



Rules for thee but not for me: Selective privacy enforcement in Chinese court judgments

Hui Zhou^{ID}, Genia Kostka^{*} ^{ID}

Institute for Chinese Studies, Freie Universität Berlin, Germany

ARTICLE INFO

Keywords:

Data protection
Privacy infringement
Cybercrime
Law enforcement
PIPL

ABSTRACT

Against the backdrop of digital transformation, governments worldwide have enacted privacy laws to safeguard personal information, yet their judicial enforcement remains understudied. This article develops a selective privacy enforcement perspective and evaluates it using an original dataset of 6,370 civil and criminal judgments collected from China. We argue that, in authoritarian legal systems, privacy laws may be enforced more effectively against private actors than against public institutions or state-connected defendants. The findings provide partial support for this perspective, principally in civil litigation. Plaintiffs fare substantially worse when suing public organizations than when pursuing claims against private actors. In criminal cases, however, defendants affiliated with public organizations receive no statistically detectable advantage over unaffiliated defendants. PIPL invocation is associated with longer prison sentences and larger fines, regardless of state affiliation. These findings cast doubt on the proposition that political connections uniformly protect privacy offenders. Instead, they suggest that privacy enforcement may depend on how the state views the threat of privacy violations. This study underscores the importance of privacy laws in protecting personal information on the one hand, while highlighting their limitations within authoritarian contexts on the other.

1. Introduction

Digital transformation is taking place across both private and public sectors worldwide, prompting governments to tackle data protection challenges resulting from the transformation. Digitalization increases the risk of data leaks, from both private and government databases. In 2024, for example, the European Parliament disclosed a massive breach on its recruitment platform, which exposed sensitive information of over 8000 staff members, including passports, employee ID cards, and criminal records [1]. Data leaks in authoritarian contexts can reach a much greater scale. In 2022, a leader of Vietnam's Ministry of Public Security revealed a case concerning the illegal acquisition and sale of thirty million personal data, which account for nearly 30% of the nation's entire population [2]. In the same year, China experienced an even more serious data leak, in which the personal information of over one billion residents was hacked and posted online for sale [3].

Additionally, privacy disputes can arise between individuals and private entities. For instance, the installation of a home security camera might inadvertently infringe on a neighbor's privacy. Routine online activities such as banking and shopping have fueled numerous

cybercrimes such as identity theft and online scams [4]. Moreover, companies collect vast amounts of user data. Inadequate security measures can result in data breaches that expose sensitive information [5]. Employees may also share or sell user data without authorization. More frequently, many businesses capitalize on user data for analytics, targeted marketing, and profit, a phenomenon which Zuboff [6,7] refers to as "surveillance capitalism."

Facing the rising threats to personal privacy, both democratic and authoritarian states have resorted to data protection laws. In 2018, Consumers International [8] pointed out that only 58% of the countries in the world had enacted specific laws to protect personal data. In less than a decade, 79% of countries have adopted privacy laws as of May 5, 2025, according to United Nations Conference on Trade and Development [9]. Liberal democratic states are the leaders in protecting personal data and privacy, but authoritarian states are quickly catching up. China, for example, enacted the Personal Information Protection Law (PIPL) in 2021, a comprehensive data protection law that addresses many shortcomings common in its previous regulatory framework, such as the absence of data processing principles and insufficient protection for data subjects [10]. Vietnam, by contrast, started implementing its

^{*} Corresponding author at: Fabeckstr. 23/25, Room 1.1123, 14195 Berlin, Germany.

E-mail addresses: hui.zhou@fu-berlin.de (H. Zhou), genia.kostka@fu-berlin.de (G. Kostka).

<https://doi.org/10.1016/j.clsr.2026.106392>

first dedicated privacy law in 2026. Despite this legislative progress, the privacy laws in authoritarian states may differ from those in liberal democracies in important ways, such as the definition of major concepts, enforcement procedures, and the strategic priority of the legislation [11, 12].

Data protection laws are important mechanisms for safeguarding personal information. The data breach notification laws in the United States, for example, significantly reduced medical identity theft reports by 3.5 reports per 100,000 people [13]. In Israel, a data protection law changed the privacy management practices of popular commercial websites [14]. While privacy legislation may be effective in preventing privacy disputes, it remains unclear as to whether it plays a role in changing litigation outcomes. Romanosky et al. [15] appear to have conducted the earliest study on data breach litigation based on large-N cases. However, their goal is to specify the conditions under which a data breach will be litigated and if so, whether the litigation results in settlement or dismissal. There are many more privacy-related questions that remain to be answered. What is the probability of winning a privacy case? Do judges treat all parties equally (e.g., individuals vs. organizations)? How do they punish perpetrators for violating personal privacy? Scholars have not yet paid sufficient attention to these questions.

This paper aims to answer the above questions based on an original dataset of 6,370 personal information lawsuits adjudicated in China between 2010 and 2024. The dataset comprises 532 civil cases and 5,838 criminal cases drawn from China Judgments Online (中国裁判文书网), a publicly available platform founded by China's central authority to publish judicial documents. With the help of AI-assisted coding, we extracted detailed information from privacy disputes. Quantitative analysis suggests that the 2021 Personal Information Protection Law (PIPL) increases the likelihood that plaintiffs will prevail against private entities, but this effect vanishes when the defendant is connected with the public sector. In other words, Chinese plaintiffs facing state-connected privacy violators are more likely to receive unfavorable outcomes such as a complete loss. Additionally, state-connected actors do not appear to have an advantage in criminal lawsuits, and whether the court invokes the PIPL matters in this setting. In summary, China's PIPL can enhance the ability of plaintiffs to confront private entities in privacy lawsuits, but not the public sector. This effect is limited to civil cases, suggesting the dual purpose of the privacy law—to shield the public sector in civil disputes and to punish all criminal offenses regardless of the identity of privacy violators.

The contribution of this study is twofold. First, this article sheds new light on information politics in China, which is widely regarded as a resilient authoritarian state [16] and a key actor in global information governance [17,18]. Our analysis reveals a pattern of selective privacy enforcement. Although the PIPL is associated with judicial outcomes in both civil and criminal cases, political connections appear to constrain its role in civil disputes. This finding suggests a distinction between institutional shielding and agent discipline: the state may shield affiliated institutions from civil liability while declining to offer similar protection in the case of criminal offenses. This finding adds nuance to existing accounts of authoritarian resilience by showing how institutional responses may vary across different types and levels of legal violations [19].

Second, China's experience provides broader insights into the implementation of privacy legislation in jurisdictions where courts operate under political constraints. Despite the prominence of the PIPL as a comprehensive privacy statute [20,21], scholars have yet to systematically examine how it is applied in privacy lawsuits. Drawing on large-scale litigation data, this article maps the landscape of privacy litigation in China and identifies factors associated with judicial outcomes. The findings demonstrate why the formal comprehensiveness of privacy legislation should be evaluated separately from its implementation and institutional constraints. This distinction also contributes to comparative debates about the enforcement of data protection laws, including those concerning European Union member states [22,23].

2. Privacy protection: a global challenge

2.1. Privacy protection in liberal democracies

Privacy is closely linked to individual freedom and sometimes even considered a human right [10,24,25]. Thus, it is taken seriously in liberal democratic jurisdictions such as the United States and the European Union. Rather than building a unified legal framework for governing data security and privacy [26], the United States adopts a fragmented approach, with various laws enacted to regulate specific sectors and industries. Major federal privacy laws include the Electronic Communications Privacy Act, the Health Insurance Portability and Accountability Act, and the Telephone Consumer Protection Act. Recent legislative efforts toward a dedicated privacy law include the American Data Privacy and Protection Act of 2022 and the American Privacy Rights Act of 2024, neither of which became law. Additionally, various state laws such as the California Consumer Privacy Act are playing a role in protecting privacy.

Much of the literature on the U.S. privacy litigation focuses on establishing standing [27–29], which concerns whether a case warrants judicial remedy. Plaintiffs can file a lawsuit if they demonstrate that the defendant's misconduct has led to an “injury in fact”, thus establishing the “standing” of the case. A data breach case typically proceeds through several stages, including (1) a plaintiff files a complaint, (2) a defendant can file a motion to dismiss (MTD) the claim, (3) the judge evaluates the MTD, (4) the plaintiff requests information from the defendant if their claim survives the MTD (otherwise the case dies here), a process known as “discovery”, (5) either the plaintiff or defendant may call for a “summary judgment” by the court without holding a trial, (6) the case proceeds to a lengthy and expensive full trial if no summary judgment is requested. However, privacy disputes rarely end up with a trial because the defendant almost always prefers to settle rather than litigate if the “discovery” stage fails to block the plaintiff [26].

The European Union represents a distinct approach to privacy protection. The 2018 General Data Protection Regulation (GDPR) is a comprehensive and unified legal framework for governing data management and privacy protection within the European Union. It greatly enhances data subjects' control over their personal data by granting them many rights and imposing many obligations on organizations. For instance, data controllers are usually required to respond promptly in the case of a data breach, and punishment including fines may apply if they fail to do so. This law influences not only organizations and individuals within the European Union but also foreign entities that process personal data of EU data subjects. The GDPR has become a prime example of data protection law and has influenced many other countries through the so-called “Brussels effect” [25,30].

In a multi-level governance framework, GDPR provisions are enforced in EU member states by their respective supervisory authorities, which inevitably leads to regional variation in the privacy protection policies and practices [31]. In Germany, for example, privacy disputes can be handled by either administrative courts or civil courts (i.e., district or local courts), depending on the type of action taken by the supervisory authority. In the case of an administrative order issued to a relevant party, the administrative court can take charge. If a fine is to be imposed or if damages are involved, the competent body would be the district or local courts. If a person has questions about the decisions of a supervisory authority, they still have a chance to lodge a constitutional complaint before the Federal Constitutional Court [32].

2.2. Privacy protection in China

Compared with democratic countries, authoritarian states have a notorious reputation for disregarding personal privacy. For instance, Villaseñor [33] contends that “in authoritarian countries, however, there is essentially no open privacy debate.” Diamond [34] also asserts that “authoritarian regimes, however, grant their citizens no effective

rights of privacy and have no compunction about demanding unfettered access to all this data.” However, this is no longer the case as many authoritarian states such as China and Vietnam have stepped up efforts to improve their regulatory framework for privacy disputes. Today, the degree of privacy regulation in China is comparable to many industrialized nations [17].

China’s case is especially intriguing when compared with the European Union and the United States. First, China is a major center for data generation and processing and also a key player in the AI race, owing to its vast domestic market, significant international presence, and the government’s supportive stance on technological innovation and digital governance. However, the lack of adequate data protection has resulted in a privacy trust crisis, with a considerable number of individuals expressing distrust in private and public organizations in handling their personal data [35]. Thus, unraveling the landscape of privacy violations may help scholars and policymakers better understand the risks and costs associated with privacy regulation. Second, China stands in stark contrast to industrialized nations and regions in many aspects, such as political system, socioeconomic development, cultural context, etc. Moreover, its privacy legislation and judicial environment also differ significantly. Therefore, China’s experience of privacy dispute resolution has the potential to inform privacy protection in less affluent and authoritarian contexts.

The cornerstone of China’s privacy protection framework is the *Personal Information Protection Law* (PIPL, hereafter), which was passed in August 2021 and took effect in November 2021. Before this law, several other pieces of legislation had already been enacted to address different aspects of personal information, including the 2015 National Security Law, the 2017 Cybersecurity Law, the 2021 Data Security Law, and the 2021 Civil Code, which contains a dedicated chapter on privacy and personal information protection.¹ As the country’s first comprehensive data protection law, the PIPL provides a legal framework for obtaining, storing, using, processing, and transferring personal information in China. Under the influence of the GDPR, this law establishes mainstream data-processing principles (e.g., lawfulness, accuracy, and data minimization), defines many key terms (e.g., data controllers and data subjects), differentiates related concepts (e.g., personal information vs. sensitive personal information), and clarifies the rights and obligations of key actors in the data ecosystem. Additionally, it addresses some privacy concerns such as the “big data swindling” phenomenon through an explicit ban on automated decision-making [36].

Why does China enact privacy laws? Jia [37] argues that the enactment of the PIPL results from both centralized strategic goals and grassroots calls for legitimacy. Specifically, extensive privacy infringements force the party-state to present itself as a “protective guardian of citizens’ rights against other intrusive actors.” Privacy laws boost the party-state’s political legitimacy while simultaneously achieving three key objectives: advancing the digital economy, protecting national security, and enhancing international standing. Whereas the law generally aligns with the global trend of privacy protection [38], scholars of privacy typically attribute China’s lawmaking effort to national security concerns rather than a genuine interest in protecting privacy [11,36]. Therefore, it is imperative to assess the effectiveness of the privacy legislation using empirical data.

3. Legal dualism and selective privacy enforcement

Law can operate as a combination of legal order and arbitrary rule. This observation was first noted by Fraenkel in his book published in 1941, in which he studied Germany in the Nazi era and found the coexistence of a normative state and a prerogative state [39]. Under most circumstances, the government could enforce laws appropriately regarding economic and civil matters. However, at other times, the government exercised power arbitrarily without being punished by the law, especially when it came to national security and race. This is what he referred to as the “dual state.” Subsequent scholarship has adopted

this concept and similar terms such as legal dualism to examine the multifaceted role of law and the judiciary in authoritarian systems [40–42].

Legal dualism provides a useful framework for studying judicial politics in authoritarian settings. A central question in this literature is whether courts enforce legal rules impartially across litigants and issue areas. Existing research identifies political connections as an important source of unequal treatment. Private actors connected to the state tend to enjoy advantages in commercial disputes [43,44]. When the government itself is involved in litigation, civil plaintiffs may also find it particularly difficult to prevail against public organizations, especially when those organizations possess substantial political or administrative influence [45]. These findings suggest that authoritarian courts can resolve routine disputes while simultaneously protecting the institutional interests of the state and politically connected actors.

Privacy litigation provides a particularly informative context for examining legal dualism. The state occupies several potentially conflicting roles in privacy governance: it enacts and enforces data protection rules, processes extensive amounts of personal information, and may itself be accused of violating privacy. Whether privacy law can be enforced effectively against the state has been a persistent concern among legal scholars and privacy observers. Before the enactment of the PIPL, scholars questioned whether the legislation would meaningfully constrain Chinese state organs [38]. After the law took effect in 2021, Yang [46] argued that China’s approach to privacy protection aligns citizens and the state against private companies. Under this arrangement, privacy law may effectively regulate private organizations while imposing substantially weaker constraints on public authority.

Comparative research on authoritarian privacy legislation provides further support for this concern. Based on an analysis of major data protection laws in China, Vietnam, Russia, and Iran, Kostka and Zhou [47] argue that authoritarian governments can adopt meaningful privacy laws for governance and legitimacy purposes while retaining broad legal exceptions that preserve state surveillance and access to personal data. Such exceptions may exempt public authorities from restrictions and obligations imposed on private entities. As a result, state agencies may collect extensive amounts of personal information or process such information without obtaining informed consent from data subjects.

Building on the literature on legal dualism and selective privacy, we develop a framework of selective privacy enforcement, with a particular focus on whether privacy law and political connections correlate with litigation outcomes. Selective privacy enforcement represents a domain-specific manifestation of legal dualism: privacy law may operate effectively in routine disputes while becoming less constraining when enforcement implicates public institutions or state-connected actors. Authoritarian governments may use privacy legislation to regulate private organizations, resolve social disputes, respond to citizens’ concerns, and strengthen political legitimacy. At the same time, the application of privacy law may become more limited when it would impose liability on the state or otherwise conflict with politically important institutional interests.

This framework is consistent with Jia’s argument that privacy legislation can enhance the Chinese government’s political legitimacy. Privacy legislation demonstrates responsiveness to citizens’ demands and presents public authority as governed by legal rules. It may thereby foster performance-based legitimacy by showing that the state is capable of identifying grievances, delivering valued protections, and controlling abuses through law [37]. These benefits do not necessarily indicate a broader commitment to judicial independence, the rule of law, or democratic governance. Rather, selected legal protections may generate public support while preserving state discretion when individual privacy claims conflict with institutional or political priorities.

The selective enforcement framework therefore leads us to expect that political connections will produce favorable treatment in both civil and criminal privacy litigation. This study draws on privacy lawsuits collected from China to test the framework. We illustrate the theoretical

expectations in the ensuing section.

4. Research hypotheses

Rather than being independent of the government, China's judiciary is part of the local governance structure characterized by the judiciary's horizontal embeddedness into the local Party-state system [48]. The Standing Committee of the Party Committee—the top decision-making body of the Chinese government at each administrative level—is composed of several political elites (e.g., party secretaries and mayors), but the leaders of the judiciary are not among them.² Therefore, it is difficult (if not impossible) for courts to stay impartial when the defendants turn out to be government agencies, especially when the latter have a higher administrative rank [45]. Second, public organizations such as state-owned enterprises (SOEs), public hospitals, and public universities have close connections with the government, not only through the financing channel but also through personnel assignment. Leaders of SOEs, for example, have administrative ranks comparable to public servants, which facilitates their career mobility across different SOEs and government agencies [49]. Even when the defendant has no direct connection with the judiciary, the patronage network within the government can also influence the defendant-court relationship. Based on this reasoning, public sector employees may be more likely to receive lenient sentences in privacy lawsuits. Hence, we formulate the first hypothesis as follows.

H1.1: In civil cases, plaintiffs suing public organizations will have a lower chance of winning than when suing private entities.

H1.2: In criminal cases, defendants affiliated with public organizations will receive more lenient sanctions than unaffiliated defendants.

In addition to the identity of defendants, we are also interested in whether the PIPL has an effect on the litigation outcome. Since its passage, no research has systematically evaluated whether the PIPL influences judicial outcomes. We anticipate that the law has had a significant effect on judicial outcomes. First, the PIPL has significantly strengthened protection for data subjects and imposed a wide range of obligations on data controllers and processors. As a result, privacy violations that previously lacked clear legal grounds may now fall within the regulatory scope of the PIPL. For instance, prior to the PIPL, data controllers might process the personal information of minors under the age of fourteen in the same manner as they did for adults. In such cases, plaintiffs seeking special protections for this group may have failed due to the absence of explicit legal provisions. However, under the PIPL, courts may be more likely to rule in favor of such plaintiffs, as the law includes multiple provisions aimed at safeguarding the rights of minors—such as requiring separate data handling rules and obtaining the consent of their parents or legal guardians. Second, the PIPL has a chapter on legal responsibilities, which includes various types of sanctions that range from compensation and admonition to fines and revocation of business licenses. These provisions provide judges with clearer guidance on what sanctions should be imposed.

The PIPL may also influence criminal cases. First, the PIPL is enacted by the national legislature. As such, violations of the PIPL may result in more severe legal consequences compared to breaches of administrative regulations or departmental rules, which hold a lower position in China's legal hierarchy. Second, criminal cases can end up with a wide range of punishments, including but not limited to detention, imprisonment and fines. The sanctions listed in the PIPL, albeit mostly administrative in nature, can also be adopted by the courts to punish perpetrators. Therefore, we expect that the PIPL will make a difference in both civil and criminal cases.

H2.1: In civil cases, the PIPL will make it easier for plaintiffs to win.

H2.2: In criminal cases, the PIPL will be associated with harsher sanctions.

Privacy and personal data protection are not absolute rights. Data protection laws commonly permit personal data to be processed without the data subject's consent for specified public purposes. Under Article 6 (1)(e) of the EU General Data Protection Regulation (GDPR), for example, processing is lawful when it is necessary to perform a task carried out in the public interest or in the exercise of official authority. Under Article 6(3), such processing must have a basis in EU or Member State law, pursue an objective in the public interest, and be proportionate to the legitimate aim pursued. Non-consensual processing under the GDPR is therefore an alternative lawful basis rather than an exemption from data protection law, and it remains subject to the GDPR's substantive and procedural safeguards.

The legal frameworks in authoritarian contexts, however, may afford state actors broader exemptions from particular obligations. Vietnam's Personal Data Protection Law of 2025 illustrates this possibility. The Law generally requires data controllers and processors to assess the impact of personal data processing, but Article 21(6) exempts competent state agencies from this requirement. The provision itself does not specify which state agencies qualify as competent or prescribe conditions governing the use of the exemption. Likewise, China's PIPL applies to state organs while qualifying some of their obligations. Article 35 requires state organs to notify individuals when processing personal information in the performance of their statutory duties. Notification is not required, however, where laws or administrative regulations require confidentiality or where notification would impede the performance of those duties [50]. In litigation, these exceptions may narrow the obligations of public sector actors and allow them to argue that their failure to notify or comply with an otherwise applicable procedural requirement was legally justified. As a result, the PIPL may produce a smaller change in judicial outcomes when a public sector organization, rather than a private entity, is the defendant.

Another explanation for the public sector's advantage has to do with the power structure of the judiciary. Despite the political-legal dichotomy in China's party-state regime, the recent decade has witnessed a weakening role of law and the strengthening primacy of politics [51], resulting in what Minzner [52] calls "China's turn against law." In judicial practice, Chinese courts operate under significant political and institutional constraints. Although recent reforms have strengthened judicial review in routine administrative disputes, courts remain cautious in politically sensitive cases and tend to rely on procedural grounds rather than undertake substantive scrutiny of government action [53]. This institutional context may limit judicial protection of privacy when individual claims conflict with asserted public interests, such as pandemic control [37]. Additionally, the PIPL is a general law on personal information protection. Its provisions on legal responsibility are mostly concentrated on administrative punishment such as fines and admonition. In reality, judges primarily rely on the Civil Code and the Criminal Law to impose sanctions, especially in criminal cases.³ Therefore, the court may issue more lenient verdicts for public defendants (e.g., lower fines or shorter prison sentences) even though the PIPL has been applied. In summary, although we expect the PIPL to influence judicial outcomes, we hypothesize that defendants with state affiliations may be largely shielded from its influence.

H3.1: In civil cases, the effect of the PIPL on win rate is greater when plaintiffs confront private entities than when they sue public organizations.

H3.2: In criminal cases, the PIPL is associated with harsher punishments for defendants from private entities than for those connected with the public sector.

5. Methods

5.1. Data and coding

Our data come from the *China Judgments Online* (CJO), a publicly available website created in 2013 by the Chinese central government to publish judicial documents nationwide. As a milestone data infrastructure project aimed at improving judicial transparency [54], this platform has published approximately 150 million documents as of 2024. It serves as a highly valuable data source widely used by scholars from various disciplines to study China-related issues. The bulk of our data was originally collected by a third-party data broker called *Make Shuju Wang* (马客数据网) based on a full-scale web scraping of first-trial cases completed in September 2021. After purchasing this dataset, we further expanded it by manually collecting first-trial privacy cases tried between October 2021 and December 2024.

As a diagnostic check, we compared the age distributions of plaintiffs in the commercially collected and manually collected samples. We selected age because it is one of the few demographic characteristics recorded with sufficient consistency in both samples and can be compared without relying on substantive coding judgments. The differences in the parties' mean ages are small, as reported in the Appendix on page A-5. This comparison increases the plausibility of merging the two samples, although it does not establish their equivalence. To account for the potential selection bias, we included a dummy variable *Data Type* in the regression analysis.

We filtered cases using keywords that are readily available in the CJO database. Examples of these keywords include the level of court and the type of document (e.g., judgment document vs. mediation document). A detailed illustration of the filtering process is provided in the Appendix on page A-4. Based on six keywords, we filtered and collected 6,370 first-trial privacy cases dating from 2010 to 2024.⁴ This sample accounts for 56.71% of all privacy-related cases hosted on the website as of October 2025. While it does not represent the full population of privacy cases in China, its size is considerably larger than in prior privacy studies. For example, Xu et al. [55] rely on 205 cases to study criminal privacy infringement, whereas Zhang et al. [56] compile 507 identity theft lawsuits taking place in Shanghai. Additionally, our dataset may include cases that are no longer available in the CJO database, as some cases have been deleted by the Chinese government in recent years [57]. Despite the large sample size, it should be noted that the application of the PIPL is rather rare in China's privacy litigation. More specifically, it was referenced in only 98 of the 1,395 cases (7.03%) tried after the PIPL took effect. The small number of decisions expressly referencing the PIPL means that estimates based specifically on its judicial application should be interpreted cautiously. Details on the distribution of the data are provided in the Appendix on pages A-6 and A-7.

We utilize large language models (LLMs) to code the judicial documents due to their increasing capabilities and growing popularity in text analysis [58]. Compared with the human coder, LLMs have several advantages, including efficiency, consistency, and immunity to fatigue. We adopt three mainstream LLMs: Meta's Llama3.2:latest (2024), Alibaba's Qwen2.5-coder (14b), and OpenAI's GPT-4o. Based on a test of 319 cases, GPT-4o outperformed the other two models by an overwhelming margin. It not only yielded results in the desired format for all 319 cases, but also achieved an accuracy rate of 97.18%. In contrast, only 23.2% and 47.96% of cases were processed by Llama3.2 and Qwen2.5, respectively. Among those cases successfully processed, their accuracy rates were 77.02% and 81.05%, respectively. For details about the performance of the three LLMs, please refer to the figure and the table on pages A-15 and A-16.⁵

5.2. Variables

Court rulings contain rich information such as the identities of concerned parties, the content of privacy disputes, legal reasoning, and

judicial outcomes. We extract a number of variables from those documents. The main variable of interest is the outcome of a lawsuit, and it depends on the type of case. For civil cases, we look at who wins the litigation. We code it as a categorical variable consisting of three possibilities from the perspective of a plaintiff: (1) a complete win, in which all of the plaintiff's claims were supported by the court; (2) a partial win, where some but not all demands were upheld by the court; and (3) a complete loss, in which all of their demands were dismissed by the court. If the case is criminal in nature and thus does not involve any plaintiffs, we examine the type and degree of criminal punishment against the privacy offender. Common criminal punishments include control (*guan zhi*, 管制), detention (*ju yi*, 拘役), imprisonment (*you qi tuxing*, 有期徒刑) and fine (*fajin*, 罚金). Additionally, some criminals are granted probation (*huanxing*, 缓刑), which is also of interest in this study.

We have two key independent variables. The first variable, *PIPL*, is a dummy variable capturing the role of the PIPL. We focus on cases in which the PIPL was applied. Cases in which the PIPL was applied were coded as 1, and otherwise as 0.⁶ The second variable, *State*, is also a dummy variable documenting the defendant's identity. In civil cases, the defendants can be either individuals or organizations. In criminal cases, however, the defendants are usually natural persons. We classify defendants into three categories—individuals, private organizations, and public organizations (or individuals connected with public organizations, in criminal cases). A critical task in the coding is to distinguish private organizations from public organizations. Several types of public organizations are deemed connected with the state, including government agencies, state-owned enterprises (SOEs), and public institutions such as public universities and hospitals. If a defendant is part of or connected with the public sector, the *State* variable takes a value of one. Otherwise, it equals zero.

We control for a host of personal characteristics and socioeconomic variables, including the plaintiff's age, gender, the claimed compensation, the number of claims (civil cases); the legal representative of both plaintiffs and defendants; the defendants' criminal background (criminal cases); and the year of the trial and the province of the court. We include the plaintiff's age and gender as pre-treatment controls for observable differences in litigant and case composition. Previous research identifies demographic variation in privacy perceptions [59, 60], digital literacy skills [61], and engagement with legal or institutional remedies (e.g., via income and hukou, see [62]). These characteristics may consequently be associated with litigation outcomes. Therefore, we include age and gender to account for this observable heterogeneity, not to estimate causal discrimination by judges. We did not include income and hukou because court documents usually do not contain such information.

5.3. Models

Since civil and criminal cases are associated with different types of outcomes, we adopt different statistical models. For civil cases, the outcome variable includes three categories for plaintiffs, a complete win, a partial win, and a complete loss. Thus, we adopt multinomial regressions so that the three outcomes can be modeled simultaneously, with the complete win serving as the reference group. The regression yields two sets of coefficients, one for the analysis of partial win relative to complete win, and the other for the analysis of complete loss relative to complete win. For criminal cases, the outcomes take various forms, including an economic value (e.g., fines), a time period measured in months (e.g., a sentence to prison), or a dummy variable (e.g., whether probation is granted). Thus, we adopt OLS for numeric variables and logit regression for probation when it is treated as a dummy variable.

$$y_{it} = \beta_0 + \beta_1 \text{State}_{it} + \beta_2 \text{PIPL}_{it} + \beta_3 \text{State}_{it} \text{PIPL}_{it} + \lambda Z_{it} + \delta_t + \gamma_i + \epsilon_{it}$$

y_{it} denotes the outcome of the litigation. The regression coefficient β_1 captures the effect of the public sector on litigation outcomes. If our

hypotheses H1.1 and H1.2 are correct, we should observe a positive value for civil cases and a negative value for criminal cases, which indicates a greater likelihood of losing a civil lawsuit and less severe sanctions in criminal cases. The regression coefficient β_2 indicates the effect of the PIPL. A negative value means that the PIPL is associated with a higher chance of a complete win for plaintiffs in civil lawsuits and more favorable treatment for defendants in criminal cases. Z_{it} represents a host of covariates including legal representatives and socioeconomic characteristics of litigants. δ_t and γ_i denote the year and provincial fixed effects, whereas ϵ_{it} refers to the error term. We cluster standard errors at the provincial level.

Additionally, our theory predicts that the effect of the PIPL depends on the identity of defendants. Thus, we introduce an interaction term between *State* and *PIPL* to account for moderating effects [63]. The quantity of interest, β_3 , captures the interactive effects between the two key independent variables. Since there are multiple dependent variables and some of them are categorical, we will use predicted values/probabilities to interpret the interactive effects.

6. Findings

6.1. Descriptive results

How likely are plaintiffs to receive a favorable result in civil privacy disputes? And how are privacy offenders punished in criminal cases? Fig. 1 provides answers to these questions. In civil cases, plaintiffs are most likely to encounter individual perpetrators ($n = 301$), followed by private organizations ($n = 165$) and public organizations ($n = 59$). Regardless of the type of defendants, the probability of achieving a complete win is always the lowest. Overall, plaintiffs fare best when confronting individual defendants, with 55% of such cases ending up with a partial win for them. The biggest challenge for privacy victims comes from public organizations: only 16.9% of cases involving public organizations result in a partial win for plaintiffs, 5.08% in a complete win, and the vast majority—78.02%—in a complete loss. Looking at all civil cases irrespective of defendant type (i.e., the *overall* category on the vertical axis), the overall win rate for plaintiffs is 51.8%, with 44.6% ending in a partial win and 7.24% in a complete win.

Panel (B) of Fig. 1 shows the outcomes in criminal cases. First, the punishment of *control* (or *public surveillance*, 管制 in Chinese)—which is the least severe among criminal sanctions—is very rare in China, with only a single recorded case. Therefore, we exclude it from the subsequent regression analysis. The most common punishment is

imprisonment, which was imposed in 5,092 out of 5,838 criminal cases. Second, probation is widely applied in China. Specifically, it was granted in 45.6% of criminal cases. Third, the average length of imprisonment is 20.75 months, with a very narrow confidence interval owing to the large sample size. However, many perpetrators may not serve in prison due to the widespread adoption of probation, with an average probationary period of 23.6 months. Although these figures help illustrate the criminal sanctions in privacy cases, it remains unclear whether imprisonment or probation is proportionate or sufficient to deter future violations.

Another important feature of criminal cases is the scale of privacy infringements. In our sample of 5,838 criminal cases, defendants are typically charged with infringing upon a large volume of personal information. The median number of compromised data is 20,000, and the mean is significantly higher at 3,068,700. In the most extreme case, the court reported 1.877 billion pieces of information illegally obtained, used or possessed by the defendant. Additionally, our dataset reveals what types of information are most likely to be compromised and how privacy offenders use the compromised information. If interested, please see results in the Appendix on page A-8.

6.2. Regression analysis

6.2.1. Civil lawsuits

The preceding descriptive analysis reveals unequal outcomes in privacy litigation. This section aims to explore the determinants of litigation outcomes. We first estimate a multinomial regression model for civil cases, as the dependent variable includes three possible litigation outcomes. Models (1) and (2) include only the key explanatory variables. Models (3) and (4) include a set of control variables such as gender and age of plaintiffs. Models (5) and (6) serve as the full models with control variables and year/province fixed effects. Throughout all six model specifications, the baseline category of the outcome variable is the complete win for plaintiffs. Thus, the regression coefficients demonstrate how a variable influences the probability of getting a partial win and a complete loss relative to a complete win. As mentioned earlier, the key independent variable is the identity of the defendant, and more specifically, their connection with the state. Defendants are classified into three categories: (1) individuals; (2) private organizations (or affiliated with a private organization); (3) public organizations (or affiliated with a public organization). The regression results are presented in Table 1.

We find a significant effect of the PIPL on civil disputes. Compared to

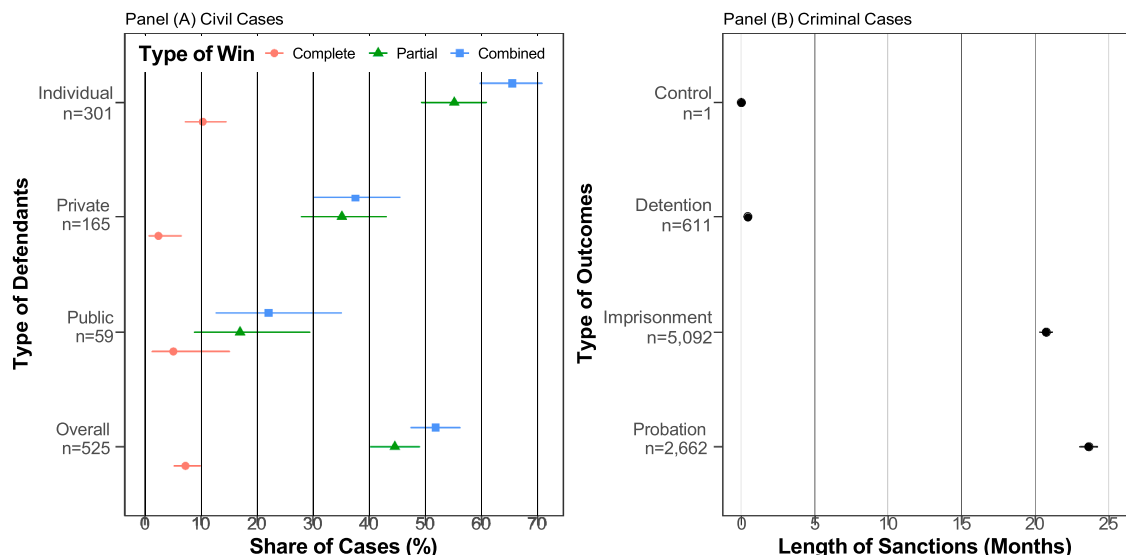


Fig. 1. The outcomes of privacy litigation in China.

Table 1

Multinomial regression of litigation outcome in civil disputes.

	Dependent Variable: Litigation Outcome (0=Complete Win)					
	Partial Win (1)	Complete Loss (2)	Partial Win (3)	Complete Loss (4)	Partial Win (5)	Complete Loss (6)
Private Organization (Individual=0)	1.025* (0.555)	2.062*** (0.552)	−0.110 (0.890)	1.262 (0.885)	1.399 (1.608)	2.625 (1.622)
Public Organization (Individual=0)	−0.481 (0.687)	1.494** (0.631)	18.448*** (0.243)	20.373*** (0.243)	11.108*** (0.288)	13.204*** (0.288)
PIPL Mention	−1.288 (1.269)	−7.797 (17.458)	11.555*** (0.00001)	−5.185*** (0.000)	13.808*** (0.00003)	−6.296*** (0.000)
Plaintiff Age			0.020 (0.022)	0.044** (0.022)	0.054 (0.044)	0.087** (0.044)
Plaintiff Gender (Female=1)			−0.673 (0.645)	−1.545** (0.657)	−1.952 (1.276)	−2.946** (1.280)
Plaintiff Agent (Practitioner=1)			−1.906* (1.015)	−2.601** (1.147)	−1.399 (1.769)	−2.217 (1.878)
Plaintiff Agent (Lawyer=1)			−0.870 (0.701)	−1.131 (0.716)	−0.577 (1.124)	−0.998 (1.130)
Defendant Agent (Practitioner=1)			0.161 (1.487)	−1.423 (1.620)	2.569 (2.686)	0.996 (2.760)
Defendant Agent (Lawyer=1)			1.062 (0.758)	1.538** (0.762)	0.378 (1.403)	0.849 (1.412)
Claimed Amount			0.177*** (0.065)	0.176*** (0.065)	0.249** (0.110)	0.250** (0.110)
Plaintiff Claims			1.445*** (0.369)	1.411*** (0.374)	3.849*** (1.008)	3.782*** (1.009)
Data Type (Manual Collection=1)			−0.241 (0.621)	−0.049 (0.637)	−2.789 (3.095)	−1.452 (3.149)
Year and Region Fixed Effects					Yes	Yes
Constant	1.685*** (0.198)	1.236*** (0.206)	−2.844** (1.356)	−4.109*** (1.383)	−1.362 (2.611)	−2.474 (2.614)
Akaike Inf. Crit.	892.986	892.986	510.048	510.048	563.242	563.242

Note: *p < 0.1; **p < 0.05; ***p < 0.01.

cases where the PIPL was not applied, those influenced by it are more likely to end up with a partial win and less likely to yield a complete loss. Thus, the PIPL appears to help plaintiffs avoid the least desirable outcome in privacy litigation. Additionally, there is compelling evidence that it is much more difficult to sue state-connected actors for privacy infringements. Take model (6) as an example. Suing a defendant affiliated with a public organization is associated with an increase in the log odds of getting a complete loss (versus a complete win) by 13.204. This suggests that when plaintiffs confront state-connected perpetrators, they are almost certain to lose rather than secure a complete win. Furthermore, model (5) suggests that suing a public organization is associated with a very high probability of getting a partial win as opposed to a complete win.

However, suing a private organization does not have a significant impact, suggesting that compared with suing individuals, plaintiffs are equally likely to get a complete win and other possibilities when confronting private organizations. Taken together, our analysis indicates that there is no substantial difference between suing a private organization and an individual for privacy infringements. However, it is significantly more difficult to sue state-connected agents. Plaintiffs in such cases face a very high likelihood of receiving either a complete loss or, at best, a partial win.

In addition to the main findings, the control variables also offer valuable insights into privacy litigation in China. First, as noted earlier, we include age and gender to account for the sociodemographic variation in privacy perceptions and behaviors, digital literacy skills, and engagement with legal remedies. Consistent with expectations, we find that litigation outcomes differ across population subgroups. First, older age is associated with a higher probability of receiving a worse outcome, although the effect is statistically significant only for the outcome of a complete loss relative to a complete win. This finding aligns with the digital divide perspective, and it suggests that older individuals are disadvantaged in courts possibly due to their lack of privacy-related knowledge and digital skills. Gender demonstrates a significant effect

on litigation outcomes in models that examine complete loss versus complete win, with females less likely to receive a complete loss. Again, this finding is consistent with the existing literature on the gender gap in privacy perceptions and behaviors. Female plaintiffs may be better at defending their privacy because they are more concerned about privacy than male plaintiffs in the first place. Additionally, it is likely that they are more familiar with privacy laws and choose to sue only when there is a good chance of winning.

Furthermore, legal representation is not significantly associated with litigation outcomes in the fully specified models, no matter whether the representative acts for the plaintiff or the defendant. This finding is somewhat unexpected because lawyers should generally be better equipped to navigate legal proceedings. One possible explanation is that privacy disputes turn primarily on documentary evidence and the facts established by the court, leaving limited scope for legal representation to shape the outcome. Additionally, we find a positive and significant relationship between the claimed amount and litigation outcome. The more a plaintiff claims, the less likely they are to achieve a complete win. This is understandable though, as courts often reduce the compensation amount to strike a balance between plaintiffs and defendants, which naturally results in outcomes short of a complete win. Finally, the number of claims made by plaintiffs also has a significant relationship with litigation outcome. The more claims a plaintiff makes, the less favorable the litigation outcome will be. This is understandable too, as increasing the number of claims will naturally decrease the probability that the court supports all claims in its final ruling.

6.2.2. Criminal lawsuits

Next, we move on to explore criminal cases, which are fundamentally different from civil cases. First, prosecutors take the place of plaintiffs. Therefore, the focus shifts to the defendants. Second, perpetrators are accused of committing crimes and therefore are subject to criminal punishments. In this analysis, we focus on various types of punishments: detention, imprisonment, and fines. Additionally, we

consider whether a defendant received probation and, if so, the length of the probationary period. A longer probationary period indicates more severe punishment. All time-related variables are measured in months. The fine variable represents an economic value, which is non-negative and right-skewed. Thus, we adopt OLS regression models for all of them. Given the nature of the fine variable, we take its logarithmic form. The regression results are presented in Table 2.

The PIPL appears to have a significant relationship with criminal punishment, particularly with regard to imprisonment and fines. When the PIPL is invoked, defendants receive imprisonment sentences that are approximately six months longer. Moreover, the fine will be 76.65% larger (i.e., $(e^{0.569} - 1) \times 100\%$). However, it remains unclear whether this effect reflects stricter punishment for private actors or for public agencies. The PIPL is also associated with longer probationary periods, but this effect is statistically insignificant. Interestingly, we do not find a significant relationship between the public sector and litigation outcome. Compared with individual defendants, employees of private and public organizations receive similar sanctions under most circumstances. The only coefficient that passes the significance test at the 0.05 level is associated with the private organization in the model of probation. These findings suggest that criminal cases are less influenced by political connections. Public sector employees may be shielded by the state only to the extent that the criminal law is not violated. Otherwise, privacy violators—regardless of their identity—will be subject to similar legal consequences.

Why is there no significant difference in the sanctioning of public sector employees and ordinary individuals in criminal cases? One possible explanation is that public sector affiliation provides less protection once privacy misconduct is treated as a serious criminal offense. At that stage, the employee may be institutionally disowned through suspension, dismissal, removal from office, or Party disciplinary sanctions. Additionally, protecting such employees may also be politically costly because leniency could signal a lack of responsiveness to misconduct that harms citizens. These explanations remain tentative, however, because our data include only published, adjudicated cases and cannot reveal whether institutional affiliation affects investigation, charging, or case publication.

Last but not least, defendants may not be punished immediately after a trial, as they can be granted probation. Probation suspends the execution of a criminal sentence and may ultimately discharge the defendant from further punishment if the defendant demonstrates good

behavior during the probationary period. Among the 5,838 criminal cases, we find that 2,662 (45.6%) were granted probation. Among defendants who received probation, the average probationary period was approximately 23.6 months. Interestingly, defendants affiliated with a private organization face longer probationary periods, which is less favorable to them. In contrast, defendants affiliated with a public organization are granted a slightly longer probationary period on average, although the effect size is minimal and statistically insignificant. Additionally, we find that the criminal history of a defendant has a significant effect on the degree of punishment. Defendants with prior convictions tend to have a longer period of imprisonment (5.51 months longer). Surprisingly, their probationary period is 7.34 months shorter. We speculate that a “selective mechanism” may be at work. Perpetrators with a criminal background are likely to be taken more seriously by the court and, as a result, are significantly less likely to be granted probation compared to those without prior offenses. However, when such individuals are granted probation, it may indicate that their case is particularly minor, and therefore, a shorter probationary period is likely to be imposed.⁷

The above analysis supports our hypothesis that the identity of defendants and the PIPL may influence judicial outcomes in both civil and criminal cases. Compared with individuals not affiliated with any organization, those connected to the public sector are more likely to prevail in civil cases. To our surprise, however, we do not see a similar pattern in criminal cases, as public sector employees receive sanctions similar to those imposed on individual perpetrators. Overall, we find that the PIPL can make a difference in both civil and criminal cases. However, it remains unclear whether this association is concentrated among private actors or public agents. This question is further examined in the next section.

6.3. Does the PIPL matter for all?

6.3.1. Civil cases

Although the PIPL matters for litigation outcomes, it may influence different defendants in different ways or to varying degrees. To explore heterogeneous effects, we interact the PIPL variable with the state connection variable. We are most interested in whether the effect of the PIPL on litigation outcomes varies with the identity of defendants. Despite its importance from the perspective of legal development, the salience of the PIPL in China's judicial practice is surprisingly low. Among the 128 civil cases and 1,267 criminal cases adjudicated by Chinese courts after the PIPL took effect, only three civil cases and ninety-five criminal cases involved the application of the PIPL.

The results from the interactive effects model are visualized in Fig. 2, with a detailed regression table available on page A-9. The three panels on the left display the probabilities of different litigation outcomes when the defendant is not connected with the government, whereas the three panels on the right correspond to state-connected defendants. We find significant interactive effects between state connection and the PIPL. First, when plaintiffs confront private defendants, whether the court invokes the PIPL can make a substantial difference. If the court does not apply the PIPL, the plaintiff is more likely to achieve a partial win ($p \approx 0.729$) than a complete loss ($p \approx 0.271$). If the court invokes the PIPL, however, the plaintiff is virtually guaranteed a partial win ($p \approx 1$), with the probability of a complete loss dropping to nearly zero. This effect is statistically significant, which is indicated by the confidence intervals.

Second, although the PIPL helps plaintiffs confront private actors, it offers little advantage when the defendant is connected with the government. Notably, the PIPL was never invoked in cases involving state-connected defendants. Of the three privacy lawsuits mentioning the PIPL, two involved individuals and one involved a private organization. The PIPL has never been invoked in any of the fifty-nine cases involving state-connected actors. Consequently, there are no predicted probabilities shown on the right side of Fig. 2, as there were no relevant data for such predictions.⁸

Table 2

OLS regressions of litigation outcomes in criminal cases.

	Detention	Imprisonment	Fine	Probation
Private Organization	-0.053	-1.274*	-0.018	3.372***
(Individual=0)	(0.052)	(0.749)	(0.086)	(1.183)
Public Organization	0.054	-1.904	0.248	0.167
(Individual=0)	(0.079)	(1.144)	(0.251)	(1.086)
PIPL Mention	-0.467***	6.122**	0.569**	2.844
	(0.148)	(2.523)	(0.219)	(4.180)
Defendant Agent	-0.868***	9.721***	1.560***	15.421***
(Practitioner=1)	(0.125)	(1.578)	(0.172)	(1.481)
Defendant Agent	-0.393***	11.430***	0.607***	-1.104
(Lawyer=1)	(0.065)	(1.241)	(0.084)	(1.022)
Defendant Age	0.006*	-0.131***	-0.005	0.123*
	(0.003)	(0.040)	(0.005)	(0.063)
Defendant Education	0.064**	-0.052	-0.019	1.617***
	(0.030)	(0.343)	(0.025)	(0.255)
Info Amount	-0.000*	0.012***	0.001	-0.009**
	(0.000)	(0.004)	(0.001)	(0.003)
Criminal Background	0.067	5.514***	0.148*	-7.344***
	(0.048)	(1.070)	(0.076)	(1.394)
Data Type	-0.063	1.477	-0.057	1.111
(Manual Collection=1)	(0.103)	(1.627)	(0.225)	(1.561)
N	2964	2962	2965	2966
Adjusted R2	0.061	0.191	0.088	0.104
Fixed effects	Yes	Yes	Yes	Yes

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

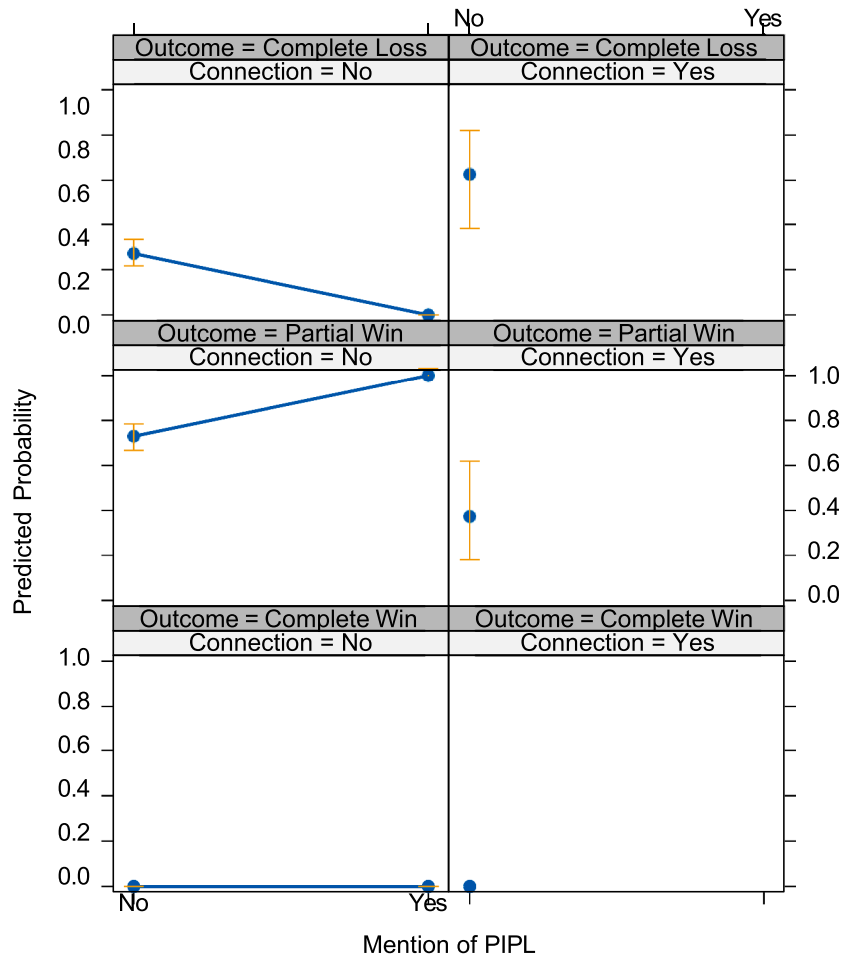


Fig. 2. Examining interactive effects in civil cases.

Among those state-connected cases without PIPL invocation, plaintiffs face a higher probability of receiving a complete loss rather than a partial win (0.627 vs. 0.373), with the probability of getting a complete win close to zero. Thus, when Chinese privacy victims resort to the court to protect their privacy, their chance of success depends largely on whether the offender is a public or private entity. The privacy law may effectively support individuals against private perpetrators, but it becomes less helpful when the perpetrator is the government or is closely affiliated with it. In such cases, courts are unlikely to invoke the PIPL, and the most likely outcome is a complete loss.

6.3.2. Criminal cases

Is the effect of the PIPL in criminal cases also dependent on the defendant? To answer this question, we reexamine criminal lawsuits with an interaction term between state connection and the PIPL. The interactive effects, presented as predicted values of dependent variables, are shown in Fig. 3, and the regression coefficients are available in the Appendix on page A-10.

We do not find significant interactive effects between state connection and the PIPL under most litigation outcomes. Indeed, none of the interaction terms demonstrate statistical significance across the four models. Litigation outcomes are generally similar regardless of whether the defendant is connected with the government or not. Importantly, the PIPL continues to demonstrate a significant impact, as the length of detention tends to go down and the length of imprisonment and probation tends to go up if the PIPL is invoked. When interpreting the results substantively, we find that imprisonment sentences are nearly six months longer and fines are 90% higher when the PIPL is invoked. In contrast, the robustness check in Table A8 with the relaxed coding of

PIPL fails to show a significant effect of the PIPL, thus suggesting that it is the invocation of the PIPL, rather than the timing of trial, that really matters.

These results suggest that unlike civil cases, criminal cases are less prone to the influence of state connection. In fact, the PIPL is more likely to be invoked in criminal cases that involve public entities. Among the 1,267 criminal cases tried after the PIPL took effect, 814 involve individual defendants, 259 involve private organizations, and 194 involve public agencies. The corresponding numbers (percentages in parentheses) are twenty-nine (3.56%), thirty-two (12.36%), and twenty-five (12.89%). Therefore, we speculate that the Chinese government is less tolerant of personal information crimes than of civil disputes, and thus tends to forgo protecting public employees in criminal contexts.

7. Discussion

These findings refine the selective privacy enforcement framework. In civil litigation, plaintiffs fare substantially worse when their claims are directed against public organizations, and the PIPL does not produce the same improvement in outcomes observed in cases involving private defendants. In criminal proceedings, however, defendants affiliated with public organizations receive no statistically detectable advantage over unaffiliated defendants. The evidence therefore provides only partial support for the expectation that political connections uniformly produce favorable judicial treatment. Instead, the consequences of state affiliation appear to depend on whether adjudication imposes liability on a public institution or punishes an individual affiliated with that institution. Our analysis is consistent with the broader literature showing that authoritarian rulers can utilize the judiciary strategically.

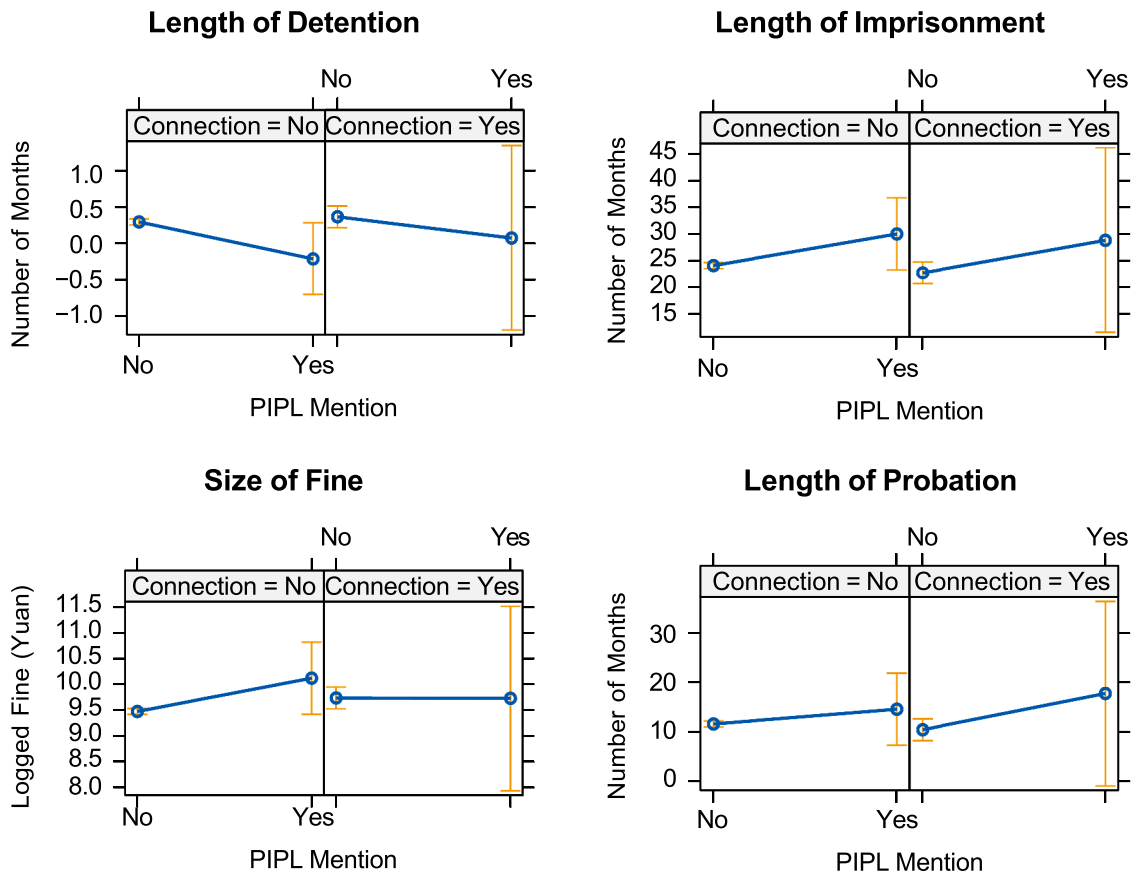


Fig. 3. Examining interactive effects in criminal cases.

It contributes to the literature by showing that uniform protection of state agents may not be the optimal strategy for regime survival.

This civil–criminal divergence suggests a distinction between institutional shielding and agent discipline. In civil cases, an adverse judgment against a public organization imposes liability directly on the institution and may challenge the conduct of a government agency, state-owned enterprise, or other state-connected body. The pattern observed here resembles findings from commercial and administrative litigation, in which institutional position and political connections can influence judicial outcomes. Criminal cases ordinarily involve natural persons rather than public organizations themselves. Punishing a public employee for unauthorized acts, especially those undertaken for personal gain, does not necessarily threaten the institution with which that person was affiliated. It may instead allow the state to characterize the violation as individual misconduct, discipline an errant agent, and demonstrate responsiveness to public concerns. Unconditional protection of state-affiliated offenders, in contrast, might have an adverse effect by undermining mass support, particularly when their conduct compromises the personal information of large numbers of citizens.

Privacy litigation is theoretically revealing because the state occupies several potentially conflicting positions: it creates and enforces data protection rules, processes extensive amounts of personal information, and may itself be involved in privacy violations. The findings indicate that authoritarian legality does not necessarily operate through uniform favoritism toward every state-connected actor. Rather, it may shield public institutions from direct civil liability while permitting the punishment of individual agents whose misconduct can be separated from official policy. This interpretation remains tentative, however. Our data comprise published, adjudicated cases and cannot show whether state affiliation affects earlier stages such as investigation, charging, settlement, or the publication of judgments. Nor can the analysis establish that the civil–criminal difference results from a deliberate state

strategy.

This framework may have implications for other authoritarian systems characterized by politically constrained courts and extensive state control over personal data. Vietnam provides a relevant case for future investigation. Its Personal Data Protection Law, the country's first comprehensive privacy legislation, took effect in 2026. As in China, Vietnamese state institutions are simultaneously data processors, regulators, and enforcement authorities. The framework developed here suggests that enforcement may depend not only on an offender's political affiliation but also on the institutional consequences of imposing liability. Authorities may be more willing to regulate private organizations or prosecute individual public employees than to impose liability directly on public institutions. Vietnam's investigation into the attempted sale of approximately thirty million personal data records allegedly originating from public sector databases illustrates the importance of this question [2], although publicly available sources do not disclose a final investigative outcome. Comparative research should therefore examine whether Vietnam and similar regimes also distinguish between shielding institutions and disciplining individual agents, rather than assuming that state affiliation always provides legal protection.

8. Conclusion

Data are important drivers of digital transformation and governance. While most countries in the world have enacted data protection laws, we still know very little about whether these laws influence the way privacy disputes are resolved. In fact, some legal scholars have doubted the effectiveness of privacy laws in safeguarding personal privacy [64,65]. Drawing on 6,370 privacy lawsuits, this paper evaluates the landscape and outcomes of privacy litigation in China. The most important take-away is that China's data protection law has a significant relationship with litigation outcomes in privacy cases. The PIPL is associated with

outcomes in both civil and criminal cases, but evidence of selective enforcement according to defendant identity emerges principally in civil litigation. In other words, authoritarian privacy laws may be enforced selectively in certain contexts, allowing the state and its agents to be less influenced by the law when involved in privacy disputes. This contrasts China's approach to privacy protection with that of Western democracies such as the United States, where the major challenge for plaintiffs seeking judicial remedies lies in the denial of standing under Article III [28,29].

China's privacy lawsuits reflect a rule-by-law approach to personal data protection [66,67], contrasting sharply with the rule of law approach common in democratic contexts. Privacy protection laws are used as tools to resolve disputes between individuals and private entities, but they are not employed as checks on state power to safeguard individual freedom and the right to privacy. As a result, the public sector may evade punishment when it infringes on personal privacy. This study highlights two major challenges in protecting privacy in authoritarian settings. On the one hand, authoritarian legislatures tend to enact laws that serve the interests of authoritarian rulers but go against the public interest. Legal loopholes in privacy laws can give public agencies a substantial advantage when they are accused of violating personal privacy. On the other hand, the judicial system in authoritarian states often lacks independence. For example, judges are not guaranteed life tenure, and the appointment of judges may depend on political leaders. For these reasons, judges in authoritarian states may refrain from issuing unfavorable rulings against those in power.

While our analysis provides valuable insights into privacy lawsuits in China, two limitations warrant special attention. First, although we find compelling evidence that the public sector and private entities are treated differently in civil lawsuits, the reason why the public sector is favored remains unclear. Multiple mechanisms may be at work. For example, courts might exploit loopholes in privacy laws to justify government conduct. Additionally, judges may protect the public sector due to concerns about its influence on their operations, funding, and appointments. Last but not least, social networks between the public sector and the courts may also contribute to unfair litigation outcomes. Further research could clarify these mechanisms by leveraging additional details from court rulings, such as the legal reasoning, the identities of judges, and the background of public defendants (e.g., their administrative rank).

Second, our analysis cannot establish that the dataset is a representative sample of privacy cases nationwide. Indeed, our analysis of the temporal and geographical distribution of privacy cases shows that many cases are concentrated in certain years (e.g., 2019 and 2020) and regions. Guangdong, for example, uploaded 63 civil cases and 556 criminal cases, significantly higher than the national average of 17 and 188. Therefore, we interpret the findings with caution. However, given the government's tendency to conceal power abuses and misconduct, we believe that underrepresented cases, if included, would likely strengthen the selective enforcement effect observed in our analysis.

9. Endnotes

1. In China, the term privacy is sometimes used interchangeably with personal information or personal data. Jia [37] discusses in detail the definition of privacy in China's context, distinguishing several meanings of privacy and focusing on the information side of privacy.

2. Regardless of the administrative levels and regions, the Standing Committee in China is composed of several leaders with a specific division of labor. The judicial system, including courts and procuratorates, is subject to the oversight of the Political-Legal Committee.

3. This statement is based on the authors' calculations using the litigation dataset compiled for this study. Among the 128 civil cases decided after the PIPL entered into force on 1 November 2021, 3 cases (2.34%) referenced the PIPL. Among the 1,267 criminal cases decided during the same period, 95 cases (7.50%) referenced it. Overall, the PIPL

was referenced in 98 of these 1,395 post-PIPL cases (7.03%). By comparison, the Civil Code was referenced in 149 of the 166 civil cases decided after it entered into force on 1 January 2021 (89.76%), while the Criminal Law was referenced in 5,835 of the 5,838 criminal cases in the full sample (99.95%). The observation periods differ because they correspond to the respective laws' dates of entry into force.

4. We focused on first-instance privacy cases and excluded other types of cases for two reasons. First, only a small fraction of privacy cases proceed to appeal or retrial. For instance, the CJO database shows 9,583 first-trial privacy cases during the period 2010–2024, but only two retrial cases. Second, appeal and other special cases differ significantly from first-trial cases in both the structure and content of their court judgments. To facilitate coding and simplify statistical analysis, it is preferable to focus on first-trial cases.

5. When using GPT-4o for coding, we experimented with coding all variables simultaneously, but the coding accuracy declined substantially compared with coding a single variable at a time. Therefore, the coding was conducted on a variable-by-variable basis to maximize accuracy.

6. We have considered a relaxed way of coding the PIPL. Instead of looking at its application, we distinguished cases based on the passage of the law. We compared the date of trial with 1 November 2021, when the PIPL took effect. Thus, cases tried on or after this date are assigned a value of one, and earlier cases are assigned a value of zero. Results are presented in Table A7 and Table A8.

7. An additional logit regression indicates that defendants with criminal backgrounds are significantly less likely to be granted probation. For details, see the regression results in Table A10 on page A-14.

8. This limitation makes it impossible to cluster standard errors at the provincial level in a multinomial model because of perfect separation. To overcome it, we recoded the dependent variable into a binary variable (lose vs. win) and clustered standard errors at the provincial level. The results remain very similar. For details, please refer to page A-13.

Funding sources

Funding: This work was supported by the Volkswagen Foundation [grant number 9D377].

Declaration of generative AI and AI-Assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used OpenAI API to code 6,370 judicial cases involving privacy infringements. During the revision and resubmission stage, the authors used ChatGPT for language assistance to increase the clarity and readability of the manuscript. After using these tools, the authors reviewed, evaluated, and edited the content as needed and take full responsibility for the content of the published article.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful to David Liu and Xianzhe Yang for their assistance with coding judicial cases. This paper was presented at the FU Privacy-China Workshop at the University of Oxford and at the 2025 Annual Meeting of the American Political Science Association (APSA). The authors thank many participants for their constructive feedback, including Miguel de Figueiredo, Lena Wesemann, Rachel Murphy, Vivienne Shue, Anton Bogs, Danqi Guo, Hao Zhang, Ming Ma, Yang Yan, Yixuan Zhang, and others. Additionally, the authors are grateful to the two reviewers for their constructive feedback.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.clsr.2026.106392](https://doi.org/10.1016/j.clsr.2026.106392).

Data availability

Data will be made available on request.

References

- [1] Noyb. noyb files two complaints against EU parliament over massive data breach. URL, <http://bit.ly/46iC7wg>; 2024. 22 August 2024.
- [2] Voice of Vietnam. The ministry of public security is investigating the sale of 30 million personal data. URL, <http://bit.ly/4nv6dC>; 2022.
- [3] Qin A, Liu J, Chien AC. Chinese police database was left unsecured long before hackers seized it. URL, <http://bit.ly/4l44G59>; 2022. 7 July 2022.
- [4] Reynolds BW. Online routines and identity theft victimization: further expanding routine activity theory beyond direct-contact offenses. *J Res Crime Delinq* 2013;50(2):216–38.
- [5] Jacobus JL, Watson BB. Clapper v. amnesty international and data privacy litigation: is a change to the law 'certainly impending'? *Richmond J Law Technol* 2014;21(1).
- [6] Zuboff S. Big other: surveillance capitalism and the prospects of an information civilization. *J Inf Technol* 2015;30(1):75–89.
- [7] Zuboff S. The age of surveillance capitalism: the fight for a human future at the new frontier of power. London: Profile Books; 2019.
- [8] Consumers International. The state of data protection rules around the world: a briefing for consumer organizations. URL, <http://bit.ly/4l4RXip>; 2018.
- [9] United Nations Conference on Trade and Development. Data protection and privacy legislation worldwide. URL, <http://bit.ly/4l9qRag>; 2025.
- [10] Weber PA, Zhang N, Wu H. A comparative analysis of personal data protection regulations between the EU and China. *Electron Commer Res* 2020;20:565–87.
- [11] Calzada I. Citizens' data privacy in China: the state of the art of the personal information protection law (PIPL). *Smart Cities* 2022;5(3):1129–50.
- [12] Cheng J, Zeng J. Shaping AI's future? china in global AI governance. *J Contemp China* 2023;32(143):794–810.
- [13] Kesari A. Do data breach notification laws reduce medical identity theft? Evidence from consumer complaints data. *J Empir Leg Stud* 2022;19(4):1222–52.
- [14] Birnhack M, Elkin-Koren N. Does law matter online? Empirical evidence on privacy law compliance. *Mich Telecommun Technol Law Rev* 2010;17:337.
- [15] Romanosky S, Hoffman D, Acquisti A. Empirical analysis of data breach litigation. *J Empir Leg Stud* 2014;11(1):74–104.
- [16] Nathan AJ. China's changing of the guard: authoritarian resilience. *J Democr* 2003;14(1):6–17.
- [17] DLA Piper Intelligence. Data protection laws of the world: an overview of key privacy and data protection laws across more than 160 jurisdictions. DLA Piper Intelligence. 2026. URL, <https://www.dlapiperdataprotection.com/>.
- [18] Mokry S, Guro J. Competing ambitions regarding the global governance of artificial intelligence: China, the US, and the EU. *Glob Policy* 2024;15(5):955–68.
- [19] Scoggins SE. Rethinking authoritarian resilience and the coercive apparatus. *Comp Polit* 2021;53(2):309–30. <https://doi.org/10.5129/001041521x15895755803929>.
- [20] Li W, Chen J. From Brussels effect to gravity assists: understanding the evolution of the GDPR-inspired personal information protection law in China. *Comput Law Secur Rev* 2024;54:105994.
- [21] Li Z. Affinity-based algorithmic pricing: a dilemma for EU data protection law. *Comput Law Secur Rev* 2022;46:105705.
- [22] De Hert P, Papakonstantinou V. The data protection regime in China: in-depth analysis. *Depth Anal* (Novemb 19 2015), Bruss Priv Hub Work Pap 2015;1(4).
- [23] Reusch Law. Reuschlaw report: actions for damages under data protection law in 2023. reuschlaw; 2023. URL, <https://www.reuschlaw.de/en/news/reuschlaw-report-actions-for-damages-under-data-protection-law-in-2023/>.
- [24] Goddard M. The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. *Int J Mark Res* 2017;59(6):703–5.
- [25] Kasap GH. Defending privacy across borders: public interest litigation in the fight against data exploitation. *Neth Int Law Rev* 2025;71:427–60.
- [26] Fenske E, Brown CW, Kosseff J. Courting consensus: how class action lawsuits shape data privacy rights and obligations in the U.S. In: Proceedings of the 23rd Workshop on Privacy in the Electronic Society; 2024. p. 172–85.
- [27] Wu FT. How privacy distorted standing law. *DePaul Law Rev* 2017;66:439.
- [28] Park S. Why information security law has been ineffective in addressing security vulnerabilities: evidence from California data breach notifications and relevant court and government records. *Int Rev Law Econ* 2019;58:132–45.
- [29] Wilson N. The aftermath of transunion: an empirical analysis of standing in federal data breach cases 2021–2023. Master's thesis. The University of North Carolina at Chapel Hill; 2024.
- [30] Bradford A. The Brussels effect: how the European Union rules the world. Oxford University Press; 2020.
- [31] Custers B, Dechesne F, Sears AM, Tani T, Van der Hof S. A comparison of data protection legislation and policies across the EU. *Comput Law Secur Rev* 2018;34(2):234–43.
- [32] Monsees C. Privacy litigation: legal recourse and applicable procedural law in data protection matters in Germany. URL, <http://bit.ly/3Gj0EqB>; 2020.
- [33] Villaseñor J. Recording everything: digital storage as an enabler of authoritarian governments. Center for Technology Innovation at BROOKINGS; 2011. URL, <https://bit.ly/44EnjXI>.
- [34] Diamond L. The road to digital unfreedom: the threat of postmodern totalitarianism. *J Democr* 2019;30(1):20–4.
- [35] Wang Z, Yu Q. Privacy trust crisis of personal data in China in the era of big data: the survey and countermeasures. *Comput Law Secur Rev* 2015;31(6):782–92.
- [36] Lee A. Personal data, global effects: China's draft privacy law in the international context. Stanford Cyber Policy Center; 2021. URL, <http://bit.ly/4l4GcsB>.
- [37] Jia M. Authoritarian privacy. *Univ Chic Law Rev* 2024;91(3):733–810.
- [38] Horsley JP. How will China's privacy law apply to the Chinese state?. URL, <https://bit.ly/4lyjxF2>; 2021.
- [39] Fraenkel E. The dual state: a contribution to the theory of dictatorship. New York: Oxford University Press; 2017. Translated from German.
- [40] Hendley K. Varieties of legal dualism: making sense of the role of law in contemporary Russia. *Wis Int Law J* 2011;29:233.
- [41] Hendley K. Legal dualism as a framework for analyzing the role of law under authoritarianism. *Annu Rev Law Soc Sci* 2022;18:211–26.
- [42] Karolewski IPI, Sata R. The dual state of judiciary in Hungary and Poland. *Z Vgl Polit* 2025;19(2):219–46.
- [43] Xu J. The role of corporate political connections in commercial lawsuits: evidence from Chinese courts. *Comp Polit Stud* 2020;53(14):2321–58.
- [44] Wang Y. Relative capture: quasi-experimental evidence from the Chinese judiciary. *Comp Polit Stud* 2018;51(8):1012–41.
- [45] Zhou H, Liu J, He J, Cheng J. Conditional justice: evaluating the judicial centralization reform in China. *J Contemp China* 2021;30(129):434–50.
- [46] Yang Z. The Chinese surveillance state proves that the idea of privacy is more 'malleable' than you'd expect. URL, <https://bit.ly/46kXtt0>; 2022.
- [47] Kostka G, Zhou H. Selective privacy: why and how authoritarian states approach data protection in the digital era. In: Reveron DS, Polatty DP, Marion S, editors. *The Oxford Handbook of Human Security*. New York: Oxford University Press; 2026. p. 561–76. chapter 35.
- [48] Ye M. The limits of judicial reforms: how and why China failed to centralize its court system. *China Q* 2023;255:753–67.
- [49] Zhang L, Liu F, Yuan X. Career mobility and performance evaluation of China's SOE managerial elites. *China Q* 2024;259:765–84.
- [50] Liu Y. Exclusion of notification and consent rules for state organs in the processing of personal information. *Shanghai Law Stud* 2021;6:123–33.
- [51] Clarke D. Law and the political system in Xi Jinping's China: the decline of the party-state distinction. In: Lynch D, Rosen S, editors. *Chinese Politics: The Xi Jinping Difference*. Routledge; 2024. p. 17–43.
- [52] Minzner CF. China's turn against law. *Am J Comp Law* 2011;59(4):935–84.
- [53] Zhang Y, Liu JZ. The rise of the Chinese judiciary and its limits: administrative litigation in the reform period. *China Q* 2025;262:330–51.
- [54] Chen L, Liu Z, Tang Y. Judicial transparency as judicial centralization: mass publicity of court decisions in China. *J Contemp China* 2022;31(137):726–39.
- [55] Xu Z, Yan S, Pan Z. Empirical analysis of sentencing for crimes of infringing citizens' personal information based on SPSS. In: 2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI). IEEE; 2021. p. 379–84.
- [56] Zhang J, Li Ke, Feng Y. Criminal sanctions on identity theft in Shanghai: an empirical case law analysis. *Int J Law Crime Justice* 2022;71:100562.
- [57] Liebman B, Stern R, Wu X, Roberts M. Rolling back transparency in China's courts. *Columbia Law Rev* 2023;123(8):2407–82.
- [58] Ferrara A. A practitioner's guide to using large language models and generative AI in economic history. Technical report. National Bureau of Economic Research; 2026.
- [59] Hoy MG, Milne G. Gender differences in privacy-related measures for young adult Facebook users. *J Interact Advert* 2010;10(2):28–45.
- [60] Tifferet S. Gender differences in privacy tendencies on social network sites: a meta-analysis. *Comput Hum Behav* 2019;93:1–12.
- [61] Antonio A, Tuffley D. Bridging the age-based digital divide. *Int J Digit Lit Digit Competence* 2015;6(3):1–15.
- [62] Whiting SH, Ma X. Validating vignette designs with real-world data: a study of legal mobilization in response to land grievances in rural China. *China Q* 2021;246:586–601.
- [63] Brambor T, Clark WR, Golder M. Understanding interaction models: improving empirical analyses. *Polit Anal* 2006;14(1):63–82.
- [64] Waldman AE. Privacy law's false promise. *Wash Univ Law Rev* 2019;97:773.
- [65] Don S. Litigation takes the stage: using litigation to solve performances in privacy law. *Hastings Sci Technol Law J* 2023;14(2):287–310.
- [66] Peerenboom R. Varieties of rule of law: an introduction and provisional conclusion. In: Peerenboom R, editor. *Asian discourses of rule of law*. 1st edition London: Routledge; 2004. p. 1–53. chapter 1.
- [67] Ohnesorge JKM. The rule of law. *Annu Rev Law Soc Sci* 2007;3(1):99–114.