant confesses but the case is more serious (with a minimum prison sentence of 5 years or more), or if the defendant pleads not guilty, the case will be randomly assigned to a judicial panel by the Chief Judge of each court. The panel will then conduct further investigative hearings, evidence examinations, and witness hearings.

2.2.1 Case Assignment to judges

The case assignment process for all provincial and criminal courts follows a set procedure. At the beginning of each budget year, the chief judge of each court forms judicial panels, each comprising two judges, through randomization.[5] These panels remain intact throughout the year until the next budget year or annual rotation. The formation of these panels occurs annually before the case assignment stage. On average, a provincial court will have approximately 11 judicial panels.

After the panel is formed and the case arrives at the court, the Chief Judge randomly assigns cases to these panels of two judges, designating one as the responsible judge and the other as a consulting judge, adhering to the "principle of randomization" (Act on Judicial Service of the Courts of Justice concerning Case Assignment, B.E. 2543 (2000)).[6] A case must be heard and adjudicated by both judges in a judicial panel.[7] To prevent the judge from developing a close connection with local people, a judge will be rotated to a new court after a four-year stay and cannot return to the same court again in the future.

2.2.2 Sentencing Practice and Judicial Discretion

In all criminal cases in Thailand, the prosecutor determines charges, for which the sentence range naturally follows from the Criminal Code of Conduct. This sentence range serves as a bound for a judge to exercise discretion over conviction and punishment.

[4] It is, at the same time, criticized of defendant's legal rights– defendants have no attorney and cannot discuss the case with external lawyers.

[5] The chief judge will also designate one judge to be a *judge on duty* for the arraignment hearing process on a daily basis. The assignment of judge on duty usually set a month in advance.

[6] There are exceptions for "specialized cases", e.g. cases involve juvenile offenders, corruption and misconduct cases, complex cases expected to take a longer time to process, all of which will be assigned to more experienced judges by their Division Chief Judge.

[7] according to the Law for the organization of the Court of Justices

Judicial discretion in Thailand can be described in two stages. First, judges determine the final verdict, deciding whether to convict, acquit, or dismiss the case. Judges then have discretion over the punishment, with sentences determined according to the court's internal sentencing guidelines, known as *Yee-Tok*. Judges are recommended to follow *Yee-Tok* closely but can deviate if approved by the Chief Judge. In the final step, judges can adjust the sentence set by *Yee-Tok* under certain conditions specified by the Criminal Code of Conduct.[8]

Judicial discretion in Thailand differs significantly from that in the US or other Western countries, where Sentencing Guidelines are public. For example, in the US context, with publicly available Federal Sentencing Guidelines, sentences are fairly predictable based on convicted charges. Consequently, it is widely agreed that prosecutors have more influence over case outcomes than other criminal justice officials, including judges.[9]

In contrast, Thailand lacks statutory sentencing practices and Thai Sentencing Guidelines differ significantly from those in other countries.[10] First, there are no uniform sentencing guidelines for all courts of the first instance; each court formulates its own guidelines. This is justified by the argument that the number of crimes varies across locations, making it preferable for each court to have its own guidelines. Second, these sentencing guidelines are not publicly available beyond the judiciary. Additionally, there is no written standard for compliance with or departure from *Yee-Tok*, unlike Western sentencing guidelines. Given these differences, judges in Thailand have substantial influence over case outcomes.

2.3 Background and the Implementation of the 2017 Judicial Reform

This section discusses the political background in Thailand leading up to the constitutional amendments and court interventions. Then, I turn to the 2017 Constitutional amendment, which is the primary focus of this paper.

Though Thailand has been a democratic nation since 1932, the country's democratic consolidation is still not complete. Thailand's political landscape has seen multiple regime transitions, swinging between democratic and authoritarian administrations as well as several distinct democratic governments. The Thai political cycle is depicted in Figure 1A of Appendix A, which emphasizes that, on average, a coup was carried out every five years to

[8] These conditions include: (i) converting imprisonment to confinement (Criminal Code of Conduct, Section 23), (ii) granting remission (Criminal Code of Conduct, Section 55), and (iii) suspending or delaying probation (Criminal Code of Conduct, Section 56).

[9] Barkow (2009); Bibas (2010)

[10] Mahakun (1977); Saengwirotjanapat (2016); Yampracha (2016)

put an end to political disputes or instability. After every coup, the nation was ruled by a military junta for a few years until constitutional democracy was reinstated.

Figure 3 displays the political cycle in Thailand since 1997, showing numerous transitions between democratic and authoritarian regimes as well as between left- and right-wing democratic parties, depicted in blue and grey respectively.[11] A noticeable pattern is that anytime a military coup results in a new government taking office, the constitution is amended to weaken the independence of the judiciary.

Prior to 2007, a fully elected legislature drafted Thailand's first constitution on October 11, 1997. The 1997 constitution was referred to as the "people's constitution" and was praised as the most democratic constitution to date since it was drafted by an elected body operating under a democratic government. This constitution included multiple sections that guaranteed judicial independence in hearing,[12] trials and decisions.[13]

The 2017 constitutional amendment, following the 2006 military coup, eliminated several sections of the 1997 constitution that had guaranteed judicial independence. Only one section, Section 197, briefly addresses judicial independence,[14] while granting Chief Regional Judges more authority to relocate judges without prior consent.

This paper focuses on the 2017 constitutional amendment, which followed another military coup on May 14th, 2014. This amendment arguably resulted in decreased judicial independence compared to previous constitutions. Specifically, only one section (Section 188) guaranteed judicial independence, and this provision was significantly shortened. It no longer prevented the transfer of judges without their consent.

This decrease in constitutional protections for judicial independence consequently led to

[11] Thai democratic politics do not have a clear separation between the left and the right wing, but are more separated by polarization of powers. In this paper, I follow this mainstream separation and categorize into two polars: liberal vs conservative. Liberal parties here include Thai Rak Thai (TRT) Party, and the two Thaksin-nominee parties (People's Power party (PPP; 2007), and Pheu Thai party (PTP; 2011)). Conservative party here refer to Democrat (DEM) Party.

[12] Criminal Code of Conduct Section 236 stipulated that "judges who abstained from the hearing could not adjudicate the case, with the exception of situations involving force majeure or other legally mandated unavoidable situation."

[13] Criminal Code of Conduct Section 249 stipulated that: (i) the trial and adjudication by judges shall not be subject to hierarchical supervision, (ii) case assignment to judges shall be in accordance with the rules of law, (iii) case recall from a judge or transfer shall not be permitted except in the case where justice in the trial and adjudication of the case shall otherwise be affected, and (iv) a judge may not be transferred without consent, with the exception of ordinary transfers stated by law/agenda, promotions, disciplinary actions, or entering the criminal court as a defendant.

[14] Criminal Code Section 197: "The trial and adjudication of cases is the exclusive power of the court. Judges have full independence in trials and adjudications. The transfer of judges without their consent is not permitted except for temporary transfers as provided by law, promotions, disciplinary actions, criminal charges, cases affecting the justice of the trial, or force majeure or other unavoidable necessity. Judges may not be political officials or hold political positions."

the amendment of the Act on Judicial Service of the Courts of Justice in 2017. Although the reform intends to "...to increase the uniformity and unbiasedness of verdicts across judge",[15] it arguably decreased judicial independence in several aspects. Below I describe details of the reform and its implementation.

Reform implementation. Prior to the reform, once a case arrived to the court, the Chief Judge randomly assigns it to the judge; the judge had full independence over his/her verdict. After the 2017 reform, court superiors were now allowed to interfere in the sentencing process by adding two additional procedures to the pre-reform criminal justice process. First, the 2017 Act defined a set of 'special cases' for which judges at the Court of First Instance and the Appeals Courts are required to report draft verdicts to the Chief Regional Judges before judgment delivery. These draft verdicts must be sent to the Office of Chief Regional Judges by post immediately upon the case's arrival at the court, and must arrive at least 15 days before the judgement delivery to ensure sufficient time for verdict revision. Chief Regional Judges are legally empowered to review the draft verdicts, write advisory reviews if they disagree with them,[16] and return the draft verdicts to the judges for final judgment delivery. Figure 4 provides a graphical summary of how this judicial reform affects the pre-reformed judicial decision-making process.

The description of 'special cases' category is reported in Table 1. These 'special cases' are generally (political) sensitive or severe cases, which include severe criminal cases,[17] severe civil cases,[18] contempt of court cases, autopsy investigation, cases that are of public interest, cases that may affect international relations, cases relating to conflict in the southern border provinces, and cases related to top individuals.[19] Understanding what offense categories are classified as 'special cases' and being able to observe these in the data is an essential component of the research design. Section 3.2 returns to these details and summary statistics of these special treated cases.

Regional discretion. Due to variations in location and the number of crimes across regions, the 2017 Act allows each regional court to have some discretion in how the reform was

[15]Media interview of the spokesman for the Court of Justice, *ThaiPost* on October 8th, 2019.

[16]It has been confirmed that an advisory reviews by the Chief Regional Judge act as a guideline or recommendation. There is no hard rule stipulate that the judge cannot deviate eventually.

[17]These cases include offenses related to national security (Criminal Code of Conduct, Sections 107-135), lèse-majesté (Criminal Code of Conduct, Section 112), sedition (Criminal Code of Conduct, Section 116), terrorism (Criminal Code of Conduct, Section 135/1-135/4), criminal cases with maximum prison sentences exceeding 10 years, and severe narcotics offenses.

[18]These cases include those where the property in dispute or the value of claims is at least 5 million baht, civil cases in which a financial institution is the plaintiff and the assets in dispute exceed 10 million baht, and property ownership disputes involving assets valued at a minimum of 200,000 baht.

[19]E.g. the prime minister, cabinet members, senators, MPs, judges, public prosecutors and senior civil servants

implemented at the regional level. In my data sample, Region 3 and 4 did not make further adjustment from the National Act, only Region 1 (central region) deviated. To be precise, Regional Court 1 issued an immediate adjustment guideline (No. 301.101/v145) on August 29th, 2017 to be enforced for all courts in the central region. Citing a high caseload and small number of judges, the Region 1 guideline specified a higher threshold than the national act for cases subject to review. For example, the cutoff for criminal cases requiring revision was set at 20 years, and drug offenses were required to be reported if the pure substance exceeded 50 grams. The details of the heterogeneity of treatment across regions will be described in Section 3.2 and Appendix A.2.

Extra enforcement. To increase the enforcement of the 2017 Act, a *2019 Act on Judicial Service of the Courts of Justice* was implemented for all regions. This new Act added an extra layer of moniotoring to the 2017 act in several dimensions. First, the previous enforcement for 'special cases' from the 2017 Act was emphasized in the 2019 Act. This resulted in each regional court implementing the policy with higher enforcement. Some regional courts, such as Region 1, issued extra announcements reminding judges to report the draft verdicts with full details. Second, the 2019 Act stipulated an additional requirement for the 'very special cases', a subset of the 'special cases', to be reported not only to the Chief Regional Judge, but also to be further reported to the President of the Supreme Court immediately once the case arrived at the court. Third, due to the high number of reported cases, most regional courts replaced the old process of reporting draft verdicts by post with an electronic system to facilitate the reporting process.[20]

As described, while the goal of the reform is to increase monitoring and reduce corruption, it potentially undermines judicial independence. Unsurprisingly, the reform has sparked public debate,[21] particularly among independent media,[22] human rights lawyers, and academics.[23] Their main concern is that reduced judicial independence could weaken the checks and balances crucial to democratic governance.

[20]The exact implementation date of an electronic report systems are unknown and different for each regional courts, but occur between 2019-2020. Here, I gather anecdotal evidence from each region where the starting date of electronic case report system implementation is found. For instance, region 1 required the verdict to be reported by email since the start of 2019 Act on Judicial Service of the Courts of Justice implementation. Region 2 employed a newly created Electronic Case Report and Management (ECRM) on May 17th, 2019. Region 4 piloted ECRM system since 2019, but fully implemented on Jan 4th, 2020 for all courts in the the region.

[21]An issue received most media coverage after a judge in Yala Provincial court who handled a case subjected to revision shot himself in a courtroom in 2019, with a statement that "...many unfairnesses occurred in the court reviewed process and my virdict is amended unfairly by the Chief Regional Judge"

[22]See iLaw: https://www.ilaw.or.th/articles/3785

[23]See Thammasat University seminars "Returning the Judgment to the Judge, Return Justice to the People" held on October 11th, 2019

3 Data

3.1 Data Description

The primary data sources consist of case-level complaint and judgment texts, summarized by court clerks. I collected this dataset from seven provincial courts outside the capital city Bangkok, distributed across three regions: central (region 1), lower north eastern (region 3), and upper north eastern (region 4).[24] Because these data were not centralized and were stored locally, the collection process depended heavily on the cooperation and infrastructure of each local court, as well as the workload of officers at the central Court of Justice (COJ), who linked and anonymized the dataset for the first time. Given its sensitive nature and the political context, data like this have not been used to empirically study judicial decisions in Thailand before. The resulting dataset is highly unique both in the context of Thailand and other developing countries.

Through an annonymized unique case ID, the data is linked with courts' registry data that contains case characteristics, defendants' characteristics, and judgment informtaion. All individual names (prosecutor, defendant, judges) are anonymized by the Thai Court of Justice and replaced by a uniquely generated random ID. Case characteristic information include annonymized case ID, court name, case filed date, charge category, and prosecutors' ID. Defendant characteristics consist of anonymized ID, age, gender, and nationality (e.g. Thai, Chinese, Laos, Burmese). Judgement information covers judgement date, judge ID (for the judge on duty and the two judicial panel), and case outcome.

The data is then linked to two other sources: (i) charge-level punishment range which I parse from the universe of Thai criminal laws,[25] and (ii) constituency-level electoral data. By linking court data to charge-level punishment range and coding the precise offense categories from court documents, I can identify the exact policy implementation a detailed suboffense level. Furthermore, by connecting the court data with electoral data, I am able to analyze the political incentives behind the reform. Detailed data source is described in Appendix B.1.

The final sample forms an unbalanced panel, including all closed criminal cases from these seven provincial courts, with judgment dates between January 1st, 2015, and December 31st, 2020.

[24] Sample distribution across region is decribed in Table 10 in Appendix B.4.

[25] There are approximatedly 500 criminal laws in Thailand, including Criminal Code, Acts, and Royal Decrees. All criminal laws can be downloaded from Office of the Council of State's official website.

3.2 Defining Treated and Control Cases

To code cases treated by the 2017 reform, I transcribe complaint and judgement texts using both automated and manual regular expressions to obtain key variables that determined if a case will be treated. These include the case-level "maximum prison sentence range", offense category, detailed information on seized drugs (such as drug size, tablets, and pure substance), and whether the case involves high-profile individuals (politicians/bureaucrats).

The reform specified two types of treated cases, distinguished by whether the defendant confesses before the case is assigned to a judge. The first type consists of "special cases" that must be reported to the Chief Regional Judge before the judgment is delivered. The second type also involves "special cases" but in these, the defendant confesses before the case assignment process. For this second type, judges are not required to submit a draft verdict before delivering the judgment. Instead, they only need to report the final verdict to the Chief Regional Judge after the judgment is delivered.

Table 1 provides a summary of all treated case category, along with number of observations. A detailed description of the treated cases and the process for scraping these variables is provided in Appendix A.2 and B.2. The number of treated cases is reported separately for cases in which draft verdicts must be reported before judgement delivery ($Treated^1_{gr}$), and cases where only the final verdict must be reported after judgement delivery ($Treated^2_{gr}$). As a reminder, whether a verdict must be reported before or after judgment delivery depends on whether the defendant confesses before the process of case assignment or not.

One of the key variables that determine a case's treatment status, the case-level "maximum prison sentence range", represents the upper bound of the prison sentence range from the most severe charge in each case. To crreate this variable, I first scraped all charges from each case. Then, I extracted the charge-level prison sentence range from the complete set of Thai criminal laws and linked these ranges to the corresponding charges. By associating each charge with its respective prison sentence range, I was able to rank all charges according to their punishment severity and identify the most severe charge (i.e., the charge with the highest prison sentence range). The upper bound of the prison sentence range for the most severe charge then serves as the case-level maximum prison sentence range. Details of how treated cases and other variables are scraped are described in Appendix B.2.

Table 1 demonstrates that whether an offense is eligible for treatment is determined by two factors: (i) offense category and (ii) charge severity (measured by a case maximum prison sentence range). I therefore create a new treatment unit, g (suboffense-level), that summarizes these two aspects of policy implementation. To explain, a case will be treated if it has a maximum prison sentence range that is above the regional-level cutoff, or if it is an offense that falls into the reported category. A case with suboffense g in region r is treated

if $Treated^j_{gr}$ turns on. This is defined as the maximum between two binary indicators:

$$Treated^j_{gr} = max\{I(Offense_o), I(ChargeSeverity_m > Cutoff_r)\} \quad (1)$$

$I(Offense_o)$ is a dummy equal to one if a case fall in to a reported offense category. $I(ChargeSeverity_m > Cutoff_r)$ is a dummy equal to one if a case has a maximum prison sentence range higher that the region-level cutoff, $Cutoff_r$, which equals 20 years for courts in region 1 (four courts) and 10 years for courts in region 3 and 4 (three courts).

There are 19 offense groups (o), and 4 maximum prison punishment bins (m). The index g summarized both offense (o) and charge severity (m) aspects. Thus, g captures variation at the suboffense level, and $Treated^j_{gr}$ varies at the suboffense-by-region level.

3.3 Sample Creation and Descriptive Statistics

To construct the baseline sample, I first exclude cases adjudicated by the judge on duty and focus on cases that are randomly assigned to the regular judges. These are cases in which the defendant confesses and the minimum prison sentence range is below 5 years. By excluding these non-randomly assigned and relatively less severe cases, I am better able to construct a comparable control group.

As described earlier in Table 1, treated cases are mostly cases with maximum prison sentence lengths above 10 years. To construct a comparable control group, I limit the non-treated cases to those with a maximum prison sentence lenghts greater than or equal to 10 years.[26] This resulted in an analysis sample of 74,678 case-by-defendant observations (69,759 defendants). The sample contains 345 judges (352 judge on duty and 457 panel judge), 226 prosecutors, and 85 crime scene districts.

Table 2 displays descriptive statistics at the case-by-defendant level. The first column shows statistics for the full sample (N=74,060). Statistics for control cases and treated cases are reported in columns 2 and 3 respectively. As the policy contains two types of treated cases (depending on whether defendant confesses prior to case assignment to the judge), column 4 reports statistics for $Treated^1_{gr}$ (N=5,271), and column 5 reports statistics for $Treated^2_{gr}$ (N=21,658).

By restricting the control cases to those with a maximum prison sentence range of 10 years, defendants and case characteristics are more comparable in pre-determined characteristics. However, there is a significant difference in the average processing time between the treated and control cases. The average processing time for treated cases is 52.2 days, nearly triple the average processing time of control cases, which is 22.3 days. This discrepancy is

[26] Distribution of cases' maximum prison sentence range can be found in Table 9 of Appendix B.4.

expected, as treated cases are typically more serious, and it may take additional time for them to undergo review by the court head.

In Table 2, cases are grouped into 9 crime categories. A case is assigned to a particular crime category if it contains at least one offense from that category.[27] There are differences in case categories between the control and treated groups. Most common in the control group are minor drugs offenses and major drugs offenses, whereas in the treated group, the most common categories are major drugs offenses and followed by minor drugs offenses. But drugs offenses make up the lion's share of both the treated and control groups.

Regarding sentences, Table 2 indicates that treated cases on average receive more severe prison sentences than control cases. Treated cases are almost twice as likely to result in incarceration compared to control cases (incarceration rate is 38.7 percent for control cases and 87.2 percent for treated cases). Moreover, among those incarcerated, the sentence lengths for treated cases are over triple the sentence lengths for control cases.

For the empirial analysis, I set prison sentence lengths to zero if cases are acquitted or received no prison sentence. Life imprisonment is set to 60 years of prison, and capital punishment is set to 100 years of prison.[28]

4 Empirical Strategy

4.1 Graphical Evidence

This section begins by graphically demonstrating the impact of the reform. Figure 5.A illustrates the average prison sentence lengths over time, with the reform period indicated by vertical lines in 2017Q3 and 2019Q1. Prior to the monitoring reform in 2017 and leading up to the additional monitoring introduced in 2019, the average sentence lengths remain fairly steady for both the control and treatment groups. In the first period following the introduction of extra monitoring in 2019, the average prison sentence lengths more than double for the treatment group but remain fairly stable for the control cases.

[27] Political sensitive offense include offenses against the royal (Criminal Code of Conduct, Section 107-112), offenses related to security of the kingdom (Criminal Code of Conduct, Section 113-135), offenses related to terrorism (Criminal Code of Conduct, Section 135/1-135/4), offenses against public officials (Criminal Code of Conduct, Section 136-146), offenses against judicial officials (Criminal Code of Conduct, Section 167-199). Malfeasance in offices (Criminal Code of Conduct, Section 136-146, 200-205).

[28] This follows from the Criminal Code of Conduct Section 91(3) that the maximum prison punishment (less severe than life imprisonment) a case could received is 50 years. It is commonly viewed that life imprisonment is a more severe punishment than 50 years of prison sentence. Also, folloing from Criminal Code of Conduct Section 52(2), a capital punishment were to be adjusted down by 1/2, it shall be reduced to life imprisonment or 50 years of prison. Thus, it is natural to set capital punishment to 100 years of prison sentence.

This figure points to the first graphical evidence of the reform. Although the monitoring reform was introduced in 2017, the effect becomes evident only after the implementation of extra monitoring and enforcement in 2019. To interpret this figure, it is important to note that there are several compositional changes over time that could drive this observed pattern. As a result, controlling for these factors are necessary for causal inference. This figure also supports the exogeneity of the refoms. For these changes to be driven by unobserved factors, one would have to believe that those confounds coincided with the exact case arrival quarters of the two monitoring regimes and treated offense categories in all regions. Given that the our sample includes several provincial courts that the spread over three different regions, this seems unlikely.

4.2 Empirical Methodology

I study the effect of case monitoring by exploiting a judicial reform in Thailand that generates exogenous variation in case monitoring at the *suboffense-by-region* level. In essence, I compare changes over time among (i) treated cases that must be reported to the Chief Regional Judge (as described in Section 3.2) and (ii) non treated cases with a maximum prison sentence of at least 10 years. Thus, the treated and control cases have arguably similar characteristics, except for the level of case monitoring. Table 3 shows that pre-determined case characteristics are balanced betweeen the treated cases and control cases.[29]

I estimate the following baseline difference-in-differences specification:

$$\begin{aligned} S_{igrt} &= \alpha_\tau Treated^j_{gr} + \beta_\tau Treated^j_{gr} \times Post_t + \gamma X_i \\ &\quad + \theta_o + \theta_m + \theta_t + \theta_r + \theta_r \times t + \alpha_j + \epsilon_{igrt} \end{aligned} \tag{2}$$

Here, S_{igrt} represents the prison sentence lengths for individual i, charged with sub-offense g, in region r and quarter t. $Treated^j_{gr}$ is a binary indicator which equals one if a case is classified as a reported category, as described in Table 1. $Post_t$ is a dummy equal to one if the case arrives at the court after 2017Q3. Control variables (X_i) include defendants' age, gender, criminal history, and number of charges per case.

The baseline difference-in-differences specification also includes: offense fixed-effects (θ_o) to control for baseline differences in case characteristics and sentencing outcome across offenses; charge severity fixed-effects (θ_m), measured from a 5-year maximum prison sentence range bin fixed-effects, to control for differences across charge severity; quarterly fixed-effects

[29] Full balancing tables are reported in Table 14, 15, and 16 of Appendix D.

(θ_t) to capture shocks to criminal justice system that are common across all offenses; region fixed-effect (θ_r) to capture time-invariant region specific effect; region-specific quarterly trend ($\theta_r \times t$) to capture court-specific shocks that affect all offenses in that region; and judge fixed-effects (α_j) to control for time-invariant judge-specific effect. Standard errors are clustered at the treatment unit, gr, level. There are 121 gr units in total. The coefficient β_τ then identifies whether cases that are treated by monitoring reform experience a differential change in sentencing outcomes.

Four identifying assumptions underlie my DiD strategy. First, the parallel trend assumption requires that punishment severity of the control and treated cases would exhibit parallel departures from their trends in the post-reform period in the absense of the reform. While Figure 5 provides the first suggestive evidence of parallel pre-trends, one might still worry about trends in punishment severity prior to the reform, particularly if such trends were differential with respect to the policy treatment, e.g. offense category or case severity. I formally test for this by including offense and charge severity fixed effects when estimating an event study specification (see Section 5.3.1). This specification also allows me to look at the dynamic impacts of the reform over the post-reform period.

Second, the DiD strategy assumes that the timing of the offense-charge severity treatment is exogenous. Although the 2017 Constitution shortened the section on court protection, there were no indications of case revisions or state intervention in the court process until the announcement of the 2017 Act on Judicial Service of the Courts of Justice on July 27th, 2017. This act stated that the effective implementation date would be for all treated case that were filed from August 26th, 2017 onwards. A test for no anticipatory effects will also be provided in the event-study analysis (see Section 5.3.1).

Third, causal inference relies on the assumption that there are no confounding factors. The assumption would be violated if the reform changed the composition of the cases, or affected the pre-judicial criminal justice process in such a way that it affects the quality of evidence presented to judges and leads to different observed sentencing outcomes (e.g. via changes in prosecutor charge decision, or the case assignment process to judges). If that is the case, it would be less convincing to claim that the results are driven by change in judicial sentencing behaviour. I discuss and formally test for each of these potential confounding factors in Section 5.3.2.

Fourth, the timing of the reform must not coincide with other events or reforms that could lead to the observed results. There are a number of reasons to support this assumption. First, there is no other criminal justice system reform or changes in punishment severity during the period of study that could potentially affect punishment severity of only the treated case. Also, when estimating the result, I include quarter fixed-effects to control for any shocks

that happen to all cases in that quarter. This would wipe out shocks that affect all offense in a quarter, e.g. financial or political cycles, or criminal justice reform that affect all cases.

5 Main Results

5.1 DID Results

I first examine the effects of exposure to superior court review on case outcomes by estimating equation (2) when the dependent variable is the prison sentence length, in months. The results are reported in Table 4. Column (1) shows the results when controlling only for quarter effects, offense effects, and charge severity effects. Column (2) adds a set of observable control variables to see if the results are robust to conditioning on the observables, and column (3) includes all controls as specified in equation (2). Being subjected to superior court review prior to judgment delivery significantly increases prison sentence lengths by 37.4 months. This is a large increase given the average prison sentence lengths of 36 months, i.e. more than 100%. Controlling for judge fixed-effects in column (4), the effect has little effect on the estimate (sentence lengths increased on average by 38.6 months). These results offer substantial, significant, and robust evidence that the reforms impact judges' behavior. The findings suggest that the state intervention in the court process, which aimed to exert more control, led to stricter punishments.

I also explore other aspects of the reform implementation. Firstly, given the nature of the reform, which was first implemented on 2017Q3 and then reinforced with extra monitoring and enforcement in 2019Q1, column (5) allows for differential treatment effects along these two treated periods. In particular, $Post_t$ is separated into $Post_t^{2017}$ and $Post_t^{2019}$. The effect of the reform is driven mainly by the implementation of extra monitoring and enforcement. This pattern is also consistent with the event study in Section 5.3.1.

Several factors may explain why the effect of the 2017 reform did not kick in until 2019. One potential explanation for the introduction of "extra monitoring" in the first quarter of 2019 is that the 2017 third-quarter reform was not strictly enforced. This lack of strict enforcement provides the rationale for the introduction of an additional layer of oversight from the President of the Supreme Court in 2019. Secondly, the 2017 reform was supposed to be implemented almost immediately after its announcement, even though each regional court may have not had sufficient time and/or infrastructure. For instance, before 2019, draft verdicts were sent to the regional courts for revision via sealed post. Combined with the very high number of reported cases, which later prompted the initiation of an electronic case reporting and monitoring system in 2019, this could made it difficult to initially implement

the reform efficiently. Thirdly, although judges' promotions are well-protected and based on seniority, the implementation of extra enforcement measures through the 2019 Act suggests the policy's importance. Lax enforcement could result in indirect consequences, such as slower promotions or transfers to remote courts. Finally, local politics could also play a role. Although a military coup occurred on May 14th, 2014, it was not until the 2019 general election that the military government participated in an election and installed a former coup leader as prime minister.[30] This turnover suggests the emergence of politicians aligned with the new junta at both national and local levels, potentially leading to stronger enforcement of the autocrat's monitoring policies at the judiciary.

Column (6) of Table 4 next separates the treatment dummy into cases report prior to judgment delivery ($Treated^1_{gr}$) and cases report after judgment delivery ($Treated^2_{gr}$). The results are large and significant only for cases that are subjected to draft verdict revision before judgment delivery, but insignificant for the latter. Taken together, these two results highlight the importance of extra monitoring and enforcement in the effectiveness of case monitoring reform.

To understand the magnitude of the estimates, the mean post-reform estimate of 38.6 months is substantial (107.2 percent increase), especially considering that our control groups consist of only serious cases. For instance, Downey and Grunwald (2023) demonstrates that when the government intervenes by explicitly requiring federal prosecutors to become more stringent—by charging and pursuing the most severe offense—prison sentence lengths increase by 4.39 months (16.9 percent increase) more in districts where the U.S. Attorney implements the policy with greater fidelity compared to districts where the policy is implemented with less fidelity.

5.2 Robustness checks

Panel A of Table 5 assesses the robustness of the results. The baseline result from Table 4, which includes the full set of controls, is shown in Column (1). To test the robustness of policy implementation across different locations, column (2) replaces region effects and region-specific trends with (provincial) court effects and (provincial) court by entry quarter fixed-effects. This adjustment accounts for the fact that each court may have its own local policies or court-specific effects, such as case randomization at the court level, that affects all cases within that court. In column (3), I further include offense-specific time trends in the baseline estimation to control for crime-specific dynamics, such as legal reforms that affect

[30] An election took place on March 24th, 2019. An election was, however, widely criticised for its fairness due to the creation of an electoral system designed to favour the junta's newly created political party, manipulation of election rules, a biased voting environment.

certain offenses nationwide. The results remain robust.

To demonstrate that the results are robust to exclusion the harshest punishments, such as capital punishment (set to 1200 months of prison sentence) and life imprisonment (set to 720 months of prison sentence), column (4) excludes cases that received capital punishment, and column (5) excludes cases that received either capital punishment or life imprisonment. These specifications, however, do not change the qualitative nature of the results.

Panel B of Table 5 assesses the robustness of column (5) of Table 4, i.e. the results corresponding to the two treatment periods. The baseline finding holds; the effect of the reform is less stark during the early implementation,but significantly larger after the 2019 extra enforcement period.

Panel C of Table 5 provide robustness of column (6) of Table 4, i.e. when there are two treatment grops. This confirms our previous conclusion that monitoring is only effective if the draft verdict is required to be reported before judgment delivery.

5.3 Identification Checks

5.3.1 Test of parallel trends and random timing assumption

In this section, I test whether the treatment and control groups would have exhibited parallel departures from their trends in the post-reform periods in the absence of treatment. The omitted period (τ=-1) is the period before the reform is fully implemented or the quarter of the reform itself (2017Q3). Estimates are displayed in Figure 6.

Prior to the reform, I find little evidence of differential group trends. For $\tau<0$, all treatment coefficients never reach statistical significance at the 5% level. Pretreatment estimates are also jointly insignificant ($F = 2.04$, $p = 0.40$). This supports both the parallel pre-reform trends and that there are no anticipatory effects; the reform is not preceded by observable changes in local crime or differential punishments.

Following the reform, the effects remain insignificant at the 5% level for the next four quarters (the magnitude of increase are between 16.6 to 39.2 months, mean=23.3). However, the estimates increase thereafter and are statistically significant in all periods (except τ=8) –sentence lengths immediately increase by about 83 months and by about 100 months in each of the following two quarters.

The dynamics of the estimates can be explained by the institutional setting of the reforms. Although the reform is implemented in 2017Q1, effects in the first phase of the reform are not significant. It was only after the introduction of extra monitoring and enforcement in the second phase of the reform (2019Q1) that estimates increased significantly. The insignificant estimate associated with τ=8 aligns with anecdotal evidence from Internal Regional Court

announcement No.301.001/v171, which announced a temporary pause in case monitoring from November 27, 2019, to December 31, 2019, due to the high caseloads and numerous public holidays during the New Year festivals.

5.3.2 Test of no confounding effects

The results thus far show that the estimates are robust to controlling for an extensive set of case and defendant characteristics, which are likely to proxy for many unobservable confounders. In this section, I provide further evidence that the results are driven by the behavior of judges and not other channels. For instance, the reform could change the composition of cases, the behaviour of prosecutors, and the behaviour of the court head in case assignment, in such a way that it leads to observed longer prison sentence. Here, I discuss each of these potential channels.

Change in number of cases. The number of cases could change, for example, if the reform caused future offenders to behave differently and commit more serious crimes, resulting in longer observed sentences. I present a number of pieces of evidence against this claim. First, during the reform period, there is no significant increase in the number of treated case. The number of cases over the case arrival date (quarter) is reported in Panel A of Figure 7 and is relatively smooth. I also conduct a formal regression test in Table 11 of Appendix C. The results suggest that the number of treated cases significantly decreased by 46.9 percent. This confirms our identification assumption that the reform did not affect offenders' behaviour, in such a way that it leads to longer prison sentence. On the other hands, it points to the potential that the reform could reduce crime. This observation, however, deserves further investigation since harsher punishment could reduce crime, if it is visible and known by the people.

Change in composition of cases filed. I examine whether the reform influenced prosecutorial behavior—for instance, whether prosecutors became aware of the reform and began filing cases in higher maximum prison sentence ranges. Prosecutors might, for example, strategically adjust charges to fall just below the case reporting threshold to avoid increased monitoring by court superiors. Similar behavior has been documented in other contexts, such as in the US context, the *Three Strikes* Law led prosecutors to charge offenders with lesser crimes not targeted by the law (Bjerk 2005). If this is happening in our case, the observed shift in prison sentence ranges may be driven by changes in prosecutorial behavior rather than changes in judicial decisions.

I again present a number of pieces of evidence against this claim. First, a graphical representation from Figure 12 (in Appendix B.4) shows number of cases for each charge

severity bins are relatively smooth. There is no significant jump at the time of the reform. Second, Table 12 of Appendix C shows a formal balance test for changes in offenses and charge severity, controlling for prosecutor fixed-effects. The results show that, while there is a change in a number of cases in certain offense groups (drugs and firearms), there is no significant change in the number of cases based on charge severity.

Change in case assignment of cases to the judge. Here, I test for potential manipulation in the case assignment process. For instance, the reform might have induced the court head to assign treated cases, which are subject to revision by a court superior, to a strict judge instead of using the random case assignment process typical of the normal period.

To test for this assumption, I first estimate judge stringency from the pre-reform period sample based on (i) prison sentence lengths (months), and (ii) share of actual prison sentence over maximum punishment,[31] which ranged between zero (most lenient) and one (most stringent).

The leave-out mean judge stringency score $(Z_{j(i)})$[32] is calculated from:

$$Z_{j(i)} = \frac{1}{n_j - n_{j(i)}} \left(\sum^{n_j} S_{j(d)} - \sum^{n_{j(i)}} S_{j(d)} \right) \tag{3}$$

where $S_{j(d)}$ is the prison sentence lengths (or share of actual prison sentence over maximum punishment) of defendant d who is incarcerated by judge j. n_j is the total number of defendants handled by judge j, $n_{j(i)}$ is the number of defendants in case i handled by judge j. $Z_{j(i)}$ thus represents the average prison sentence lengths in other cases judge j has handled. The pre-reform sample that is used to construct an index has 228 judges, each of whom has presided over an average of 113 randomly assigned court cases.

I aggregate this case-level stringency score[33] into judge-level stringency score. Figure 8 displays the distribution judge-level stringency measures, constructed from both sentence lengths and share of actual over maximum punishment. I utilize this measure to categorize judges as either 'strict' if their stringency score is above the median, or 'lenient' if the score is below the median.

I then test whether judge stringency measure in the pre-reform period predicts the probability of a judge being assigned a treated case after the reform implementation. The results are reported in Table 13 of Appendix C. Column (1) reports the result at the judge-level.

[31] See Appendix B.3 for variables construction.

[32] E.g. Kling (2006); J. Doyle (2007); Maestas, K. J. Mullen, and Strand (2013); French and Song (2014); Dobbie and Song (2015); Aizer and J. J. Doyle J. (2015); Bhuller et al. (2020)

[33] The distributions of case-level judge stringency measures (controlling for fully interacted quarter and court dummies) are reported in Figure 15 of Appendix D.

Columns 2 and 3 report the results for the case-level regression. The results show that pre-reform judge stringency does not significantly predict judge assignment to treated case, implying that the reform does not change the assignment of cases to the judge.

5.4 Uniformity across Judges

As each judge's draft verdict is reviewed by the Chief Regional Judge, the succes of the reform should imply lower verdict variation within region. To address this hypothesis, Figure 16 depicts the distribution of judge stringency (control for a fully interacted region and year fixed effect) before and after the reform, with the corresponding summary statistics provided in Table 18 in Appendix D. We can see from Figure 16 that the distribution of both measures of judge stringency shift to the right (mean increased from 34.1 to 36.8). Interquartile range for sentence lengths stringency decreased from 11.2 to 7.8 (Table 18 in Appendix D)[34]. This suggest that the reform increase court punishment and consistency across judges.

To support this evidence, I also explore differential responses to the reform by pre-reform judge stringency. I estimate equation (2) on prison sentence lengths, separately for the strict judge and lenient judge subsample. Strict judges are defined here as those with pre-reform judge stringency scores above the median, and lenient judges are judges with pre-reform judge stringency scores below the median. Judge stringency scores are calculated as described in Section 5.3.2.

As shown, I find stark differences in effects between strict judge and lenient judge. Columns (1) and (2) of Table 6 Panel A show that, by being subjected to superior court review prior to judgment delivery, cases handled by lenient judges have prison sentence lengths significantly increased by 79.5 months. However, being subjected to superior court review does not significantly increase prison sentence lengths for cases that were handled by strict judges.

To assess the robustness of the result, Columns (3) and (4) of Table 6 present the estimation results from using alternative measures of judge stringency, constructed from the share of prison sentence over maximum punishment. The results are consistent with columns (1) and (2), where it shows that treated cases handled by lenient judges were treated significantly more stringent, but insignificantly more stringent for strict judges.

As the policy was implemented differently by region, it is natural to explore the results for each region. However, the number of treated case will be very small if we were to consider each region separately for lenient and strict judges. Keeping this caveat in mind, Panel B of

[34]The fact that interquartile range for share over maximum sentence lengths stringency increased from 0.0019 to 0.021 implies that although distribution of judge stringency shift to the right, the increase in stringency at a lower distribution is smaller or bounded by the maximum punishment.

Table 6 estimates a similar model as Panel A, but separating the $Treated^{1}_{gr}XPost_t$ dummy in equation (2) for each region. The results are aligned with Panel A–treated cases are given longer sentences by lenient judges, though it is insignificant in some regions due to small samples. The results however give a mixed sign, and are insignificant when considering the sample of strict judges.

In interpreting these results, it is important to note that some regions could have few harsh judges when measured according to the whole sample. If that is the case, most judges in that region will be classified as 'lenient' judge in our measure. As the verdict is reviewed by the Chief Regional Judge, it is important to understand how the Chief Refional Judge can influence the behavior of judges in the region.

To better understand the heterogeneity of judges within a region, I categorize judges as either lenient or strict based on their stringency scores compared to the regional median. To be precise, a judge is classified as lenient if their stringency score is below the regional median, and strict if the score is above the regional median. Table 17 in Appendix 64 reports the result with this new judge classification. The results mirror those found in Table 6 with significant, large positive treatment effects for lenient judges and insignificant, small estimates for strict judges. Looking at estimates that allow for differential treatment effects across region in Panel B, the reform significantly increase sentence lengths in region 1 by 90.5 percentage points (column 1) and 110.1 percentage points (column 3). Though estimates are insignificant for regions 3 and 4, they are positive and large, i.e. more than 100% in all regions.

Taken together, these results imply that the reform increases the uniformity of the treated verdict, in the sense that it changes the behaviour of lenient judges to become more stringent, while little to no effect is seen for the strict judges.

6 Mechanisms

The previous sections established that the reform increased punishment stringency and uniformity across judges. However, without knowing a defendant's actual guilt, it is difficult to determine whether the observed uniformly harsher punishments are driven by increased monitoring (if the original verdict was too lenient) or increased manipulation (if the original verdict was fair). Also, several literatures had questioned whether sentence uniformity leads to unfairness. For example, criminal justice scholars commonly believe that while sentencing guidelines increase uniformity, they do so at the cost of bias (Grunwald 2015; Ogletree 1988). They argue that uniformity can reduce the fairness of sentences by limiting the judges' ability to consider all relevant case characteristics, resulting in deviations from ideal outcomes.

In this section, I examine the political incentives driving the reform, aiming to clarify whether it served as (i) a monitoring tool to enhance the accuracy of punishments, (ii) an instrument for advancing specific policies or securing political support, or (iii) a means of manipulation to consolidate power and facilitate corruption by the military regime.

First, our findings suggest that the reform may have served as a monitoring tool to address the principal-agent problem within the courts, thereby enhancing fairness and efficiency in judicial decision-making. For instance, if verdicts prior to the reform were overly lenient, the introduction of monitoring could lead to stricter and more accurate judgments.

Second, the reform could also be viewed as a tool by the government to achieve specific policy objectives. These objectives might include policies to satisfy voters, similar to a democratic context. In the U.S., for instance, Democratic voters are generally expected to support the legalization of abortion, making it natural to expect government intervention to achieve this policy goal. In the context of Thailand, the "war on drugs" implemented by former Prime Minister Thaksin Shinawatra since 2003 (leader of a left-wing party) has been a significant policy that has gained him huge political support over the years. Other policy objectives might involve autocratic policy to stabilize social and political unrest, addressing riots, protests, crime control, or restricting freedom of speech.

Third, the reform could also be interpreted as a tool for an autocratic government to manipulate the courts and consolidate power. Anecdotal evidence in our case suggests that judicial strictness could either increase or decrease due to such manipulation. A notable example of harsher treatment, as revealed by the media, is a murder case in the southern border province of Yala. Initially, the provincial judge acquitted the defendant due to insufficient evidence, but the case was later reclassified as a national security offense (one of the reform treated case categories), with the Chief Regional Judge advocating for capital punishment.[35] On the other hand, there are cases involving high-profile politicians where the reform appears to result in more lenient judgments by the criminal court. These are national security offenses (one of the reform treated case category) that were either acquitted or seemingly treated with greater leniency by the criminal courts.[36]

Given these arguments, it is crucial to understand the drivers behind our findings: why does the reform generate uniformly harsher prison sentences? In the next subsection, I leverage the rich heterogeneity in the data to explore the potential incentives of the reform

[35]The judge later committed suicide in the courtroom in 2019, claiming that his verdict had been altered by the Chief Regional Judge. This bring media awareness to further investigate in the case.

[36]According to Bloomberg, former Thai Prime Minister Thaksin Shinawatra is facing indictment over charges linked to his time in office, including those potentially related to national security offenses. Despite the seriousness of these charges, the former Prime Minister receive lenient treatment or even acquittal by the criminal court in Bangkok (Bloomberg News, June 2024).

by analyzing the heterogeneous effects based on case characteristics and local politics.

6.1 Heterogeneity across Case Categories

In this section, I explore the differences across various case categories to understand which types of crime the government prioritizes for control and which receive more lenient treatment. Due to small treated sample if we were to consider each offense group separately, I first estimate equation (2) on prison sentence lengths allowing for differential treatment effects along five offense groups[37]: (i) offenses involving politicians and bureaucrats, (ii) offenses related to social/political stability, (iii) severe drug offenses, which include production/import/export/selling offenses, (iv) minor drugs offenses, which include posession/use of drugs, and (v) other offenses.

Table 7 displays estimated treatment effects from the estimation of equation (2) allowing for differential treatment effects along offense groups under various specifications. Column (1) shows results with quarter effects, quarter-of-the-year effects, offense effects, but otherwise no other controls. Column (2) adds a set of observable controls, column (3) includes all controls in equation (2), and column (4) adds judge fixed-effects. The results show that the reform significantly decreased prison sentence lengths by 160.9 percentage points (more than 100%) for cases involving politician and officer malfeasance in office, while it significantly increased prison sentence lengths by 87.1 percentage points for severe drug offenses. While more imprecisely estimate, the sign of the estimates are positive for social stability and politically sensitive cases, negative for minor drugs offenses and other treated case.

The results suggest that the government intervention in the court was with an aim to achieve more social control throgh higher punishment for severe drugs, social stability, and politically sensitive cases. However, cases involving politician and officer malfeasance in office are treated significantly more favorably by judges.

6.2 Heterogeneity across Political Alignments

In this section, I focus on political support at the constituency level. Using the 2019 general election in which an autocratic government ran for election, I measure local autocratic support from autocratic parties vote share at the constituency (c) level.[38] These electoral data are linked to the case records by subdistrict of the crime scene. The idea behind this heterogeneity analysis is that crimes occurring in areas with strong autocratic support might

[37] A full description of these offenses are described in Appendix B.3.

[38] There are two autocratic parties: Palang Pracharath Party, and Bhumjaithai. These are the two political parties whose leaders were previously part of the 2014 military junta.

influence judges to align more closely with autocratic-induced monitoring policies.

I estimate the heterogeneous effect of monitoring on judicial decision making across constituencies by interacting the term $Treated^1_{gr} \times Post_t$ with the measure of local political alignment. All continuous variables measuring political support are standardized to have a mean of zero and standard deviation equal to one. I also control for linear trends based on these local political characteristics. Thus, the interaction term here captures differential changes from the underlying trend of prison sentence as a function of predetermined constituency characteristics.

Table 8 shows the triple difference results from interacting the reform with each political alignment measure. Column (1) shows the baseline result with no interaction term. Column (2) reports the result when interacting $Treated^1_{gr} \times Post_t$ with the autocratic vote share. The effect of the monitoring reform is 21.6% larger if crimes were to occur in constituencies that are one standard deviation above the mean of autocratic support with respect to constituencies with mean autocratic support. This implies that the effect of the monitoring reform on sentence outcomes is stronger in areas that are more political aligned with an autocratic government.

7 Conclusion

It has long been claimed that governments attempt to increase social control through court-imposed punishments. This article provides the first causal evidence of the impact of a state intervention in the courts through monitoring. I find that severe and politically sensitive cases subjected to superior court monitoring received significantly longer prison sentences after the reform. These effects are primarily driven by previously lenient judges becoming more stringent post-reform. Additionally, judges' behavior varied depending on the offense category. While severe drug offenses and social stability cases received longer prison sentences, cases involving politicians and public officers' malfeasance received significantly more favorable treatment. The results are also driven by crimes in more autocratically aligned neighborhood.

Overall, my results indicate that the reform successfully achieved its intended consequence, increasing verdict uniformity and severity of court punishments for severe and politically sensitive cases that were subject to draft verdict revision before judgment delivery. However, this unequal application of justice across offense category raises several questions about the functions and fairness of the criminal justice system with respect to case categories.

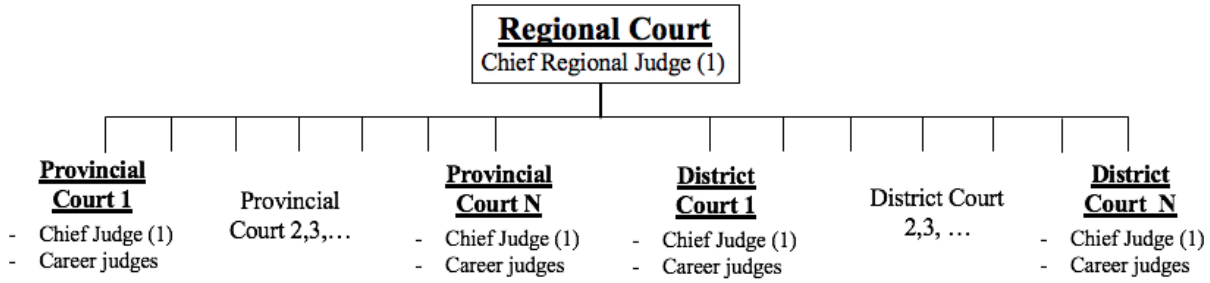
My findings may have broader implications for understanding judicial politics in other societies, particularly regarding the role of courts under authoritarian regimes. Judicial in-

terventions are prevalent in countries like South Africa, Israel, or Myanmar where democracy is not fully consolidated. The impacts of these interventions on social stability and the economy deserve further exploration. Just as autocratic interventions in the Thai courts were used to regain social control, similar actions may contribute to the weakening of democracy in other regions.

Even in established western democracies where there is the formal separation of powers, the politicisation of judicial appointments can undermine judicial independence (Mehmood 2022). Additionally, with the rise of populist parties in Europe over the last two decades, judicial independence has been severely reduced in countries such as Hungary, Poland, and Turkey. Despite this trend, the literature on the economics of crime and political economy has generally overlooked the threats to judicial independence and the role of judicial institutions in shaping the rule of law. This gap in research highlights the need for a deeper understanding of how judicial actions influence political and economic stability globally.

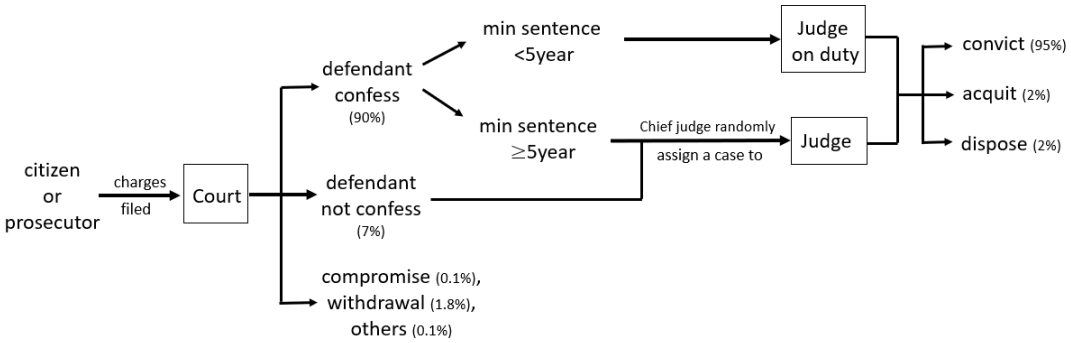
Figures

Figure 1: Structure of each Regional Court



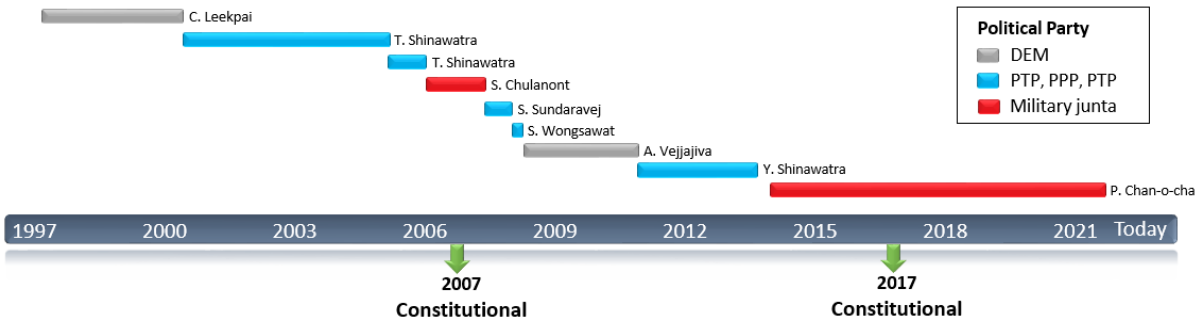
Note: There are 9 regional courts in Thailand. Each regional court has a similar structure as displayed above: (i) oversees a number of provincial courts and district courts (ii) presided by one Chief Regional judge and one Vice Chief Regional Judge where both are appointed by the Office of the Judicial Comission (JC). Other judges in provincial courts and district courts are recruited from a merit-based system. Our sample contains 7 provincial courts, overseen by 3 regional courts.

Figure 2: Flow chart of criminal cases in Thailand.



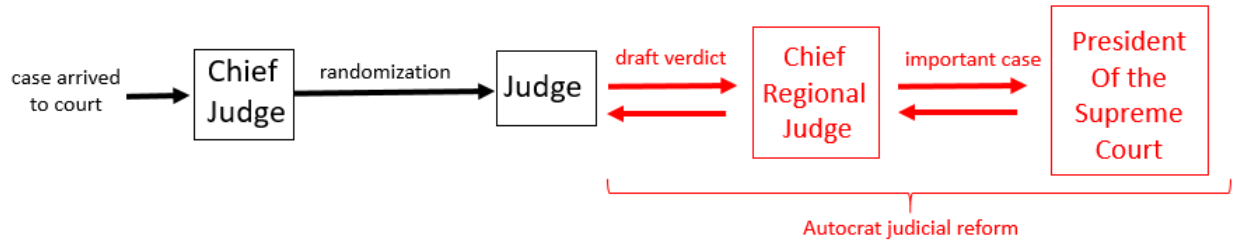
Note: The process is summarized from the Thai Criminal Procedure Code. Source: Author's calculation.

Figure 3: Timeline of the thai political turnover (1997-2022).



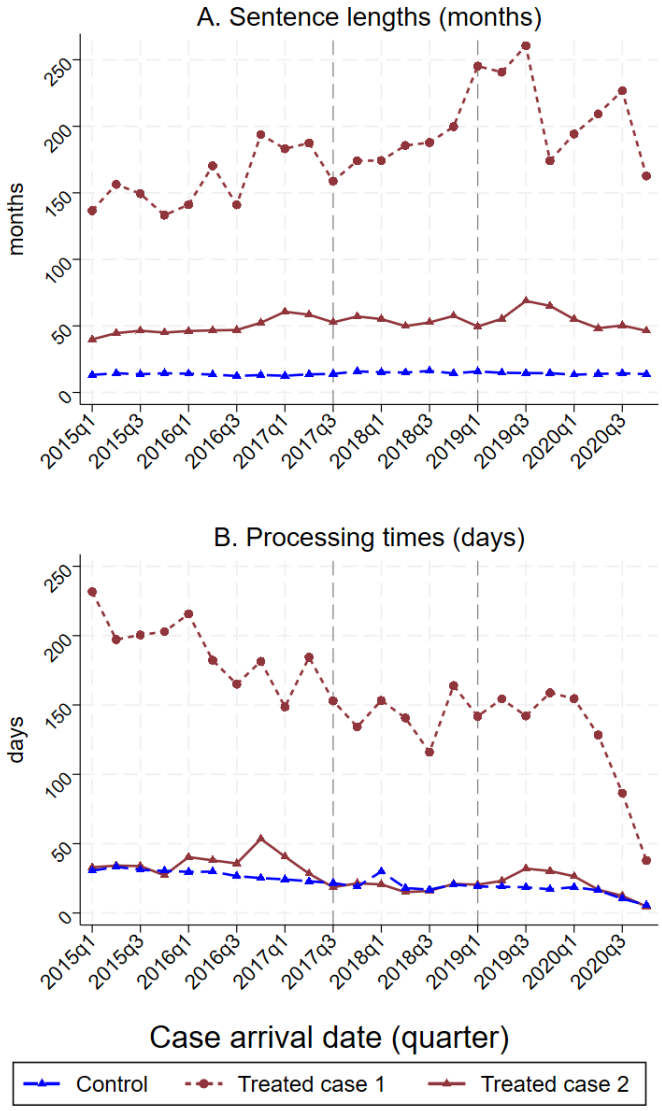
Note: This includes autocratic periods (red), and democratic periods (Democrat party, grey; Thaksin-supported party, blue) periods.

Figure 4: Case assignment process before vs. after the 2017 reform.



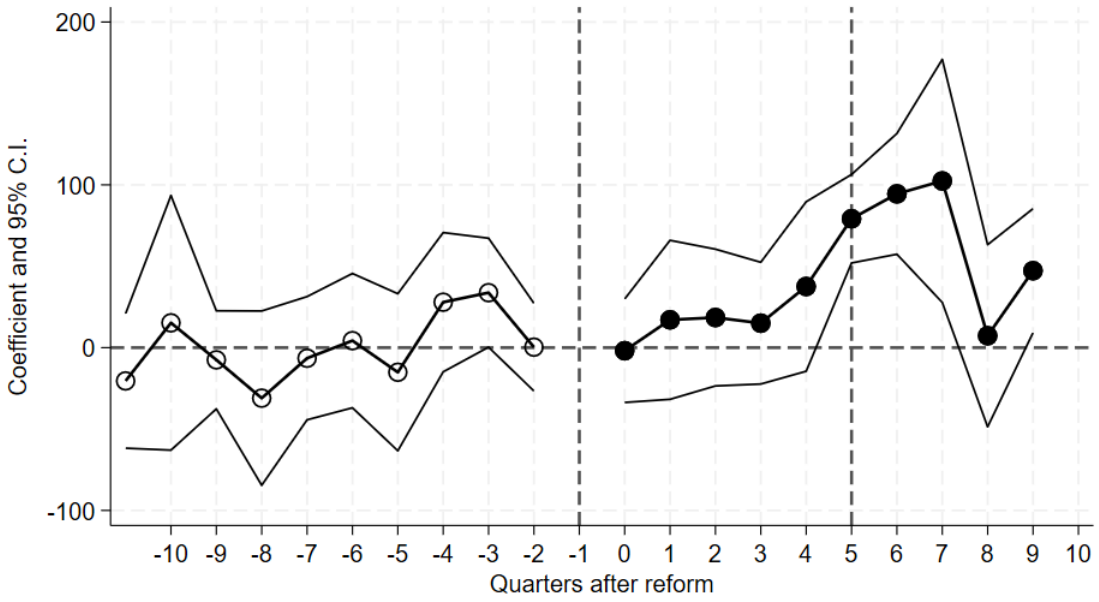
Note: Pre-reform case assignment process is described in black. The red part of the figure is added by the 2017 judicial reform.

Figure 5: Graphical representation.



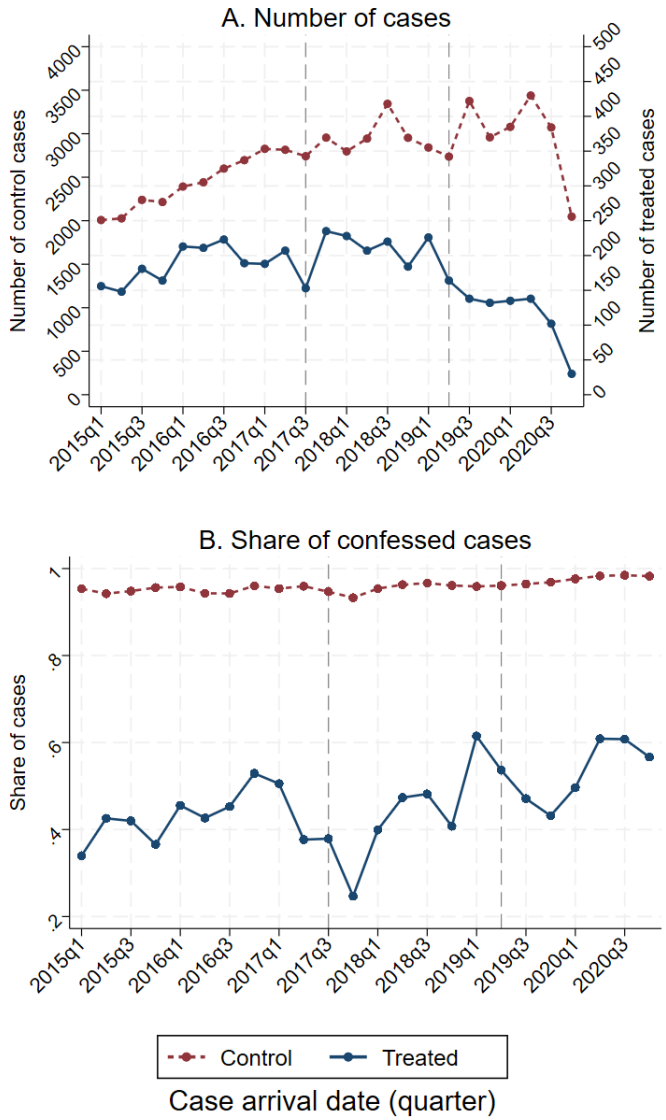
Note: Vertical line represents days of regulation enforcement for special case: Aug 26th, 2017 (2017Q3). Regulation is enforced for very special case on: Feb 21th, 2019 (2019Q1). The horizontal axis represents time that case arrives to the court. The last figure shows harsh punishment rate, conditional on similar sentence range. *Treated case 1* are cases that draft verdict must be reported to the Chief Regional Judge before judgment delivery. *Treated case 2* are cases that only final verdict must be reported to the Chief Regional Judge after judgment delivery.

Figure 6: Event study



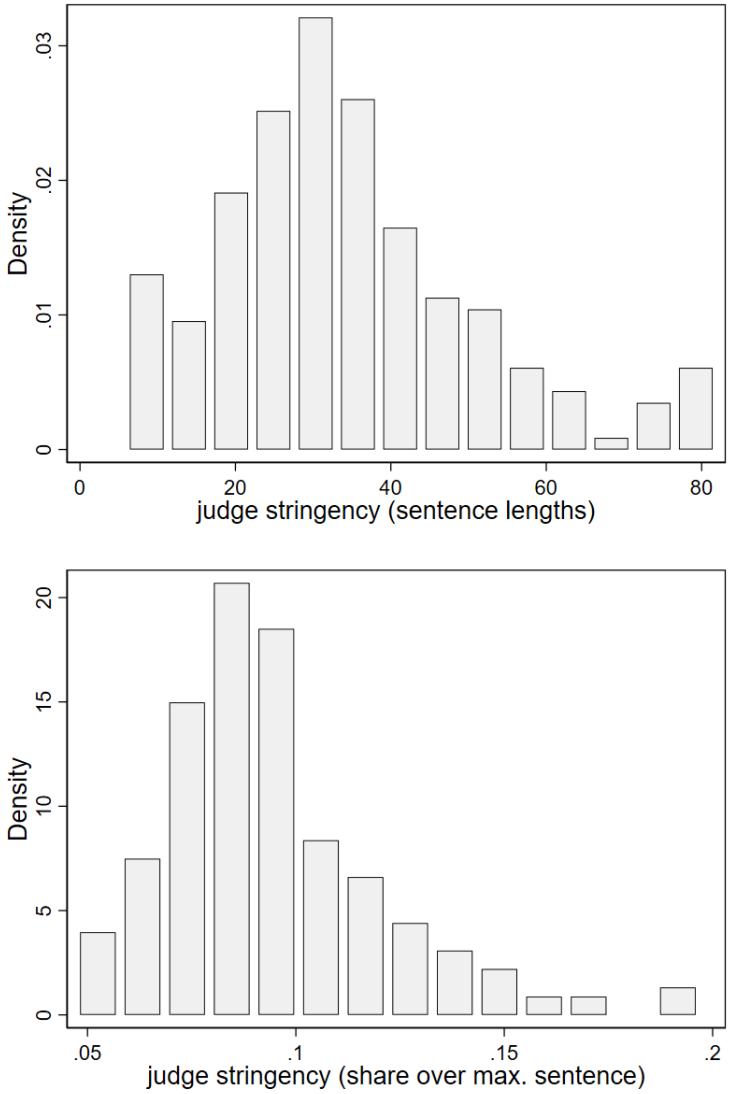
Note: The figure shows DD coefficients and 95% confidence intervals from estimating equation (2) on prison sentence lengths (mean = 36). Standard errors are clustered at *gr* level. Treatment is defined as cases that are subjected to superior court revision prior to judgment delivery. Control cases are those that do not subjected to superior court monitoring. Sample includes cases with maximum punishment range greater than or equal to 10 years of prison sentence. The dashed vertical line is the reference quarter or quarter of the reform itself, 2017Q3 (the reform is implemented on August 26, 2017 for region 3/4; August 28,2017 for region 1), and the year of extra monitoring (2019Q1).

Figure 7: Change in criminal behavior and the reforms.



Note: Treated case here are cases that draft verdict must be reported to the Chief Regional Judge before judgment delivery. Observation=number of case.

Figure 8: Judge stringency distribution (judge-level)



Note: This figure shows the distribution of "judge-level stringency score", which is aggregated from "case-level judge stringency score" (calculated according to equation 3) for each juge. The baseline "pre-reformed sample" is used for estimation. Distribution of "case-level judge stringency score" before aggregation is reported in Figure 15 in Appendix D.

Tables

Table 1: Definition of Treated Cases after 2017 Judicial Reform

Treated case category	Treated case 1: monitor pre-judgment		Treated case 2: monitor post-judgment	
	N	Avg. max. sentence rage (year)	N	Avg. max. sentence length (year)
A. By charge severity:				
1. Cases with max. prison				
>20 years for courts in region 1	2350	88.12	10595	61.73
>10 years for courts in region 3 & 4	2212	59.94	9722	46.44
B. By offense:				
2. National security/international relations/terrorism	44	42.00	636	18.33
3. Drugs related cases				
Offense=produce,import,export	24	86.67	143	69.79
Offense=sell				
>50 g. of pure substance for courts in region 1	563	98.47	310	92.16
>20 g. of pure substance for courts in region 3 & 4	563	88.55	372	92.18
Total drugs offense	1137	93.73	768	88.33
4. Cases involve top bureaucrats/politicians	113	6.89	0	.
5. Organized crimes, environment crimes, human trafficking, fishery act, etc.	342	26.17	0	.
Total	5052	73.76	20930	53.81

Note: N represents defendant-case number of observations. Life imprisonment is set to 60 years (720 months) of prison sentence. Capital punishment is set to 100 years (1200 months) of prison sentence. A single treated case can be in more than one treated case category.

Table 2: Summary Statistics

	Full sample	Control cases	Treated cases		
	(1)	(2)	(3) all treated	(4) treated 1	(5) treated 2
A. Defendants:					
Average age	32.232	32.197	32.296	34.025	31.876
Male (0/1)	0.916	0.932	0.889	0.877	0.892
Thai (0/1)	0.996	0.997	0.995	0.985	0.997
Known criminal history (0/1)	0.067	0.069	0.063	0.045	0.067
Number of charges per defendant	1.571	1.670	1.395	1.546	1.358
Number of defendants	70338	45205	26406	5231	21256
B. Judges & Prosecutors:					
Number of primary judges	346	341	330	297	326
Number of prosecutor	480	174	322	106	234
C. Case characteristics:					
Average processing time (days)	33.097	22.301	52.233	163.740	25.096
Defendant confess (0/1)	0.917	0.946	0.866	0.410	0.976
Defendant not confess (0/1)	0.062	0.034	0.113	0.578	0.000
D. Local characteristics:					
Number of crime scene (district)	85	85	82	81	81
Number of crime scene (constituency)	10	10	10	9	10
LeftWin2011	0.658	0.630	0.708	0.648	0.723
RightWin2011	0.034	0.045	0.014	0.032	0.010
AutocratWin2019	0.336	0.347	0.316	0.360	0.305
E. Offense category:					
Violent	0.067	0.075	0.054	0.215	0.015
Property	0.035	0.038	0.031	0.015	0.035
Fraud	0.003	0.003	0.003	0.013	0.000
Sex	0.012	0.010	0.017	0.054	0.008
Drugs(major)	0.295	0.103	0.636	0.432	0.686
Drugs(minor)	0.487	0.682	0.142	0.066	0.160
Politician/bureaucrat involved	0.005	0.004	0.007	0.029	0.002
Political stability	0.024	0.001	0.065	0.063	0.065
Others	0.070	0.084	0.045	0.115	0.028
F. Verdicts:					
Convict, def. confess (0/1)	1.000	1.000	1.000	1.000	1.000
Convict, def. not confess (0/1)	0.911	0.900	0.917	0.917	.
G. Sentence:					
Prison (0/1)	0.562	0.387	0.872	0.808	0.887
Avg. prison sentence (months) , cond. on incarceration	57.406	22.670	84.533	212.825	56.462
Life imprisonment (0/1)	0.003	0.000	0.007	0.035	0.001
Capital punishment (0/1)	0.000	0.000	0.001	0.003	0.001
Observations	74678	47734	26929	5271	21658

Note: There are 73,886 case-by-defendant. Treated1 are cases that draft verdict must be reported to the Chief Regional Judge before judgment delivery. Treated2 are cases that only final verdict must be reported to the Chief Regional Judge after judgment delivery.

Table 3: Covariate balance

	Control cases				Treated cases				DID	
Case characteristics	Before	After	β_{post}	*p-value*	Before	After	β_{post}	*p-value*	β_{post}	*p-value*
Sample Size	25643	43191			2294	2758				
A. Defendant characteristics (0/1)										
Average age	31.5	32.4	-0.300	[0.068]	33.7	33.7	-0.311	[0.803]	-0.423	[0.625]
Male (0/1)	0.913	0.926	-0.020	[0.002]	0.881	0.894	-0.031	[0.093]	0.001	[0.959]
Thai (0/1)	0.997	0.997	0.000	[0.930]	0.989	0.982	0.004	[0.687]	-0.003	[0.578]
Criminal history (0/1)	0.045	0.082	-0.004	[0.401]	0.028	0.056	-0.014	[0.362]	-0.008	[0.321]
Charges per defendant	1.59	1.56	0.112	[0.008]	1.56	1.55	0.141	[0.004]	0.007	[0.857]
District poverty index	0.576	0.524	-0.019	[0.214]	0.551	0.596	-0.066	[0.159]	0.020	[0.582]
B. Offence group (0/1)										
Fraud offense	0.003	0.005	0.008	[0.041]	0.007	0.023	0.010	[0.072]	-0.000	[0.971]
Sex offense	0.016	0.009	-0.002	[0.592]	0.096	0.053	-0.008	[0.620]	-0.036	[0.181]
Violent offense	0.075	0.050	-0.000	[0.953]	0.299	0.186	0.019	[0.576]	-0.082	[0.088]
Property offense	0.054	0.043	0.002	[0.870]	0.028	0.029	0.001	[0.940]	0.018	[0.256]
Drugs offense	0.774	0.850	0.016	[0.273]	0.444	0.615	-0.006	[0.872]	0.092	[0.032]
Firearms	0.110	0.096	0.022	[0.035]	0.177	0.139	0.028	[0.289]	-0.025	[0.479]
Traffic	0.317	0.271	0.050	[0.036]	0.046	0.070	0.031	[0.296]	0.031	[0.222]
Malfeasance in office	0.005	0.003	-0.000	[0.338]	0.041	0.014	-0.009	[0.243]	-0.005	[0.309]
Political sensitive	0.005	0.003	-0.002	[0.109]	0.031	0.017	-0.009	[0.593]	-0.005	[0.628]

Note: This table reports results on predetermined case characteristics. Each row in each panel is a separate regression. Column (3) and (4) report the results of regressing these characteristics on a dummy that indicates whether a case arrives after the reform or not for the control cases, whereas Column (7) and (8) report those for the treated cases. All regressions control for quarter, court, court trend, and judge fixed effects. Standard errors are clustered at charge severity-court bin level

Table 4: Estimate of the Baseline equations

	Sentence lengths(months)					
	(1)	(2)	(3)	(4)	(5)	(6)
$Treated^1_{gr}$	121.387***	129.920***	130.067***	128.927***	127.890***	131.155***
	(18.515)	(22.203)	(21.858)	(21.949)	(21.737)	(21.406)
$Treated^1_{gr} \times Post_t$	39.172***	37.545***	37.366***	38.566***		39.135***
	(7.597)	(7.417)	(7.540)	(7.360)		(7.522)
$Treated^1_{gr} \times Post^{2017}_t$					14.660*	
					(8.801)	
$Treated^1_{gr} \times Post^{2019}_t$					69.555***	
					(9.817)	
$Treated^2_{gr}$						1.477
						(9.526)
$Treated^2_{gr} \times Post_t$						2.360
						(2.945)
Observations	72309	72190	72190	72190	72190	72190
R-square	0.382	0.388	0.389	0.395	0.399	0.395
Mean DV.	36	36	36	36	36	36
Offense fe.	Yes	Yes	Yes	Yes	Yes	Yes
Severity bin fe.	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly fe.	Yes	Yes	Yes	Yes	Yes	Yes
Region fe.	No	No	Yes	Yes	Yes	Yes
Region trend	No	No	Yes	Yes	Yes	Yes
Judge fe.	No	No	No	Yes	Yes	Yes
Control vars.	No	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable = prison sentence length. Life imprisonment is set to 60 years (720 months) of prison sentence. Capital punishment is set to 100 years (1200 months) of prison sentence. $Treated^1_{gr}$= Monitor pre-judgment. $Treated^2_{gr}$= Monitor post-judgment. Standard errors are cluster at the gr level (121 clusters)

Table 5: Robustness Table

	Sentence lengths (months)				
	(1)	(2)	(3)	(4)	(5)
	Baseline	Court trend	Add offense trend	Exclude capital punishment	Exclude capital punishment & life imprisonment
Panel A					
$Treated^1_{gr}$	128.927***	129.005***	129.558***	127.240***	109.412***
	(21.949)	(22.031)	(21.957)	(22.628)	(23.596)
$Treated^1_{gr} \times Post_t$	38.566***	38.481***	36.624***	36.016***	33.005***
	(7.360)	(7.278)	(7.314)	(7.543)	(7.197)
Panel B					
$Treated^1_{gr}$	108.4***	108.5***	108.7***	108.7***	108.7***
	(4.63)	(4.62)	(4.62)	(4.62)	(4.62)
$Treated^1_{gr} \times Post^{2017}_t$	19.21**	19.10**	18.32**	18.32**	18.32**
	(2.60)	(2.60)	(2.50)	(2.50)	(2.50)
$Treated^1_{gr} \times Post^{2019}_t$	54.40***	54.32***	53.06***	53.06***	53.06***
	(6.80)	(6.86)	(6.72)	(6.72)	(6.72)
Panel C					
$Treated^1_{gr}$	109.242***	109.338***	110.195***	110.195***	110.195***
	(22.001)	(22.057)	(22.541)	(22.541)	(22.541)
$Treated^2_{gr}$	-1.160	-1.018	-0.510	-0.510	-0.510
	(8.485)	(8.548)	(9.199)	(9.199)	(9.199)
$Treated^1_{gr} \times Post_t$	35.013***	34.841***	33.704***	33.704***	33.704***
	(7.287)	(7.239)	(7.184)	(7.184)	(7.184)
$Treated^2_{gr} \times Post_t$	2.956	2.653	2.824	2.824	2.824
	(2.610)	(2.575)	(2.384)	(2.384)	(2.384)
r2	0.44	0.44	0.44	0.44	0.44
N	71976.00	71976.00	71976.00	71976.00	71976.00

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Baseline results are presented in Column (1) and estimates with full controls. Column (2) replaces court region FE and court region trend by court FE and court trend. Column (3) adds offense group FE to the baseline estimate. Column (4) excludes capital punishment (1200 months prison sentence). Column (5) excludes capital punishment (1200 months prison sentence) and life imprisonment (720 months prison sentence).

Table 6: Heterogeneity by Judge Stringency

	Measure of Judge Stringency: Sentence lengths (months)		Measure of Judge Stringency: Share over max. sentence	
	(1) Lenient judges	(2) Strict judges	(3) Lenient judges	(4) Strict judges
Panel A: All regions				
$Treated^1_{gr}$	95.312***	145.249***	93.608***	157.499***
	(18.168)	(22.644)	(17.921)	(21.391)
$Treated^1_{gr} \times Post_t$	79.034***	10.185	46.708***	8.546
	(20.022)	(8.822)	(14.988)	(8.472)
r2	0.392	0.392	0.358	0.412
ymean	26	39	29	39
N	17893	32905	21165	29633
Panel B: By region				
$Treated^1_{gr}$	96.238***	145.093***	93.409***	158.108***
	(18.049)	(22.138)	(17.489)	(21.109)
$Treated^1_{gr} \times Post_t \times Reg1$	83.083***	26.983*	96.959***	13.675
	(15.144)	(15.639)	(24.173)	(10.679)
$Treated^1_{gr} \times Post_t \times Reg3$	119.805***	-76.617***	63.573	85.036*
	(38.807)	(19.787)	(48.042)	(46.028)
$Treated^1_{gr} \times Post_t \times Reg4$	31.807	-13.819	18.936	-15.779
	(33.419)	(26.246)	(26.001)	(27.762)
r2	0.397	0.395	0.363	0.414
ymean	26	39	29	39
N	17893	32905	21165	29633

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable = prison sentence length. Life imprisonment is set to 60 years (720 months) of prison sentence. Capital punishment is set to 100 years (1200 months) of prison sentence. Each column is a subsample regression. Lenient judges are judges with stringency score below the median. Strict judgese are judges with stingency score greater than or equal to the median. All columns control for observables and unobservables as described in equation (2).

Table 7: Heterogeneity by Case category

	Sentence lengths(months)			
	(1)	(2)	(3)	(4)
$Treated^1_{gr}$	116.972***	124.761***	124.959***	123.676***
	(19.455)	(23.045)	(22.707)	(22.866)
$Treated^1_{gr} \times Post_t \times Officer_o$	-167.606***	-166.114***	-167.461***	-168.623***
	(27.553)	(30.306)	(30.186)	(30.619)
$Treated^1_{gr} \times Post_t \times Stability_o$	41.750	38.845	39.250	39.502
	(29.998)	(29.522)	(29.501)	(29.505)
$Treated^1_{gr} \times Post_t \times SeriousDrugs_o$	84.684***	82.669***	82.098***	83.429***
	(22.297)	(22.570)	(21.914)	(22.003)
$Treated^1_{gr} \times Post_t \times MinorDrugs_o$	-13.141	-14.611	-16.300	-16.059
	(56.864)	(55.669)	(56.498)	(55.093)
$Treated^1_{gr} \times Post_t \times Others_o$	-26.781	-27.499	-26.832	-25.606
	(20.911)	(20.481)	(20.895)	(20.816)
Observations	72309	72190	72190	72190
R-square	0.397	0.402	0.404	0.409
Mean DV.	36	36	36	36
Offense fe.	Yes	Yes	Yes	Yes
Severity bin fe.	Yes	Yes	Yes	Yes
Quarterly fe.	Yes	Yes	Yes	Yes
Region fe.	No	No	Yes	Yes
Region trend	No	No	Yes	Yes
Judge fe.	No	No	No	Yes
Control vars.	No	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable = prison sentence length. Life imprisonment is set to 60 years (720 months) of prison sentence. Capital punishment is set to 100 years (1200 months) of prison sentence. Each column is a subsample regression. All columns control for observables and unobservables as described in equation (2).

Table 8: Heterogeneity by political support

	Sentence lengths(months)	
	(1)	(2)
$Treated^1_{gr}$	128.927***	129.144***
	(21.949)	(21.596)
$Treated^1_{gr} \times Post_t$	38.566***	37.537***
	(7.360)	(8.095)
$Treated^1_{gr} \times Post_t \times$ AutocratVote$_c$		21.606**
		(10.515)
Observations	72190	72078
R-square	0.395	0.397
Mean DV.	36	36
Offense fe.	Yes	Yes
Severity bin fe.	Yes	Yes
Quarterly fe.	Yes	Yes
Region fe.	Yes	Yes
Region trend	Yes	Yes
Judge fe.	Yes	Yes
Control vars.	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable = prison sentence length. Life imprisonment is set to 60 years (720 months) of prison sentence. Capital punishment is set to 100 years (1200 months) of prison sentence. The table shows the results of OLS estimation of equation (3), augmented with an interaction term between $Treated^1_{gr} \times Post_t$ and $AutocratVote_c$, which capture local autocratic alignment (measured by 2019 general election vote share of autocratic parties). Continuous variables measuring constituency characteristics are normalized to have a mean of zero and standard deviation equal to one. Standard errors are clustered at the gr level (121 clusters)

References

Adena, Maja et al. (2015). "Radio and the Rise of The Nazis in Prewar Germany". *The Quarterly Journal of Economics* 130.4, pp. 1885–1939.

Aizer, Anna and Joseph J. Doyle Jr. (2015). "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges". *The Quarterly Journal of Economics* 130.2, pp. 759–803.

Alesina, Alberto and Eliana La Ferrara (2014). "A Test of Racial Bias in Capital Sentencing". *American Economic Review* 104.11, pp. 3397–3433.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson (2019). "Politics in the Courtroom: Political Ideology and Jury Decision Making". *Journal of the European Economic Association* 17.3, pp. 834–875.

Ash, Elliott et al. (2021). "In-Group Bias in the Indian Judiciary: Evidence from 5.5 Million Criminal Cases". In: pp. 47–47.

Axbard, Sebastian (2024). "Convicting Corrupt Officials: Evidence from Randomly Assigned Cases". *Review of Economic Studies*.

Axbard, Sebastian and Zichen Deng (2024). "Informed Enforcement: Lessons from Pollution Monitoring in China". *American Economic Journal: Applied Economics* 16.1, pp. 213–252.

Barkow, Rachel E. (2009). "Institutional Design and the Policing of Prosecutors: Lessons from Administrative Law". *Stanford Law Review* 61.4, pp. 869–921.

Becker, Gary S. and George J. Stigler (1974). "Law Enforcement, Malfeasance, and Compensation of Enforcers". *The Journal of Legal Studies* 3.1, pp. 1–18.

Besley, Timothy and Stephen Coate (2003). "Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach". *Journal of Public Economics* 87.12, pp. 2611–2637.

Bhuller, Manudeep et al. (2020). "Incarceration, Recidivism, and Employment". *Journal of Political Economy* 128.4, pp. 1269–1324.

Bibas, Stephanos (2010). "The Need for Prosecutorial Discretion". *Temple Political & Civil Rights Law Review*.

Bjerk, David (2005). "Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion under Mandatory Minimum Sentencing". *The Journal of Law and Economics* 48.2, pp. 591–625.

Borcan, Oana, Mikael Lindahl, and Andreea Mitrut (2017). "Fighting Corruption in Education: What Works and Who Benefits?" *American Economic Journal: Economic Policy* 9.1, pp. 180–209.

Di Tella, Rafael and Ernesto Schargrodsky (2003). "The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires". *The Journal of Law & Economics* 46.1, pp. 269–292.

Dobbie, Will and Jae Song (2015). "Debt Relief and Debtor Outcomes: Measuring the Effects of Consumer Bankruptcy Protection". *American Economic Review* 105.3, pp. 1272–1311.

Downey, Mitch and Ben Grunwald (2023). "The Costs of Top-Down Control: Discretion and Turnover of Federal Prosecutors". *Working Paper*.

Doyle, Josep (2007). "Child Protection and Child Outcomes: Measuring the Effects of Foster Care". *American Economic Review* 97.5, pp. 1583–1610.

Ernest, Liu, Wenwei Peng Yi Lu, and Shaoda Wang (2022). "Judicial Independence, Local Protectionism, and Economic Integration: Evidence from China". *NBER Working Paper* 30432.

Ferraz, Claudio and Frederico Finan (2008). "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes". *The Quarterly Journal of Economics* 123.2, pp. 703–745.

— (2011). "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments". *American Economic Review* 101.4, pp. 1274–1311.

French, Eric and Jae Song (2014). "The Effect of Disability Insurance Receipt on Labor Supply". *American Economic Journal: Economic Policy* 6.2, pp. 291–337.

Gehlbach, Scott et al. (2024). "Is There Really a Dictator's Dilemma? Information and Repression in Autocracy". *SSRN Electronic Journal*.

Ginzburg, Tom and Tamir Moustafa (2008). *Rule by Law: The Politics of Courts in Authoritarian Regimes*. Cambridge University Press.

Grunwald, Ben (2015). "Questioning Blackmun's Thesis: Does Uniformity in Sentencing Entail Unfairness? - ProQuest". *Law & society review* 49, pp. 499–534.

Jia, Ruixue, Masayuki Kudamatsu, and David Seim (2015). "Political Selection in China: The Complementary Roles of Connections and Performance". *Journal of the European Economic Association* 13.4, pp. 631–668.

King, Gary, Jennifer Pan, and Margaret E. Roberts (2014). "Reverse-Engineering Censorship in China: Randomized Experimentation and Participant Observation". *Science* 345.6199, p. 1251722.

Kling, Jeffrey R. (2006). "Incarceration Length, Employment, and Earnings". *American Economic Review* 96.3, pp. 863–876.

La Porta, R. et al. (2004). "Judicial Checks and Balances". *Journal of Political Economy* 112.2, pp. 445–470.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand (2013). "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt". *American Economic Review* 103.5, pp. 1797–1829.

Mahakun, Wicha (1977). "Phupipaksa Kub Karn Chai Dullapinij Naikarn Kumnod Tote (Judges and the Exercising of Sentencing Discretion)". *Dullapah*, pp. 54–69.

Martinez-Bravo, Monica et al. (2022). "The Rise and Fall of Local Elections in China". *American Economic Review* 112.9, pp. 2921–2958.

Mehmood, Sultan (2022). "The Impact of Presidential Appointment of Judges: Montesquieu or the Federalists?" *American Economic Journal: Applied Economics* 14.4, pp. 411–445.

Mullen, Brian, Rupert Brown, and Colleen Smith (1992). "Ingroup Bias as a Function of Salience, Relevance, and Status: An Integration". *European Journal of Social Psychology* 22.2, pp. 103–122.

Ogletree, Charles J. (1988). "The Death of Discretion? Reflections on the Federal Sentencing Guidelines". *Harvard Law Review* 101.8, pp. 1938–1960.

Poblete-Cazenave, Rubén (2023). "Do Politicians in Power Receive Special Treatment in Courts? Evidence from India". *American Journal of Political Science* 00, pp. 1–18.

Rozenas, Arturas and Denis Stukal (2019). "How Autocrats Manipulate Economic News: Evidence from Russia's State-Controlled Television". *The Journal of Politics* 81.3, pp. 982–996.

Saengwirotjanapat, Surasit (2016). "Karn Chai Dullapinij Nai Karn Kumnod Tote (Exercising Sentencing Discretion)". *Research studies funded by the Office of the Court of Justice.*

Shayo, Moses and Asaf Zussman (2011). "Judicial Ingroup Bias in the Shadow of Terrorism". *The Quarterly Journal of Economics* 126.3, pp. 1447–1484.

Simonov, Andrey and Justin Rao (2022). "Demand for Online News under Government Control: Evidence from Russia". *Journal of Political Economy* 130.2, pp. 259–309.

Vannutelli, Silvia (2023). "From Lapdogs to Watchdogs: Random Auditor Assignment and Municipal Fiscal Performance". *NBER WP* 30644.

Wen, Jaya Y (2024). "State Employment as a Strategy of Autocratic Control in China". *Review of Economics and Statistics (forthcoming).*

Xu, Guo (2018). "The Costs of Patronage: Evidence from the British Empire". *American Economic Review* 108.11, pp. 3170–3198.

Yampracha, Supakit (2016). "Understanding Thai Sentencing Culture". PhD thesis. University of Strathclyde.

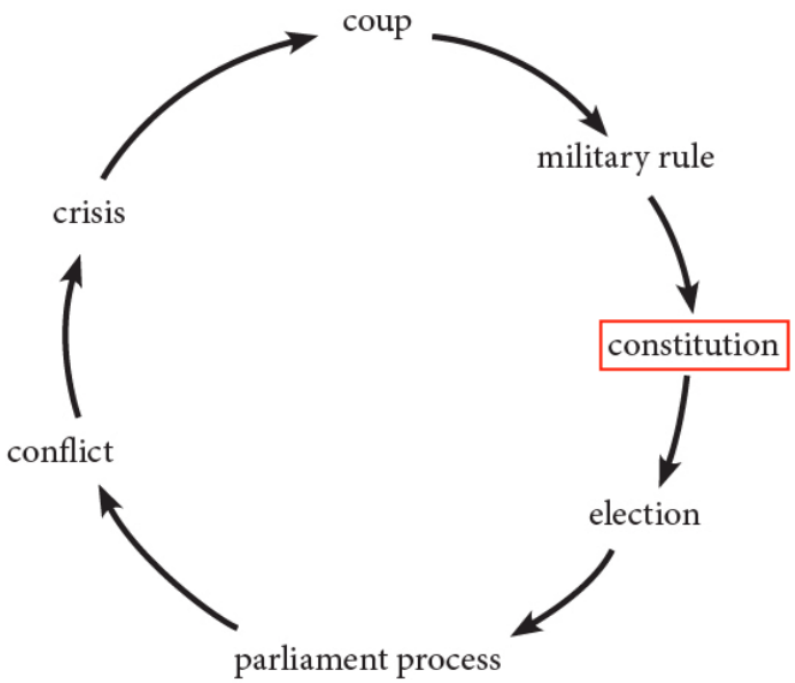
Yanagizawa-Drott, David (2014). "Propaganda and Conflict: Evidence from the Rwandan Genocide". *The Quarterly Journal of Economics* 129.4, pp. 1947–1994.

Appendix

A Background

A.1 Political Cycle

Figure 9: Coup political cycle



A.2 Treated Case Definition

The implementation of the 2017 judicial reform is stated according to the "Regulations on Reporting Important Cases in the Courts of First Instance and Appeallate Courts to the President of the Supreme Court, and reporting cases and examining case files in the Office of the Chief Justice of the Region B.E. 2560 (2017)".[39] Documents are publicly available on the Court of Justice website.

Article 7 stipulates that cases that judges must report draft verdict to the Chief Regional Judge before judgment delivery include:

1) National security offense (Criminal Code, Sections 107-135; including Sections 112 and 116).

2) Terrorism offense (Criminal Code, Sections 135/1-135/4, or other laws).

3) Severe cases that have a maximum prison sentence range over 10 years, life imprisonment, or death sentence.

4) Drug-related offenses are subjected superior court review if

4.1) offenses are related to posessing for sell, or sell:

- Narcotic drugs Category 1: greater than 1,000 tablets/80 g./20 g. of pure substance, either one.

- Narcotic drugs Category 2 (specifically opium, morphine, cocaine): with a calculated quantity of 100 g. or more.

- Narcotic drugs Category 5 (specifically marijuana or hemp): with a dry weight of 10 kg. or more, or a quantity of fresh marijuana or hemp plants weighing 30 kg. or more.

4.2) offenses are related to manufacturing, importing/exporting:

- Narcotic drugs Category 1 [40] 5) Contempt of court case.

6) Autopsy investigation case.

7) Cases involving famous individual, e.g., Prime Minister, Ministers, MPs, Senators, people holding positions in independent organizations according to the constitution, judges and also extends to the case of civil servants from the level of Director-General and above. Soldiers or police officers ranked General, or higher. Director of State Enterprises Personswho were granted diplomatic immunity.

8) Cases whose circumstances are of interest to the public

9) Cases that may affect international relations

10) Cases related to security in the southern border provinces etc.[41]

[39] Published on: https://opsc.coj.go.th/th/content/category/detail/id/8/cid/1142/iid/132412

[40] except it is done by dividing or packaging according to the Narcotic Act of B.E. 2522, Section 65, paragraphs three and four.

[41] The reform also include civil cases, not covered by our data sample. These include:

1) Civil cases involving property disputes of 5 million baht or more, case that plaintiff is a financial institution, cases that involve more than 10 million bath or more of disputed assets.

2) Cases involved real estate possession rights that have objections and funds greater than of

The 2017 Act also open rooms for regional adjustment for ease of implementation. In our court sample, the Regional Court 1 issued additional adjustments to the 2017 judicial reform through the "Regulations of the Office of the Chief Justice Region 1 regarding reporting and examining case files" to be implemented concurrently with the 2017 judicial reform. Regional Court 3 and 4 do not make adjustment to the national Act. These adjustments for Regional Court 1 are summarized as follows:

3.) Criminal cases are subjected to reported if it has a maximum prison sentence range over 20 years, instead of 10 years as stated in the 2017 Act on Judicial Service of the Courts of Justice.

4) Drugs related cases are subjected to reported if

4.1) offenses are related to posessing for sell, or sell:

- Narcotic drugs Category 1: greater than 50 g. of pure substance.

- Narcotic drugs Category 5 (specifically marijuana or hemp plants): with a dry weight greater than or equal to 50 kg

In 2019, an extra monitoring of special cases are implemented according to the "Judicial Regulations of the Court of Justice on Reporting Important Cases in the Court of First Instance and the Court of Appeal to the President of the Supreme Court and Reporting Cases and Examination of Case Files in the Office of the Chief Justice of the Region B.E. 2562 (2019)".[42] Article 7 which stipulate superior court revision for "special case" remained, with added Article 5 for "very special cases". Article 5 is described as follow,

Article 5 prescribes that the President of the Court of Appeal and the Chief Justice of the Court of First Instance Chief Justice of various courts including the Chief Justice of the Region Has a duty to report to the President of the Supreme Court on "very important cases" Important cases here refer to:

1.) National security offense (Criminal Code, Sections 107-135; including Sections 112 and 116).

2.) Terrorism offense (Criminal Code, Sections 135/1-135/4, or other laws), human trafficking offenses, or the Fisheries Act that are of public interest.

3.) Security-related cases in the southern border provinces.

4.) Cases involving famous individual, e.g., Prime Minister, Ministers, MPs, Senators, people holding positions in independent organizations according to the constitution. or judges.

5.) Civil cases filed with the state and disputed assets of 1,000 million baht or more or civil cases with disputed assets of 5,000 million baht or more.

6.) Cases that may affect international relations, etc.

equal to 200,000 baht.

[42]Published on: https://opsc.coj.go.th/th/content/category/detail/id/8/cid/1142/iid/132413

B Data

B.1 Data Sources

The data used in this paper comes from a variety of sources. Here, I describe each data sources, its coverage, and characteristics.

Complaint and Judgment texts: This is the main data sources used in this study. The data is linked with case records through annonymized case ID. The case records include anonymized defendant and case characteristics. From the complaint and judgment texts, I extracted information on charges, suboffenses, prison sentence lengths (months), seized drugs (size, tablets, and pure substance), and the subdistrict of the crime scene. The data is provided by the Thai Court of Justice.

Criminal Code of Coduct, Acts, and Royal Decrees: The data is available from the website of Office of the Council of the State. For each law/Acts/Decree, I extract charge-level punishment range. These data are used to linked with case records at the charge-level and obtain maximum punishment range, which is the main variable for charge severity and treatment determination.

Election data: I collect 2011 and 2019 general election result from the Election Commission of Thailand.[43] These data contain votes of each candidate by constituency. I link this data to case records at the sub-district level.

Others: district-level poverty index is downloaded from the National Statistical Office of Thailand website.

B.2 Code Treated Case

The number of treated cases are reported in Table 1. Here, I described how each treated cases are coded from complaint and judgment text.

Code Prison Sentence Range.–As described in the previous section, treated cases are cases with maximum prison sentence range over 10 years, life imprisonment, or death sentence. A key challenge here is to identified case-level "prison sentence range", which is not directly observed in the data. I identified case-level prison sentence range from several sources.

First, case-level prison sentence range is identified from the complaint text. This include both from the plaint cover and plaint text.[44] I extract charges from complaint text, match each charges with the sentence range specified by the Criminal Code of Coduct, Acts, and Royal Decrees. Case-level maximum prison sentence range are maximum sentence range

[43] https://www.ect.go.th/ect_th/th/db_119_ect_th_download_14

[44] See example from Figure 10

from all charges of each case. Our sample contains 106,742 case-by-defendant observations. There are 109,483 case-by-defendant observations (95,128 cases) with non-missing complaint data. Out of this sample, charges and prison sentence range can be found from 56,509 cases.

To complement the first source of sentence range, I extract charges from the judgement text. Charges from the judgement are a subset of charges from the complaint text by default as the judge cannot convict the defendant with charges beyond those that are filed by the prosecutor in the complaint. The maximum prison sentence range from the judgement is thus less than or equal to the actual unobserved maximum prison sentence range. Out of 106,742 case-by-defendant observations, judgement text exist for 106,225 case-by-defendant observations (68,663 cases), and sentence range can be identified from 63,276 cases.

The maximum prison sentence range variable is identified as the maximum prison sentence range of these two data sources. Out of 106,742 case-by-defendant observations, case-level can be identified from 79,689 case-by-defendant observations (74.65%). Distribution of the coded cases are displayed in Table 9

I believe this is a fairly good match rate for several reasons. First, since I coded the data starting with the most severe laws down to the least severe laws, the missing cases are likely those with low maximum prison sentences from less severe laws, which will eventually be excluded from the study. Second, there are a large number of cases with no specified prison sentence range, and these cases are included in the missing data. Therefore, it is reasonable to assume that the missing cases contain very few instances with prison sentences of ten years or more. Third, I cross-checked the data with the annual number of treated cases that were available, and the numbers were fairly close.[45]

Code national security/international relations/terrorism offense.– These include offenses relating to the security of the kingdom (criminal code, section 107-135); terrorism (criminal code, section 135/1-135/4); Money Laundering Act, Emergency Decree on Public Administration in Emergency Situation, B.E. 2548.

Code drugs related offense.–As described in Section A.2.1, the conditions for reporting a case's draft verdict to the Office of the Regional Judge also depend on specific suboffenses in drug-related cases. The treatment cutoff is determined by factors such as the type of drug offense, drug type, quantity, weight, and purity. I extracted this detailed information using keyword searches and cross-checked it with linked registry data at a broader level, such as drug offense group, drug type, and charge severity.

[45]The number of draft verdict reported to the Office of Chief Regional Judge in region 8 is 2,345 in 2019, and 2,139 in 2020 (Source: Annual Report of the Office of Chief Regional Judge in region 8). It was 360 cases in 2009, 695 cases in 2010, and 664 in 2011 (Source: report from Legislative Institutional Repository of Thailand for draft Judicial Act, 2011)

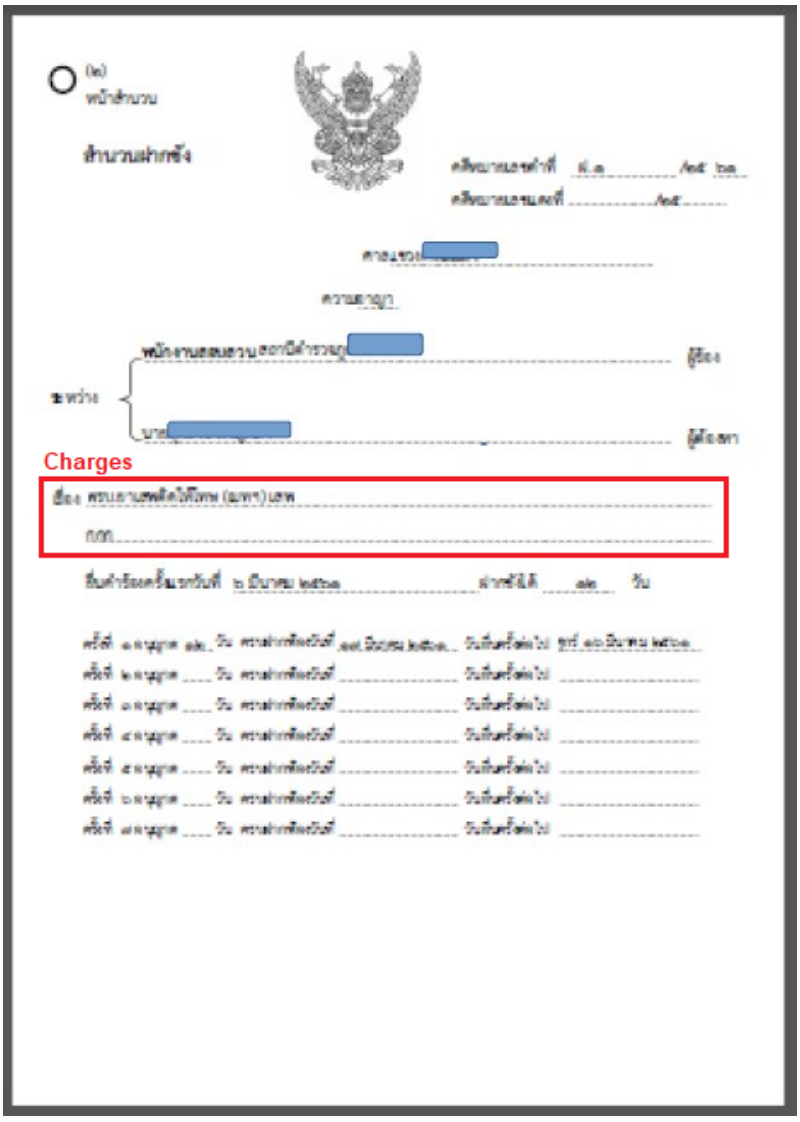
Code cases involved top bureaucrats/politicians.–As described in Section A.2.1, the conditions for reporting a case's draft verdict to the Office of the Regional Judge also depend on whether a cases involves top individual, e.g. prime minister, politicians, bureaucrats, judges, police, soldier. I extracted this detailed information using keyword searche of occupation and rank position.

Table 9: Distribution of maximum prison sentence (Obs.=defendant by case)

	Freq.	Percent
<1 year	256	0.29
1	569	0.64
2	1,248	1.40
3	1,930	2.17
4.5	121	0.14
5	9,350	10.50
7	2,161	2.43
7.5	771	0.87
10	42,476	47.69
10.5	1,298	1.46
12	2	0.00
15	7,476	8.39
20	987	1.11
22.5	7	0.01
life sentence	17,006	19.09
cap. punishment	3,412	3.83
Total	89,070	100.00

Note: This table reports the distribution of maximum prison sentence range of each case. The case-level maximum prison sentence range here is obtained from the the most severe charged of scraped from: (i) complaint texts, (ii) judgment texts, and (iii) plaint cover.

Figure 10: Example of Charges scraped from Plaint Cover.



B.3 Variable Description

Offense category

In this section, I described how treated cases are categorized for the analysis of Section 6.1.

(i) Offense involved public officers and bureaucrats: malfeasance in office offense (Criminal Code, Section 147-166), malfeasance in judicial office (Criminal Code, Section 200-205), Act on Election of Local Council Members or Local Administrators, B.E. 2002, cases involve bureaucrat and public officers.

(ii) Politically sensitive offense: offenses relating to the security of the kingdom (Criminal Code, Section 107-135), terrorism (Criminal Code, Section 135/1-135/4), Money Laundering Act, Emergency Decree on Public Administration in Emergency Situation, B.E. 2548, offenses against officials (Criminal Code, Section 136-146), offenses against judicial officials (Criminal Code, Section 200-205).

(iii) Severe drugs offense: drugs offense (produce, import, export, distribute).

(iv) Minor drugs offense: drugs offense (posess, use).

Sentence outcome

Actual sentence over maximum punishment: To construct this variable, the maximum punishment of each case is determined by the Criminal Code of Conduc as follow:

For a case convicted with single-offense, maximum punishment is equal to the upper bound of the prison sentence range.

For a case convicted with multiple-offense, maximum punishment follows from the Criminal Code of Conduct (Section 91): (i) cases with maximum prison sentence range below 3 years can receive a maximum punishment of 10 years, (ii) cases with maximum prison sentence range between 3 to 10 years can receive a maximum punishment of 20 years, (iii) cases with maximum prison sentence range above 10 years can receive a maximum punishment of 50 years.

Finally, the ratio is set to 1 for cases that received life imprisonment and capital punishment.

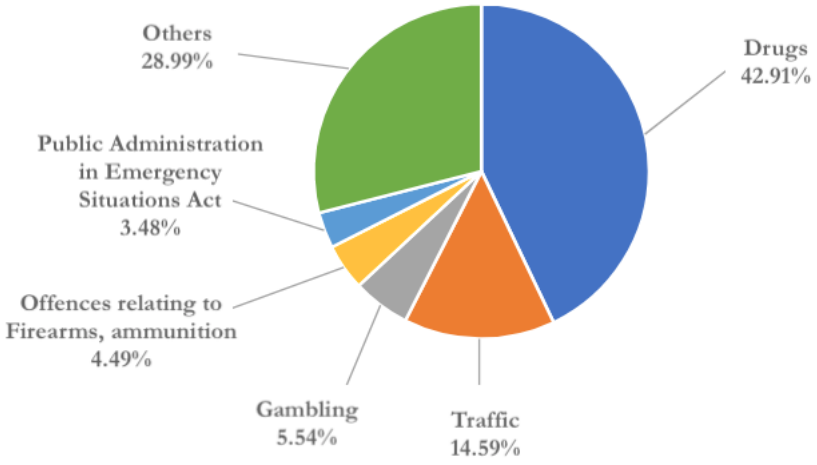
B.4 Sample Coverage

Table 10: Sample coverage by Court

(1) No.	(2) Provincial Court Name	(3) Region	(4) Observation	(5) Number of Constituency
1	Chainat Provincial Court	1	3,320	2
2	Ayutthaya Provincial Court	1	15,702	5
3	Lopburi Provincial Court	1	13,478	4
4	Angthong Provincial Court	1	7,079	2
5	Amnatcharoen Provincial Court	3	5,557	2
6	Loey Provincial Court	4	6,814	4
7	Udonthani Provincial Court	4	21,936	10
			73,886	29

Note: Observation=case by defendant. Court data (Column (1)-(4)) is merged with electoral data at the subdistrict level to obtain Number of Constituency per court (Column (5)).

Figure 11: Composition of the Thai Criminal cases from all provincial court in Thailand.



Source: Court of Justice Annual Statistics.

Figure 12: Number of observations (case by defendant), separated by maximum prison sentences.

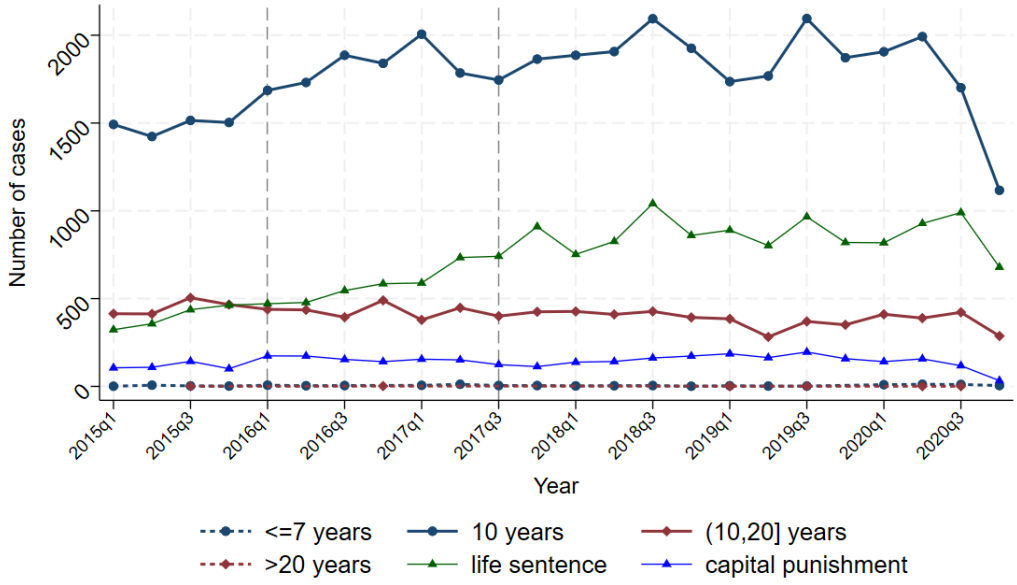


Figure 13: Number of observations (case by defendant), separated by maximum punishment and offense group.

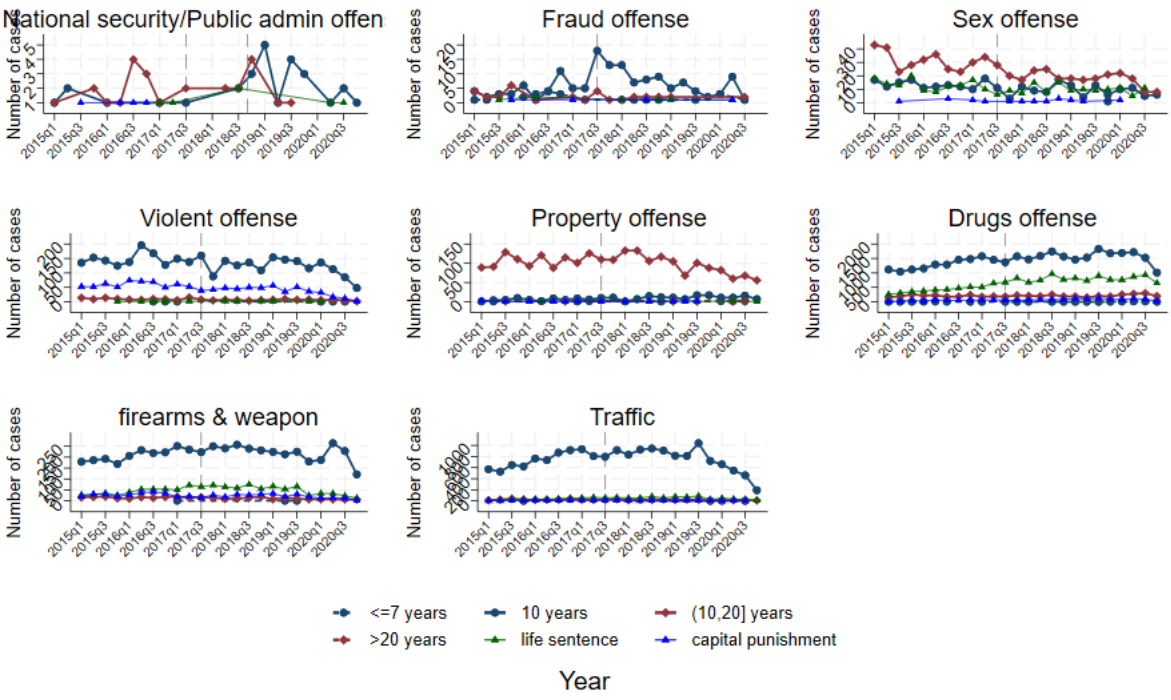
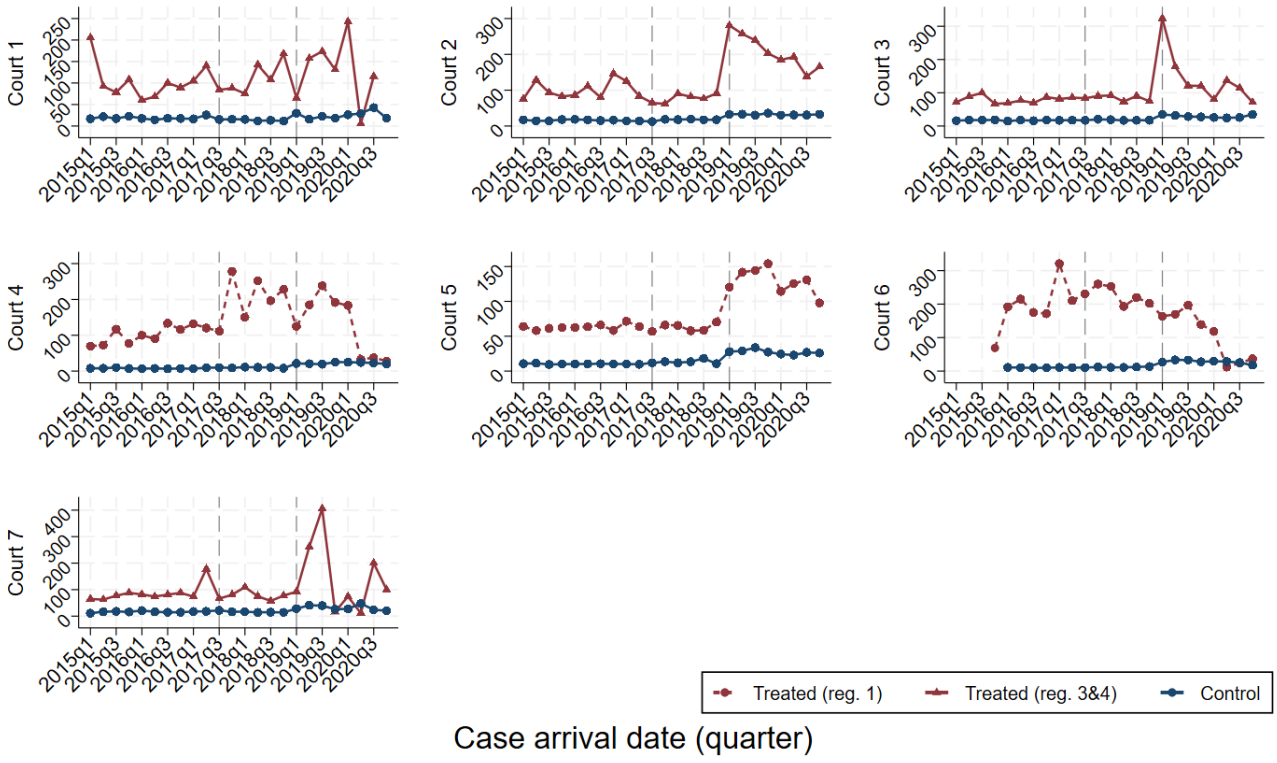


Figure 14: Average prison sentence lengths by courts



Note: Our sample contains 7 provincial courts, distributed in 3 region. Region 1(court 1,2,3,7); Region 3 (court 4); Region 4 (court 5,6)

C Identification Checks

Table 11: Effects of the reform on number of cases

	Dependent var: Number of cases				
	(1)	(2)	(3)	(4)	(5)
Treated1	-17.479***	-13.742***	-13.951***	-11.990***	-14.121***
	(1.149)	(1.297)	(0.758)	(0.959)	(1.128)
Treated1XPost	-5.658***	-8.252***	-6.442***	-8.036***	-7.973***
	(1.625)	(2.013)	(1.177)	(1.283)	(1.396)
N	3789	3513	4116	3909	3909
R-square	0.075	0.287	0.059	0.097	0.136
Mean DV.	18	19	17	17	17
Quarterly fe.	Yes	Yes	Yes	Yes	Yes
Court fe.	Yes	Yes	Yes	Yes	Yes
Court trend	Yes	Yes	Yes	Yes	Yes
Offence fe.	No	Yes	No	Yes	Yes
Quarter fe.	No	Yes	No	Yes	Yes
Prosecutor fe.	No	No	No	No	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: In Column (1)-(2), number of cases are collapsed to court-offense-treated case dummy-quarter level. In Column (3)-(5), number of cases are collapsed to prosecutor-offense-treated case dummy-quarter level. Standard errors are clustered at the offense level.

Table 12: Effects of the reform on charge composition.

	Mean		(1)		(2)		(3)	
Case characteristics	Before	After	β_{post}	*p-value*	β_{post}	*p-value*	β_{post}	*p-value*
Sample Size	27937	45949						
A. Case category: offence group (0/1)								
Fraud offense	0.003	0.006	0.002	[0.227]	-0.020	[0.087]	-0.017	[0.088]
Sex offense	0.022	0.012	-0.010	[0.036]	-0.059	[0.042]	-0.056	[0.046]
Violent offense	0.093	0.059	-0.035	[0.046]	-0.084	[0.177]	-0.076	[0.173]
Property offense	0.052	0.042	-0.009	[0.243]	-0.021	[0.675]	-0.023	[0.655]
Drugs offense	0.747	0.836	0.094	[0.001]	0.222	[0.036]	0.204	[0.042]
Firearms	0.115	0.098	-0.016	[0.232]	0.130	[0.000]	0.137	[0.000]
Traffic	0.295	0.259	-0.034	[0.174]	-0.130	[0.532]	-0.146	[0.485]
Malfeasance in office	0.008	0.004	-0.002	[0.053]	-0.022	[0.154]	-0.023	[0.229]
Political sensitive	0.007	0.004	-0.002	[0.016]	-0.026	[0.125]	-0.028	[0.155]
B. Charge severity: maximum prison sentence (0/1)								
10 years	0.610	0.569	-0.041	[0.382]	0.296	[0.128]	0.287	[0.140]
11-15 years	0.137	0.111	-0.026	[0.122]	-0.017	[0.762]	-0.016	[0.784]
16-20 years	0.021	0.009	-0.013	[0.072]	-0.045	[0.070]	-0.041	[0.079]
>20 years	0.000	0.000	-0.000	[0.803]	0.000	[0.212]	0.000	[0.281]
life imprisonment	0.178	0.262	0.087	[0.068]	-0.037	[0.604]	-0.037	[0.627]
death sentence	0.050	0.044	-0.006	[0.584]	-0.197	[0.120]	-0.193	[0.118]
Quarter fe.			No		Yes		Yes	
Court fe.			No		Yes		Yes	
Court trend			No		Yes		Yes	
Judge fe.			No		Yes		Yes	
Prosecutor fe.			No		No		Yes	

Note: This table reports results on balance test for changes in offense and charge severity. Column (3) controls for prosecutor's fixed-effects.

Table 13: Effects of the reform on case assignment

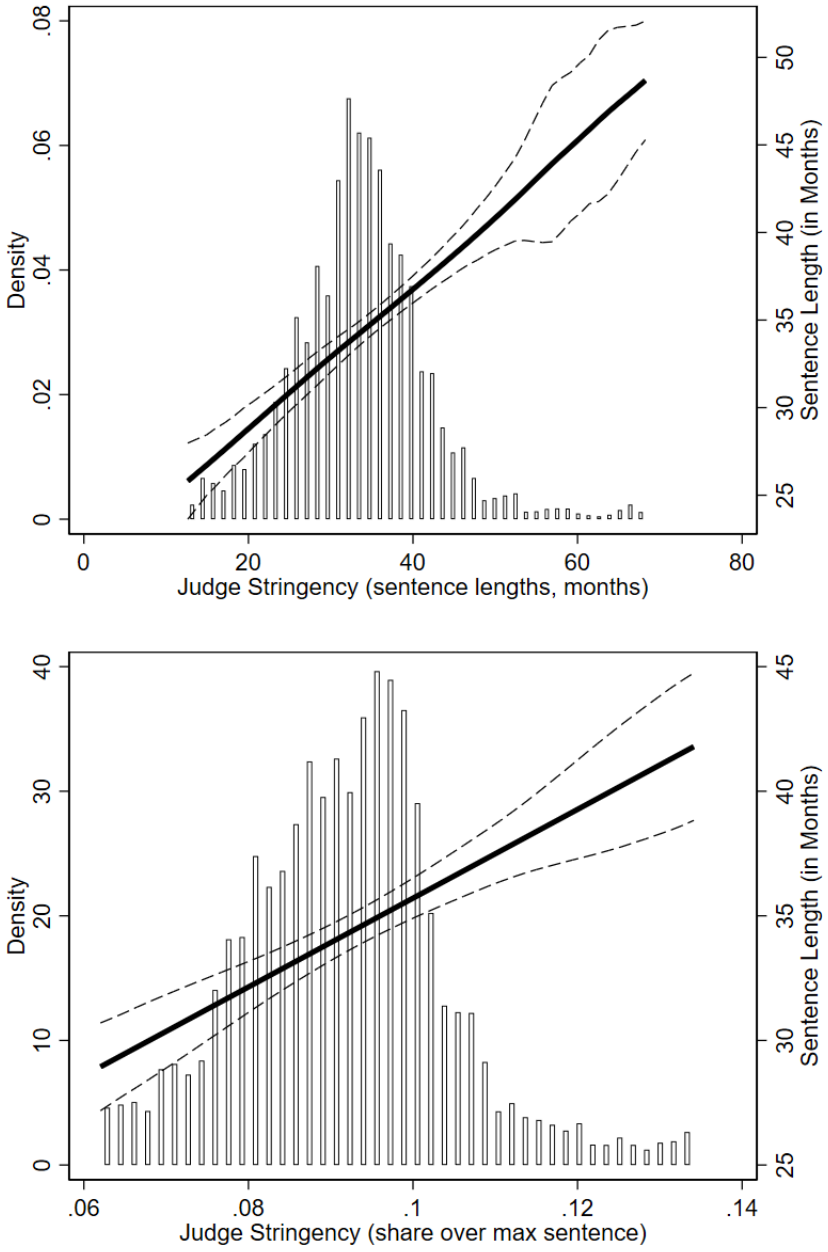
	judge-level sample	case-level sample	
	(1) Share of treated case per judge after the reform	(2) Dummy if a case is treated after the reform	(3) Dummy if a case is treated after the reform
Panel A: Judge Stringency (sentence lengths, months)			
judge_stringency1	-0.00028	0.00012	0.00007
	(0.00018)	(0.00011)	(0.00011)
r2	0.0169	0.0219	0.1595
N	150	22612	22588
Panel B: Judge Stringency (share over max. punishment)			
judge_stringency2	0.19335	0.13778	0.09405
	(0.28837)	(0.10446)	(0.09313)
N	150	22612	22588
R-square	0.011	0.022	0.160
Court X Quarter fe.		Yes	Yes
Control Vars.		No	Yes

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Column (1) regress "share of treated case each judge received after the reform" over judge stringency created by sample before the reform. Column (2) regress "dummy if a case is treated after the reform" over judge stringency created by sample before the reform, controlling for a fully interacted court-by-quarter FEs. Column (3) add control variables to Column (2). Standard errors of Column (2)-(3) are two-way clustered at the judge and defendant level.

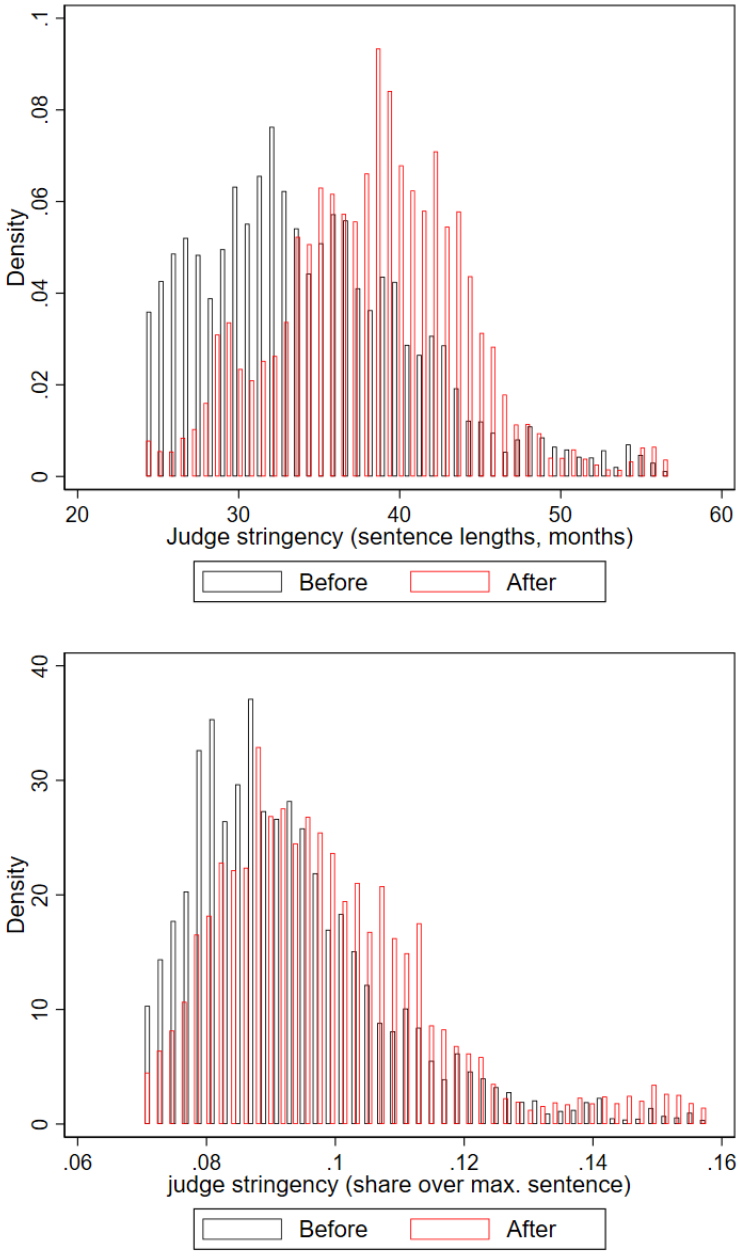
D Additional Figures, Tests, and Tables

Figure 15: Case-level judge stringency distribution.



Note: This figure depicts the relationship between prison sentence lengths (months) in the focal case and judge stringency. The baseline "pre-reformed sample" is used for estimation. Histogram shows the density of judge stringency (leaving out the top and bottom 1%). The solid line shows a Kernel-weighted local polynomial regression of prison sentence on judge stringency, while the dashed lines show 90% confidence bands. Prison sentence and judge stringency are residualized using court-by-quarter FEs and mean-standardized.

Figure 16: Judge stringency distribution (case-level)



Note: This figure depicts the density of judge stringency (leaving out the top and bottom 1%) at the case-level. The solid line shows a Kernel-weighted local polynomial regression of prison sentence on judge stringency, while the dashed lines show 90% confidence bands. Prison sentence and judge stringency are residualized using region-by-year FEs and mean-standardized.

Table 14: Covariate balance for Control Cases.

	Mean		(1)		(2)		(3)	
Case characteristics	Before	After	β_{post}	*p-value*	β_{post}	*p-value*	β_{post}	*p-value*
Sample Size	25643	43191						
A. Defendant characteristics (0/1)								
Average age	31.5	32.4	0.906	[0.003]	-0.315	[0.062]	-0.300	[0.068]
Male (0/1)	0.913	0.926	0.013	[0.008]	-0.017	[0.019]	-0.020	[0.002]
Thai (0/1)	0.997	0.997	0.000	[0.936]	0.000	[0.883]	0.000	[0.930]
Criminal history (0/1)	0.045	0.082	0.037	[0.000]	-0.002	[0.569]	-0.004	[0.401]
Charges per defendant	1.59	1.56	-0.025	[0.571]	0.118	[0.021]	0.112	[0.008]
District poverty index	0.576	0.524	-0.052	[0.002]	-0.037	[0.021]	-0.019	[0.214]
B. Offence group (0/1)								
Fraud offense	0.003	0.005	0.002	[0.243]	0.009	[0.097]	0.008	[0.041]
Sex offense	0.016	0.009	-0.006	[0.134]	-0.003	[0.242]	-0.002	[0.592]
Violent offense	0.075	0.050	-0.024	[0.125]	-0.007	[0.310]	-0.000	[0.953]
Property offense	0.054	0.043	-0.011	[0.252]	0.003	[0.645]	0.002	[0.870]
Drugs offense	0.774	0.850	0.081	[0.002]	0.021	[0.158]	0.016	[0.273]
Firearms	0.110	0.096	-0.014	[0.340]	0.019	[0.058]	0.022	[0.035]
Traffic	0.317	0.271	-0.045	[0.095]	0.070	[0.033]	0.050	[0.036]
Malfeasance in office	0.005	0.003	-0.002	[0.086]	0.003	[0.419]	-0.000	[0.338]
Political sensitive	0.005	0.003	-0.002	[0.080]	-0.002	[0.221]	-0.002	[0.109]
Quarter fe.			No		Yes		Yes	
Court fe.			No		Yes		Yes	
Judge fe.			No		No		Yes	

Note: This table reports results on predetermined case characteristics for the control cases. Each row in each panel is a separate regression. Column (1), (2), and (3) report the results of regressing these characteristics on a dummy that indicates whether a case arrives after the reform or not for the control cases. Column (2) control for quarter, court,court trend. Column (3) adds judge fixed effects. Standard errors are clustered at the charge severity-court bin level

Table 15: Covariate balance of Treated cases

	Mean		(1)		(2)		(3)	
Case characteristics	Before	After	β_{post}	*p-value*	β_{post}	*p-value*	β_{post}	*p-value*
Sample Size	2294	2758						
A. Defendant characteristics (0/1)								
Average age	33.7	33.7	0.379	[0.571]	-0.186	[0.885]	-0.311	[0.803]
Male (0/1)	0.881	0.894	0.008	[0.599]	-0.046	[0.028]	-0.031	[0.093]
Thai (0/1)	0.989	0.982	-0.002	[0.716]	0.001	[0.882]	0.004	[0.687]
Criminal history (0/1)	0.028	0.056	0.032	[0.000]	-0.008	[0.634]	-0.014	[0.362]
Charges per defendant	1.56	1.55	0.041	[0.482]	0.149	[0.003]	0.141	[0.004]
District poverty index	0.551	0.596	0.071	[0.106]	-0.002	[0.984]	-0.066	[0.159]
B. Offence group (0/1)								
Fraud offense	0.007	0.023	0.002	[0.536]	0.010	[0.336]	0.010	[0.072]
Sex offense	0.096	0.053	-0.044	[0.108]	-0.020	[0.259]	-0.008	[0.620]
Violent offense	0.299	0.186	-0.111	[0.029]	0.009	[0.809]	0.019	[0.576]
Property offense	0.028	0.029	0.004	[0.671]	0.011	[0.478]	0.001	[0.940]
Drugs offense	0.444	0.615	0.178	[0.000]	0.003	[0.952]	-0.006	[0.872]
Firearms	0.177	0.139	-0.030	[0.361]	0.017	[0.525]	0.028	[0.289]
Traffic	0.046	0.070	0.029	[0.207]	0.049	[0.148]	0.031	[0.296]
Malfeasance in office	0.041	0.014	-0.008	[0.115]	-0.005	[0.252]	-0.009	[0.243]
Political sensitive	0.031	0.017	-0.007	[0.459]	-0.009	[0.460]	-0.009	[0.593]
Quarter fe.			No		Yes		Yes	
Court fe.			No		Yes		Yes	
Judge fe.			No		No		Yes	

Note: This table reports results on predetermined case characteristics for the treated cases. Each row in each panel is a separate regression. Column (1), (2), and (3) report the results of regressing these characteristics on a dummy that indicates whether a case arrives after the reform or not for the control cases. Column (2) control for quarter, court,court trend. Column (3) adds judge fixed effects. Standard errors are clustered at the charge severity-court bin level

Table 16: Covariate balance, DiD

	(1)		(2)		(3)	
Case characteristics	β_{post}	*p-value*	β_{post}	*p-value*	β_{post}	*p-value*
A. Defendant characteristics (0/1)						
Average age	0.379	[0.570]	-0.406	[0.625]	-0.423	[0.625]
Male (0/1)	0.008	[0.598]	0.001	[0.962]	0.001	[0.959]
Thai (0/1)	-0.002	[0.715]	-0.003	[0.630]	-0.003	[0.578]
Criminal history (0/1)	0.032	[0.000]	-0.008	[0.299]	-0.008	[0.321]
Charges per defendant	0.041	[0.481]	0.008	[0.847]	0.007	[0.857]
District poverty index	0.071	[0.104]	0.027	[0.463]	0.020	[0.582]
B. Offence group (0/1)						
Fraud offense	0.002	[0.535]	-0.001	[0.826]	-0.000	[0.971]
Sex offense	-0.044	[0.106]	-0.037	[0.178]	-0.036	[0.181]
Violent offense	-0.111	[0.028]	-0.082	[0.099]	-0.082	[0.088]
Property offense	0.004	[0.670]	0.018	[0.254]	0.018	[0.256]
Drugs offense	0.178	[0.000]	0.097	[0.030]	0.092	[0.032]
Firearms	-0.030	[0.359]	-0.026	[0.478]	-0.025	[0.479]
Traffic	0.029	[0.205]	0.040	[0.146]	0.031	[0.222]
Malfeasance in office	-0.008	[0.113]	-0.007	[0.157]	-0.005	[0.309]
Political sensitive	-0.007	[0.457]	-0.005	[0.575]	-0.005	[0.628]
Quarter fe.	No		Yes		Yes	
Court fe.	No		Yes		Yes	
Judge fe.	No		No		Yes	

Note: This table reports results on predetermined case characteristics for the difference between treated and control group. Column (1), (2), and (3) report the results of regressing these characteristics on a dummy that indicates whether a case arrives after the reform or not for the control cases. Column (2) control for quarter, court,court trend. Column (3) adds judge fixed effects. Standard errors are clustered at the charge severity-court bin level

Table 17: Heterogeneity by within-region Judge Stringency

	Measure of Judge Stringency: Sentence lengths (months)		Measure of Judge Stringency: Share over max. sentence	
	(1) Lenient judges	(2) Strict judges	(3) Lenient judges	(4) Strict judges
Panel A: All regions				
$Treated^1_{gr}$	89.078***	152.561***	93.046***	154.266***
	(18.177)	(21.453)	(18.893)	(21.116)
$Treated^1_{gr} \times Post_t$	74.364***	5.195	59.560***	6.875
	(12.463)	(9.396)	(15.152)	(9.199)
r2	0.411	0.391	0.369	0.407
ymean	29	38	30	38
N	18579	32219	20686	30112
Panel B: By region				
$Treated^1_{gr}$	89.311***	152.897***	92.717***	154.651***
	(17.840)	(20.989)	(18.037)	(20.848)
$Treated^1_{gr} \times Post_t \times Reg1$	90.444***	18.050	109.891***	13.118
	(11.636)	(16.819)	(16.939)	(12.675)
$Treated^1_{gr} \times Post_t \times Reg3$	189.685	44.404	74.820	49.426
	(178.706)	(35.415)	(83.358)	(36.475)
$Treated^1_{gr} \times Post_t \times Reg4$	43.660*	-22.971	18.736	-15.189
	(25.625)	(25.459)	(25.755)	(26.574)
r2	0.414	0.394	0.376	0.409
ymean	29	38	30	38
N	18579	32219	20686	30112

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable = prison sentence length. Life imprisonment is set to 60 years (720 months) of prison sentence. Capital punishment is set to 100 years (1200 months) of prison sentence. Each column is a subsample regression. Lenient judges are judges with stringency score below median score of the region. Strict judgese are judges with stingency score greater than or equal to median score of the region. All columns control for observables and unobservables as described in equation (2).

Table 18: Summary statistics of Judge Stringency

Variable	(1) Before	(2) After	(3) β_{post}	(4) *p-value*
Panel A. Judge stringency = sentence lengths				
Mean	34.125	38.656	4.603	[0.000]
IQR (Q75-Q25)	11.213	7.828		
IQR (P90-P10)	22.065	15.809		
Panel B. Judge stringency =share over max sentence				
Mean	0.092	0.100	0.008	[0.000]
IQR (Q75-Q25)	0.019	0.021		
IQR (P90-P10)	0.040	0.043		

Note: Column (1) and (2) report the distribution of judge stringency measures (mean and inter quartile range), before and after the 2017 judicial reform. Column (3) provides the coefficient for the difference.