Plea Leniency and Prosecution Centredness in China’s Criminal Process

Xin He

Faculty of Law, University of Hong Kong, Hong Kong, China
Email: xfhe@hku.hk

(First published online 29 March 2023)

Abstract

China’s criminal proceedings have been recognized as being “investigation centred.” I argue that the rise of the Plea Leniency System has led to “prosecution centredness.” Analysis of the operation and consequences of plea leniency shows how the procuratorate has overshadowed the police and further marginalized the courts. In plea leniency, the defendant has little chance of being acquitted and the legal profession provides little defence. While this paradigm shift signals further leniency in criminal justice, rights protections make way for efficiency and crime control. As such, plea leniency has profound implications for the operation of the criminal justice apparatuses, defendants, defence lawyers, and the mode of crime control in China.

Keywords: plea leniency; prosecution-centredness; China’s criminal process

Until recently, China’s criminal proceedings have been recognized as being “investigation centred” (zhengcha zhongxin zhuyi 侦查中心主义). A popular saying captures the relationship between the three criminal apparatuses: the police cook the rice, the prosecutors deliver the rice, and the judges eat the rice. In this sequential yet collective process, the cook – the police – take centre stage. Once the rice is cooked, judges will eat all of it. Although constraints occur at both the prosecution phase and the trial phase, they have little effect on the course towards conviction. A guilty verdict is all but inevitable.

Realizing the grave consequences of investigation centredness – the miscarriage of justice being the most conspicuous – the state has called for a shift to “trial centredness” (shenpan zhongxin zhuyi 以审判为中心). In an historic decision in 2014 (Plenum Decision hereafter), China’s Communist Party (CCP) promoted the concept of trial centredness in criminal justice. Academics may debate

1 Trevaskes 2010a; McConville et al. 2011; Mou 2020; Miao 2021; Chen, Ruihua 2017a.
2 Fu 2003.
3 CCP 2014.

© The Author(s), 2023. Published by Cambridge University Press on behalf of SOAS University of London. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/S030574102300022X Published online by Cambridge University Press
what this means, but the general consensus is that the courts shall play a prominent role and evidence other than oral confession shall be key to conviction.4

However, almost a decade after the call, trial centredness has yet to take shape; instead, in a different form, the role of the confession has gained new momentum. Criminal trial reforms and the official discourse have been dominated by the plea leniency system (renzui renfa congkuan zhidu 认罪认罚从宽制度, literally, “leniency for acknowledging guilt and accepting punishment”). In a drive for efficiency, plea leniency has been rolled out at breakneck speed. Less than two years after a 2016 pilot programme, it was enacted into the Criminal Procedural Law (CPL hereafter).5 As an overarching principle of the “mini constitution,” it has since pervaded the whole criminal process. It is used in an array of crimes, from minor personal injury, to property damage, drug trafficking, rape and murder. Moreover, it is applied at all phases of the process – during the police investigation, the public prosecution and the court trial. Two years into its enactment, the plea leniency system was implemented in 86.8 per cent of all criminal cases.6 During this time period, the duration for handling cases shortened and the appeal rate dropped precipitously.

The plea leniency system is dominated by the prosecutors: they not only obtain a guilty plea from the suspects/defendants but they also ensure that the suspects/defendants accept their specific sentence. The prosecutors also play a central role in liaising with lawyers in facilitating this consent. Once all the details regarding the crime and the punishment have been agreed, the trial is nothing more than a procedural formality: the court merely rubber stamps the confession affidavit.

Few studies in the English-speaking world have regarded the introduction of the plea leniency system as paradigm shifting. In Jeremy Daum’s comprehensive introduction, he only suggests that the reforms “have the potential to undermine the already limited protections afforded to China’s criminal defendants and risk encouraging false confessions.”7 Yuguang Lu argues that with plea leniency, China’s criminal procedure has “entered a new stage.”8 Yet, following his detailed comparisons between plea leniency and plea bargaining, he does not regard the change as systemic. Similarly, Yuhao Wu focuses on measuring the duration of the process and the amount of leniency but pays little attention to the changing roles of the three apparatuses.9 Enshen Li concentrates on the role of the mandatory lawyers10 and whether the system can boost legitimacy.11

By contrast, papers in Chinese are replete with discussions of this topic, apparently because it is politically hot. For obvious reasons, leading authorities in China’s criminal process unanimously hail the system.12 Fan Chongyang 樊崇义 calls it “a historic transition in China’s criminal process.”13 Hu Yunteng 胡云腾, a former vice-president of the Supreme People’s Court (SPC), asserts that the system will benefit all litigation participants.14 A Xinhua report acclaims it to be “monumental.”15 As one can imagine, these supportive statements are more propaganda than analysis. Most Chinese studies embrace the central authorities’ position. Some do suggest that the procuratorate has become dominant or has shouldered the bulk of the responsibility in the criminal process; however, these commentators often adopt a doctrinal or comparative perspective, without offering actual evidence.16

---

4 Chen, Weidong 2016.
5 NPC 2018.
6 Jiao 2021.
7 Daum 2021.
8 Lu 2021, 69.
9 Wu 2020a; 2020b.
10 Li, Enshen 2022b.
11 Li, Enshen 2022a.
13 Ibid.
14 Ibid.
16 Wan 2019; Yan 2020; Xiong 2019.
This article analyses the developments, operation and consequences of the plea leniency system. I argue that the rise of plea leniency has led to “prosecution centredness” in China’s criminal process. The police’s power has declined while the procuratorate’s power has expanded considerably. The police may still dominate the investigatory period but the procuratorate has acquired some power here and clearly dominates the post-investigatory period. This change has further marginalized the role of the courts. I will also demonstrate that plea leniency fundamentally differs from plea bargaining. While it signals further leniency in criminal justice, rights protections are sacrificed for efficiency and crime control. Plea leniency has profound implications for the operations of the criminal justice apparatuses, defendants, defence lawyers and the mode of crime control in China.

Investigation Centredness and Its Discontents

Criminal justice in China is conducted by three state apparatuses: the police, the procuratorate and the courts. The sequential process is often depicted as a relay race, or a conveyor belt, that turns criminal suspects into convicted prisoners. The CPL defines the relationship between the three as a “division of responsibility, mutual coordination, mutual constraint” (分工负责，互相配合，互相制约). Nicknamed the “iron triangle,” there is more coordination than constraint underlying this relationship. Rather than form a triangular framework of checks and balances, the three apparatuses collaborate to ensure criminal convictions. There are few mechanisms available that allow them to “check” one another. Never are the courts an independent institution in fact finding or decision making.

Traditionally, of the three apparatuses, the police played the dominant role. While they were once responsible only for the preliminary investigation, they gradually gained authority over the entire criminal justice regime. The police had enormous formal and informal power to deprive suspects of personal liberty for a significant amount of time and to dispose of the relevant property. Only in some formal situations, such as arrests, did the police need the procuratorate’s endorsement, which was given in at least 80 per cent of cases. Despite China’s commitment in international treaties, in no situation was habeas corpus needed.

The evidence collected by the police also has an overwhelming impact on court decisions. The criminal trial has long centred on dossiers. The courts are there to confirm the written evidence prepared by the police. Although the CPL stipulates conditions under which illegally obtained evidence must be excluded, the conditions are too difficult to prove in most cases. To exclude evidence, defendants or their lawyers have to prove that it was obtained by force or torture. For defence lawyers not participating in the investigation process, this is a formidable undertaking. Even if there is a chance that the prosecutors might adjust the charges and fix some procedural or evidential defects, the courts invariably confirm the predetermined guilty outcomes. The minimum sentence invariably is no less than the detention period defendants receive, so as to prevent police mistakes or state compensation. When there are procedural or evidential defects, the procuratorate and the court are not supposed to mitigate or drop the charges, but rather they should fix them. Police investigation – the starting point of the criminal process – is often regarded as its endpoint.

One consequence of investigation centredness is the lack of rights protection for defendants. Most notably, wrongful convictions or miscarriages of justice have haunted China’s criminal justice
system: in some extreme cases, the actual murderer has emerged years after the legal “murderer” was executed, or the “murdered” victim has returned from the grave.24 Investigation centredness is widely regarded as a leading cause of such cases.25 In an empirical analysis, Lena Zhong and Mengliang Dai find that the top three reasons for wrongful convictions are all related to police conduct.26 Accordingly, the CCP has prohibited the extraction of confessions through torture during the investigation stage of the criminal trial.27 Evidence obtained through torture or other illegal means must be excluded. For the first time, the CCP has specified that instances of torture or fabrication should be punished. Judges, prosecutors and police will be held accountable for life for negligent or wrongful acts that result in wrongful convictions.

Furthermore, the Party has set the goal of achieving trial centredness. Trial centredness is not court centredness; it means the evidence is at the centre of the criminal trial. It stresses that the facts and evidence of cases shall stand the test of law. Court hearings, instead of the dossiers, shall be the central link in the fact-finding process.28 Additionally, evidence presented by the police and the prosecution shall be determined to be factual only after it has been tested in court.

**The Plea Leniency System**

Although trial centredness targeted investigation centredness, it would have further diminished the procuratorate’s influence within the criminal justice process. The Supreme People’s Procuratorate (SPP hereafter) thus has tried to maintain its dominant position.29 The plea leniency system, although only mentioned in passing in the Party’s (2014) Plenum Decision, has since been heavily promoted as a means to improve efficiency, social stability and over-crowded detention centres, which are unable to detain minor offenders for too long.

The plea leniency system effectively states that if a defendant acknowledges guilt and accepts the proposed sentence (either a specific sentence or a sentencing range), he or she will in return be treated leniently. Acknowledging guilt means that the defendant/suspect confesses to the offence honestly and does not dispute the charges. Accepting sentence means that the defendant feels contrition and is willing to accept the requisite penalty. The litmus test is the attitude and behaviour of the defendant: is the apology sincere and the compensation to the victim adequate?30

The plea leniency system originates from the “expedited procedure” for minor criminal cases. This procedure applies to defendants who are facing less than three years of imprisonment. In such cases, the court investigation and deliberation are curtailed, leaving only the announcement of judgment.31 The plea leniency system addresses concerns about efficiency. The SPP argues that given China’s increasing criminal caseloads, complicated and straightforward cases should follow different channels. Only by expediently processing minor cases can adequate resources be allocated to deal with serious and complicated ones. This argument was endorsed by Meng Jianzhu 孟建柱, the-then head of the Central Political and Legal Committee: “Leniency in substance and simplification in procedure will better allocate resources and provide both leniency and rights protection. It reduces confrontation and facilitates the confessions of both the suspects and defendants, making them cooperate with the judicial apparatuses in lawfully handling the cases.”32 In 2016, a pilot scheme was rolled out in 18 cities.33

---

24 He, Jiahong 2016.
26 Zhong and Dai 2019.
27 CCP 2013.
29 Yu 2021.
30 SPC, SPP, MPS, MSS and MoJ 2019, Arts. 6–7.
31 NPC 2018, Art. 224.
32 Xing 2016.
33 NPC 2016.
Is the concern for efficiency in criminal cases a real and pressing one? While the courts’ overall caseloads have increased, the criminal caseload has plateaued at around one million annually since the late 2010s.\textsuperscript{34} In addition, as Zuo Weimin points out, the handling of criminal cases is already efficient thanks to the launch of the expedited procedure.\textsuperscript{35} There is little room for further truncation. According to him, under the expedited procedure, an average court trial lasts only 5 minutes, 90 per cent are conducted in under 10 minutes, and it takes on average 6.4 days for the court to process a case.\textsuperscript{36}

Other scholars have questioned whether plea leniency can protect human rights.\textsuperscript{37} Some have been concerned about whether plea leniency is compatible with trial centredness.\textsuperscript{38}

All of these concerns have been set aside, however. On the eve of the amendment of the 2018 CPL, the SPP stated in an interim report that the pilot scheme had achieved “stunning” success: “Until October 2018, about 50 per cent of the criminal cases had adopted the scheme in the pilot cities. The prosecution duration was shortened to 26 days, 70 per cent of [prosecutions] were processed through the expedited procedure, under which 95 per cent of the court decisions were handed out immediately after the trial.”\textsuperscript{39}

Supported by such figures, the scheme was officially incorporated into the CPL (2018). However, the plea leniency system is more than just an alternative route. Listed as one of the general provisions, it has become a fundamental principle in criminal proceedings.\textsuperscript{40} It is applied throughout all phases of the criminal process and covers all kinds of crimes. In 2019, the SPP, the Supreme People’s Court (SPC hereafter) and three ministries jointly issued detailed regulations on its implementation (Guiding Opinions hereafter).\textsuperscript{41}

Traditionally, China’s primary policy in criminal justice has been “leniency for those who confess; severity for those who resist” (\textit{tanbai congkuan, kangu congyan} 坦白从宽，抗拒从严). However, this vague policy is not always enough to induce suspects to confess: many do not trust that the authorities will honour promises of leniency. In many cases, the policy has been regarded as a trick used to force confession and self-incrimination. According to a cynical saying, “one stays imprisoned forever once one confesses” (\textit{tanbai congkuan, laodi zuochuan} 坦白从宽，牢底坐穿). An empirical study by Wang Fang and Guo Liang, which was based on more than 6,000 judgments in intentional injury cases, shows that confessions obtained under the policy did not lead to a more lenient sentence, because “confession is a legally obligated duty.”\textsuperscript{42} The plea leniency system is meant to overcome this mistrust.\textsuperscript{43} Under the system, the type and length of sentence are usually specified in the confession affidavit. In addition, the Guiding Opinions (2019) stipulate that under the plea leniency system, a defendant who confesses should receive more leniency than one who only admits guilt.\textsuperscript{44} In practice, leniency terms can reduce the original sentence by about 10 to 30 per cent.\textsuperscript{45} The promises of the prosecutor, with specific sentencing terms written into the confession affidavit and witnessed by defence lawyers, are more credible.

To address concerns over rights protection, China has made it mandatory for every defendant to have access to either a professional or legal aid lawyer (the lawyer-for-all scheme).\textsuperscript{46} Authorities will

\begin{itemize}
\item \textsuperscript{35} Zuo 2017.
\item \textsuperscript{36} Ibid., 168.
\item \textsuperscript{37} Chen, Ruihua 2017b.
\item \textsuperscript{38} Gu and Xiao 2017.
\item \textsuperscript{39} Zheng 2020.
\item \textsuperscript{40} NPC 2018, Art. 15.
\item \textsuperscript{41} SPC, SPP, MPS, MSS and MoJ 2019.
\item \textsuperscript{42} Fang and Liang 2019.
\item \textsuperscript{43} Zuo 2017.
\item \textsuperscript{44} SPC, SPP, MPS, MSS and MoJ 2019.
\item \textsuperscript{45} Daum 2018.
\item \textsuperscript{46} SPC and MoJ 2018.
\end{itemize}
assign a duty lawyer, from among those who are on standby in all detention centres, to suspects and defendants who do not hire their own lawyer. When prosecutors process a case, they shall consult these lawyers. The CPL also clarifies that, *inter alia*, the authorities shall allow these lawyers to review dossiers and meet suspects or defendants. This is to ensure that when the suspects or defendants opt for plea leniency, they understand its nature, processes and consequences.

It is amazing how widely and promptly this system has been applied. In January 2019, 20.9 per cent of all criminal cases were processed under the system. By 2020, that rate had skyrocketed to 86.8 per cent; from January 2020 to November 2021, 85 per cent of criminal cases followed this route. By 2021, the courts confirmed 96.85 per cent of the sentences proposed by the procuratorate. The appeal rate for these cases was 3.5 per cent, 20.51 per cent lower than before the system was implemented. The system has been touted by Zhang Jun, the SPP president, as a “modernized governance approach to control crime with Chinese characteristics.”

**Prosecution Centredness**

In its traditional role as the middleman between the police and the courts, the procuratorate already had considerable power over those two branches. The plea leniency system allows the procuratorate to grab more power from both the police and the courts. This change can be evinced in its relationships with all participants of the criminal justice system – the police, courts, defendants, defence lawyers and victims.

**Overshadowing the police**

Plea leniency has given the procuratorate more power than the police in criminal proceedings. Although plea leniency applies to all three phases in the criminal justice system, leniency offers can only be made at the prosecutorial phase. This shift has taken away some of the police’s “teeth.” The police can invoke plea leniency or educate suspects on how to plead, but the Guiding Opinions expressly prohibit the police from granting leniency or recommending sentencing. If the suspects do plead, the police can only record the plea in the dossiers and transfer them over to the procuratorate. Only after the prosecutors step in can offers of leniency be made.

In addition, plea leniency allows prosecutors to constrain the police’s investigatory power. Traditionally, prosecutors intervened only once the police had finished investigation, interrogation and gathering evidence: when the rice had already been cooked. Plea leniency offers the prosecutors another reason to determine whether to approve the arrest recommendation proposed by the police or not. Moreover, the prosecutors intervene earlier, usually immediately following the arrest or detention of a suspect. Once the prosecutors step in, “the police no longer need to do further work gathering substantial evidence.” From that point onward, collecting evidence is the prosecutors’ job: they become not just one of the cooks, they become head chef. Chen Guoqing, a vice-president of the SPP, has stated that the procuratorate now has part of the investigatory responsibility.

Although the police still take the lead during the investigation phase, the plea leniency system gives the procuratorate a greater say. “During the review period for the approval of arrest or on

---

50 Zheng 2020.
51 SPC, SPP, MPS, MSS and Moj 2019, Art. 23.
52 NPC 2018, Art. 81.2.
53 Lu 2021, 61–62.
54 Ibid., 62.
major cases, the police shall carefully listen to the opinions and suggestions of the procuratorate with regard to plea leniency and conduct relevant work accordingly. This is the opposite of the traditional process in which the procuratorate follows the lead of the police and helps to fix any defects. Now, both the police and procuratorate have the power to interrogate defendants/suspects.

The procuratorate has enormous discretion over matters during the post-investigatory phase – far more than the police. Prosecutors can rephrase, redact and frame the elements of the crime, as well as communicate with the judges, defence lawyers and victims. They have the power to initiate charges, determine detention periods, conditions and, more importantly, whether or not there is any culpability. They determine whether to press or waive charges, and they can recommend sentences. Although the police may have initiated the criminal proceedings, the prosecutors complete them. The police have almost no say over prosecution or adjudication outcomes. According to Zhaohua Yan, the police’s authority to initiate the process or withdraw cases is under regulatory scrutiny. “They cannot determine a case’s trajectory and outcome under plea leniency.” However, the prosecutors’ power to grant leniency is more flexible. No longer the middleman between the police and the courts, the prosecutors are now the dominant actor.

In 2019, the SPP president officially stated that the procuratorate should bear “the dominant responsibility” in the criminal process. In other words, the procuratorate now has the dominant “role” and the plea leniency system has been cited as evidence of this change. Why have the police given up so much power? First, a more regulated police power is necessary in order to reduce wrongful convictions. Xi Jinping has attempted to use cases of wrongful convictions as evidence of the need to consolidate control over the justice agencies. Second, the police are no longer responsible for collecting evidence once the prosecutors step in; this reduces the chances of the police making mistakes. Arguably, it is rare for a wrongful conviction to occur under the plea leniency system as the defendants have voluntarily pleaded guilty and accepted their punishment.

Marginalizing the courts

The procuratorate is not shy of asserting its dominance over the courts. Chen Guoqing, vice-president of the SPP, openly states that the procuratorate has replaced the courts in assuming “much” of the judicial responsibility. Traditionally, the prosecutors would make a decision on whether to bring charges. Now, under plea leniency, they determine whether or not there is culpability, the type of crime and the exact sentence. Thus, prosecutors have become the “judges.” The courts have been further marginalized under the plea leniency system. Traditionally, the courts reserved the power to make the final decision. Under plea leniency, a judge only verifies the voluntary nature, authenticity and legality of the guilty plea. Is there violence or duress involved? Are the defendants mentally healthy? Is there effective communication between the defendant and the procuratorate? Yet, the prosecutors have already done this before the case comes to court. The CPL stipulates that the courts shall adopt the recommendations of the procuratorate except under five circumstances. All of the circumstances are difficult to prove – for example,

---

56 SPC, SPP, MPS, MSS and MoJ 2019, Art. 24.3.
57 Ibid., Art 8.2.
58 Yan 2020, 42.
60 Zhang 2019, 9.
61 Nesossi 2017.
63 SPC, SPP, MPS, MSS and MoJ 2019, Art. 39.
64 NPC 2018, Art. 201.
if the defendant has not voluntarily admitted guilt; it is hard to prove voluntariness if the guilty plea has been made under the threat of a harsher sentence. Even according to the US Supreme Court, an action is generally regarded as voluntary only if one understands the pros and cons of a choice.65 Changyong Sun and Wenjun Tian’s empirical study found that the trial is little more than a formality.66 Although only 23.73 per cent of judges regarded the sentences recommended by the procuratorate to be correct, 95 per cent followed the recommendations.67 Judges tend to adopt recommendations to maintain a good relationship with the prosecutors and to avoid protests or appeals later on. Moreover, Enshen Li finds that lawyers do not even appear in almost two-thirds of trials.68

The most crucial component of a trial is the evidence. Under normal circumstances, few witnesses would appear in court.69 However, under plea leniency, there are even fewer witnesses. Since the prosecutors have reached an agreement with defendants and their lawyers, any negotiation between the lawyers and witnesses has already been incorporated into the confession affidavit. There is no point in calling more witnesses to the stand. The prosecutors’ burden of proof is exempted.70 All other evidence has been sifted, scrutinized, arranged or rephrased by the prosecutors and there is no need to challenge the prosecutorial recommendations.

With the procuratorate encroaching on the courts’ power, one might expect resistance from the courts. Yet this is not the case. Plea leniency conforms to, if not maximizes, the institutional interests of the courts in several ways: it reduces the numbers of appeals by defendants and protests by the procuratorate; it reduces the length of trial; most decisions can now be handed out immediately after the trial, a goal the courts have long aimed for; some leniency pleas even lay out the means of implementation; and the number of closed cases for individual judges has tripled, or even quadrupled.71

While the criminal caseloads remain stable, there has been an explosion in civil and enforcement dockets. The average caseload per judge has been increasing ever since the judicial reforms in 2014, and particularly following the “judges’ quota” reforms, which have nearly halved the number of judges and which mean that only the court staff under the judge’s track can adjudicate cases.72 Plea leniency not only delivers efficient relief from an overbearing workload but it also sets aside the most complicated or organized crimes in which some of the defendants refuse to plead in exchange for leniency through the regular trial channel.

Plea leniency also spares judges the risk of appeal; by eradicating the possibility of either factual or legal errors in the proceedings, judges become exempt from life-long responsibility.73 Since 2014, judges have been held accountable for their mistakes for life. Plea leniency provides an ideal rescue from that risk. Plea leniency also appears to promote the legitimacy of the system as a whole. Thus, the courts are happy to cooperate with the procuratorate in implementing it. Empirical studies have shown that judges even urge defendants to confess and to plead for leniency.74 As long as neither the victims nor their families complain, both the judges and the prosecutors are happy. In this situation, both win.

In most cases, the courts merely rubber-stamp whatever the procuratorate has presented – they eat all the rice cooked up by the procuratorate. They have no reason to conduct a fully-fledged cross

65 *Brady vs The United States*, 397 US 742, 1970.
66 Sun and Tian 2021.
67 Ibid., 11.
68 Li, Enshen 2022a, 89.
69 McConville et al. 2011, 246; Mou 2020, 193; Guo 2021, 194–96.
70 Sun 2018, 181.
71 He, Xin 2021a; Ng and He 2017a.
72 He, Xin 2021b.
73 Lu 2021, 61.
74 Chen, Ruihua 2017b.
examination of witnesses or to seriously question any evidence. Excluding illegally or improperly obtained evidence and altering suggested sentences has become increasingly rare. Under normal circumstances, the trial time had been short; under plea leniency, it has become even shorter – the judge simply ticks the boxes. Trials have become more administrative than judicial; their judicial power has been stripped, or usurped, by the procuratorate.

**Overwhelming the defendants**

With the plea leniency target rate set at 70 per cent by the SPP in 2019, it is natural that the procuratorate pressures defendants to plead guilty in exchange for leniency. The capability imbalance between defendants and prosecutors epitomizes the contrast between “repeat players” and “one-shotters.” Following Marc Galanter’s analysis, the prosecutors are repeat players, with repeated involvement in similar cases, and the defendants/suspects are one-shotters – usually first-timers in the process. One-shotters are unable to compete against repeat players in terms of knowledge and resources.

Armed with the power to grant leniency, the procuratorate can exert immense pressure on the defendants. This is a versatile tool. In addition to the discretion to reduce sentences by 10–30 per cent, the prosecutors can grant probation, bail or even withdraw the charge; they can increase or shorten the length of detention; they can schedule an earlier or later trial date. Non-pleading defendants are more likely to be left waiting in custody at a detention centre with little idea of how long the process will take. Pleading defendants may get bail or be tried within a few days. Their attitude, including any hesitation, reluctance or eagerness to admit guilt, makes a difference. The pressure on the defendants to plead guilty in exchange for leniency is huge and most can do little to resist.

The procuratorate can also withdraw any leniency which has been granted. If defendants regret and reverse their original confession during the trial, the confession affidavit will be revoked, and the case will follow the regular route. Thus, most defendants are reluctant to reverse their plea. Whether they are guilty or not is not the issue: there is little to gain from reversing a plea but much to lose.

In a guiding case issued by the SPP, the suspect was found to have committed the offence of breaking, entering and stealing. He had pleaded guilty and had accepted the proposed sentence of 25 months in jail, but he then appealed, citing that the penalty was too heavy. This sparked a protest from the local procuratorate which argued that “any defendant who provides neither new evidence nor new facts reneges on his commitment to the confession and the proposed sentence. The appeal is solely based on the fact that ‘the sentence was harsh,’ without proper legal basis. This move wastes the state’s resources and the original condition of leniency has vanished.”

Under this pressure, the appeal court remanded the case to the original basic-level court, revoked the original 25-month sentence and handed out a new sentence of 31 months.

**Controlling defence lawyers**

Even with the support of defence lawyers, the imbalances between the defendant and the prosecutors persist. First, the skills of the lawyers, especially those of the public defenders, pale in comparison to those of the prosecutors. Duty or legal aid lawyers are not usually highly skilled and many are inexperienced novices. They are selected and salaried by local justice bureaus. Privately hired
lawyers might not be specialists in all areas of criminal law. The prosecutors, who regularly deal with criminal charges, understand the law inside out, while these lawyers may be unfamiliar with the particular sections of the criminal law and process.

Second, and more importantly, is the fact that lawyers lack independence and professionalism. The relationship between Chinese lawyers and state officials is “symbiotic.”81 Privately hired lawyers rely on state officials for case referrals and favourable decisions, as do duty and legal aid lawyers. Although there is a small number of rights defence lawyers, most are pragmatic deal brokers. Originally, judges were in a better position to control lawyers, since they monopolized the decision-making power and controlled which cases were brought to trial. Under plea leniency, it is the prosecutors who have this control and who make decisions that are nearly impossible to appeal. The lawyers have to rely on the prosecutors to make a living. Any belligerent, confrontational or uncooperative behaviour can be detrimental to their interests or even sabotage their careers. Cooperating or showing kindness, on the other hand, can earn them reduced charges or lesser sentences for their clients.82

This tendency is also illustrated in the SPP’s guiding cases, which must be followed by prosecutors in similar cases across the country. Guiding cases are intended to exemplify plea leniency’s superiority in terms of fairness and rights protection. However, one of the guiding cases does not even mention the presence of lawyers,83 possibly because there were none involved.84 In another case, the role of the defence lawyer was limited to that of a notary public witnessing the hearing.85 In another, the lawyer helped to convince the defendant to accept the deal proposed by the prosecution.86

Only in one case did the defence lawyer play a minimal role.87 The defendant had fatally injured his friend. The prosecutors had consulted the defence lawyer on whether the defendant had voluntarily turned himself in (zishou 自首), whether there was any evidence of self-defence and whether the victim’s family had forgiven the defendant. The defence lawyer suggested that the defendant had acted in self-defence. The procuratorate then explained that the evidence showed that the defendant, after being hit on his head with a wine glass, had taken a knife from the kitchen to stab the victim with: “This is an instigated response, instead of self-defence.”88

The legal analysis in the guiding case is not problematic. The point, instead, is that the prosecutors had dominated the entire proceedings. They physically controlled the evidence; the defence lawyer had nothing new to add. The guiding case also showed that the lawyer had immediately accepted the procuratorate’s narrative: “after communicating with, and seeking the opinion of the lawyer on duty, the prosecutor, under the presence of the lawyer on duty, explicated the detailed elements and the legal basis of the sentences.”89 The lawyer cast no doubt on whether any evidence should have been excluded, or whether there had been any alternative interpretation. Nor did he suggest any alternative opinion on the proposed sentence.

Traditionally, there were “three difficulties” for criminal lawyers.90 Under the plea leniency system, two have been resolved – the difficulties in accessing case files and in meeting suspects/defendants. However, the symbiotic relationship has changed little. Indeed, allowing access to case files and defendants is meaningless as both are controlled by the procuratorate. When prosecutors have only a few weeks to complete a plea leniency case, they put pressure on the defence lawyers. The lawyers may have access to the dossiers and suspects, but they are often reminded to be expedient:

81 Liu 2011.
82 Li, Enshen 2022b.
83 SPP 2020, No. 81–84.
84 Ibid., No. 84.
85 Ibid., No. 81.
86 Ibid., No. 83.
87 Ibid., No. 82.
88 Ibid.
89 Ibid.
90 Li, Yuwen 2014; Liu and Halliday 2016.
the defendants have already been informed about their rights. In many cases, the confession affidavit is prepared before the lawyers are even notified. Many lawyers help the procuratorate to convince or urge the defendants to plead guilty. Thus, instead of actively defending clients, lawyers become “explainers,” “persuaders” or even just “observers.” According to another survey, 89 per cent of lawyers admitted that their role was confined to seeking “lenient sentencing outcomes.” Li finds that in 99.67 per cent of his surveyed cases, the lawyers raised no objection to the prosecution’s case. Some lawyers are there only to witness the “signing ceremony,” having been downgraded to a notary public. Their presence becomes the best evidence that the admission of guilt has been “voluntary.” In most situations, their presence is ceremonial and symbolic, to fulfil procedural requirements. They work more for the procuratorate than for the defendants.

Does legal representation have any effect on case outcomes? Yuhao Wu found that despite an increase in lawyer representation for defendants in plea leniency cases from less than 10 per cent to nearly 50 per cent, the overall case outcomes did not change. The positive effect of court-appointed lawyers even disappeared after the lawyers-for-all programme was implemented.

**Pacifying victims**

Granting leniency to defendants may aggrieve their victims, who not only want to see the murderers, rapists, thieves, burglars, etc. duly punished by the law but also, in many circumstances, want financial compensation. This is similar to criminal reconciliation in which the defendant may get leniency if he financially compensates the victim and is forgiven. Usually, more leniency is justified with more compensation. Thus, pacifying the victims has become part of the procuratorate’s job.

The plea leniency system equips prosecutors with more resources. The prosecutorial relief fund, a scheme to help victims who do not receive proper compensation from defendants, tripled in size from 2018 to 2021, the period during which plea leniency was established.

In the above-mentioned case of the man who fatally wounded his friend, the prosecutor learned that the compensation the defendant had given the victim’s family was insufficient to cover the family’s losses. The prosecutor then paid several visits to the family, offering not just condolences but also judicial remedies. The family promised to take the matter no further.

Here, the prosecutor not only demonstrated care but also provided the victim’s family with funds. By showering goodwill and/or material benefits, he effectively ended the dispute. Ensuring that the defendant does not appeal is one matter; ensuring that the victim’s family does not complain or petition is another. Thus, the judicial relief fund under plea leniency further empowers the procuratorate.

**Differences between Plea Bargaining and Plea Leniency**

Plea bargaining means that, through negotiation between the prosecution and the defendant, usually via their lawyers, the prosecutors may “drop some charges or accept a plea to a lesser crime” if the defendant admits to others. This has facilitated the criminal trial process. Once a deal is made,

---

91 Li, Enshen 2022b.
92 Ibid.
93 Hu 2018, 115, quoted in Li, Enshen 2022a, 89.
94 Li, Enshen 2022a, 89–90.
95 Wu 2020c.
96 Ng and He 2017b.
98 SPP 2020, No. 82.
99 Hessick 2021, 23; see also Fisher 2000, 864.
a jury or judge trial is unnecessary. Thus, it has contributed to a reduction in trials in many jurisdictions. In the federal courts of the United States, for example, the plea bargaining rate was more than 97 per cent in 2018. Explaining the “triumph” of plea bargaining, George Fisher argues that prosecutors have the strongest incentives and power to make it happen, and judges “see it as in their interests.” Legislators then allocate more sentencing authority to prosecutors, which “tilt[s] the terms of battle.” Defendants, although armed with a “nominal” right not to plead, are hopelessly undefended.

The development paths of plea bargaining and plea leniency are strikingly similar. Two differences are crucial, however. First, the reference point for each is different. Defendants who choose not to bargain under the common law system are subject to a jury trial or, in rarer instances, a judge trial. Going to trial is a gamble: one might win big or lose big; plea bargaining is only a means to hedge the bets. Nonetheless, acquittal is likely. For example, in the US, juries acquit one-third to 16 per cent of defendants. Moreover, these rates do not count cases in which the prosecution drops the charges. In other words, if the defendant is determined to pursue a jury trial, acquittal is likely. For the prosecutors, going to trial also means that the jury could “spurn their painstakingly assembled cases.” Thus, defendants, with their lawyers behind them, may wage a meaningful battle against the prosecutors.

In contrast, conviction rates in criminal trials in China are extremely high, even with the recent emphasis on trial centredness. Official statistics state that of the 1.1 million criminal cases closed in 2020, only 1,040 produced a not-guilty verdict. Of those 1,040 cases, 384, or almost 40 per cent, were private prosecutions. These were for minor crimes that did not need the involvement of the procuratorate. In other words, it is extremely rare, if not impossible, for a defendant to be acquitted or released after being charged by the procuratorate. To not plead for leniency is neither a gamble nor a hedge; it is destined to failure.

In plea bargaining, if a judge detects any duress, ignorance, the absence of a defence lawyer or any complaint from the victim, the deal will be rejected. The court process allows the defendants another opportunity in the “arraignment.” In contrast, even if Chinese judges find evidential or procedural discrepancies in a trial, they will not free the defendant. In China, an acquittal means that the procuratorate has blundered. Owing to the mutual supervisory relationship between the courts and the procuratorate, both institutions have to tolerate each other’s shortcomings and find a way to resolve or avoid undesirable outcomes. Yu Mou describes how the procuratorate, upon learning that a court is about to issue an acquittal, would propose an exchange: to drop some potential protest by the court in other cases, in order to replace the acquittal with a conviction. Both institutions understand that they should maintain a cooperative relationship.

Another related difference between plea bargaining and plea leniency is that plea bargaining is preconditioned on professional and independent lawyering. While defence lawyers may not be able to truly represent their clients’ interests in plea bargaining, as shown in Albert Alschuler’s classic study, their mere presence makes defendants less likely to plead guilty. Fisher demonstrates that in Massachusetts in the 18th century, when poor defendants had more legal representation, plea

100 Hessick 2021, 32.
102 Friedman 1998.
103 Hessick 2021, Ch. 3.
104 Friedman 1998, 192.
105 Fisher 2000, 865.
107 Chen, Ruohui 2017b, 38.
109 Mou 2020, 183.
bargaining rates dropped precipitously. He calls this “the power of representation.”  

In plea leniency, however, defence lawyers side with the prosecution in persuading, if not coercing, the defendants to admit guilt and to accept the proposed sentence; they are not in a position to negotiate with the prosecution. They are unwilling, or unable, to find new evidence that proves the defendant’s innocence. In rare circumstances, they may find faults in the evidence prepared by the police or other prosecutorial recommendations. Nonetheless, they cannot reverse the inevitable course towards conviction. If negotiation exists at all, there is no balance of power between the two sides.

In sum, plea leniency is dominated by the procuratorate. Indeed, the very wording of plea leniency distinguishes it from plea bargaining: there is no bargaining; it is a “plea for leniency.” The defendant is to earn leniency with an admission of guilt, a show of contrition and by offering evidence for conviction. The officials may, or may not, show mercy. Wu shows that no leniency is offered in felony charges, regardless of the plea, only one-sided pleading instead of two-way bargaining. The procuratorate does not negotiate – it is the arbitrator or, to put it bluntly, the “judge.” It grants leniency according to its own discretion set by the law. The confession affidavit with sentencing recommendations is no contract: it is a pending judgment for punishment. If justice is negotiated in plea bargaining, it is imposed and coerced in plea leniency. Sun and Tian find that the prosecutor’s recommended sentences are a Hobson’s choice. The “negotiation” time given to duty lawyers never exceeds ten minutes. In essence, plea leniency gives discounted sentences for an early guilty plea. Unlike plea bargaining, no charges are dropped or altered. Despite the CCP’s efforts to diminish the role of confession in the criminal process, the rise of plea leniency suggests that confession remains central to the system.

Consequences

The rise of the plea leniency system is likely to have an immense impact on the operation of China’s criminal justice. According to the official rhetoric, both social stability and efficiency have been achieved with shortened trials and lower appeal rates. However, there are consequences beyond the official rhetoric. The system may be efficient, but this efficiency comes at the expense of defendants’ rights.

Since the procuratorate has become the dominant player among the three apparatuses, it enjoys enormous and almost unconstrained discretion. Under the current arrangement, the courts have lost any meaningful supervision over it. If the prosecutor’s constrained discretion can become a source of unfairness in plea bargaining in the common-law world, the procuratorate’s unrestrained discretion in plea leniency can only be worse. The procuratorate can easily strike an agreement with the defendants by manipulating the evidence. For example, prosecutors might choose to disclose only unfavourable evidence to urge defendants to confess and accept the suggested sentence. Neither the CPL nor the Guiding Opinions stipulate mechanisms to guarantee full disclosure let alone any penalty to dissuade prosecutors from evading full disclosure. The procuratorate can intentionally, or subconsciously, conceal or tamper with any evidence unfavourable for the defendants. It can also manipulate timings: for example, favourable evidence might only come to light after a plea deal has been struck. On the other hand, prosecutors can also drop charges, in

112 Wu 2020a.
113 Baldwin and McConville 1979.
114 Sun and Tian 2021, 5.
115 Ibid., 9.
117 Belkin 2013.
118 Hessick 2021, Ch. 3.
119 Liang and He 2014.
part or in full, in respect of the suspects’ “meritorious service,” a step which is largely subject to their interpretation. This can terminate the criminal proceedings, even before the courts have become involved.

The plea leniency system also creates incentives for false confessions. There are no mechanisms to verify whether the acknowledgment of guilt is accurate or reliable in cases in which the defendant has confessed; the defendants can end up taking the blame for others. Owing to information asymmetry and the harsh conditions imposed in detention and interrogation, admissions of guilt may not be a deliberate decision over the strength of evidence.

As with the defence lawyers, the courts can be even less useful in challenging the police or prosecutorial defects since the rise of plea leniency. Traditionally, if the evidence was impaired, the courts would ask the procuratorate to fix it. Similarly, the procuratorate also required the police to ensure the evidence was sound. The mutual supervisory relationship between the three apparatuses, although skewed as already shown, offered defendants a certain degree of protection from the abuse of power by officials. However, under the plea leniency system, proposed criminal charges are rarely rejected and the incentives for further investigation have declined. As long as the defendant admits guilt and accepts the proposed sentence, there is no reason to present further evidence. Confession carries more weight than any other type of evidence and determines the case outcome. Both the police and the procuratorate will only become more unscrupulous and spend less time gathering evidence.

Furthermore, without a fully developed factual record, when an innocent person is wrongfully convicted, it is hard to determine what went wrong. Neither the police, the prosecutors, the judges, defence lawyers or anyone else in the criminal justice system will take the blame for the error, which will be squarely placed upon the defendant who “voluntarily” confessed.

Indeed, plea leniency has effectively abolished the appeal system. In the guiding case mentioned above, the defendant was sentenced to an extra six months for appealing after taking plea leniency, a 24 per cent increase in jail time. It is an entrenched legal principle in Chinese criminal procedural law that no appeal shall incur an added penalty. This is to safeguard the defendant’s right to appeal. The plea leniency system has undermined this principle. Arguably, the system may still be compatible with the principle: if the original sentence is a result of lenient treatment, no “penalty” is added if the leniency term is taken away. However, in reality, it acts as a deterrent to any defendant who has confessed. For them, the first trial has to be the final trial. To be fair, there is also a trial penalty in plea bargaining. However, as shown, when a conviction is imminent in China’s criminal process, few defendants appeal. This is why, across the country, the appeal rate has fallen to as low as 2.3 per cent. In Yuguang Lu’s study of 388 plea leniency cases, only four were appealed and none of the decisions were reversed in appeal. Thus, the legal principle and the defendants’ rights have taken a back seat to concerns over efficiency and stability. Indeed, as long as confession remains a core value of the system, appeals may be regarded as dispensable.

Plea leniency marks a further type of leniency in China’s criminal justice policy. Ever since Xiao Yang’s tenure as president of the SPC, “balancing severity and leniency” has been moving towards “swift and severity.” The thrust is to treat serious crimes harshly while treating minor crimes leniently. The trend has been towards leniency, except in cases of crimes against state interests – for example, those involving rights lawyers, terrorists and separatists. Execution rates have plummeted. With plea leniency extending to nearly 90 per cent of all criminal cases, the trend is

120 SPC, SPP, MPS, MSS and MoJ2019, Part III.
121 Biddulph, Nesossi and Trevaskes 2017, 110.
122 Lu 2021, 61–62.
123 Lu 2021.
124 Trevaskes 2010b.
125 Liebman 2015; Li, Yuwen 2014.
126 Johnson and Miao 2016.
clear. Confession in exchange for leniency has long been part of China’s criminal justice policy, but the plea leniency system has taken this one step further. The Guiding Opinions stipulate “additional” reductions in sentencing beyond what is covered under the “confession for leniency policy.” The move towards leniency manifests not just in substantive laws but also in criminal procedures. Originally, criminal reconciliation was the major vehicle towards leniency.\textsuperscript{127} Now that reconciliation has joined forces with plea leniency, the trend has become more systemic. Legally speaking, criminal reconciliation is only supposed to be applied in minor cases, but plea leniency applies in all crimes.\textsuperscript{128} Criminal reconciliation has been driven partially by private forces, but plea leniency derives solely from the might of state power. In contrast to the “swift and severe” policy from before Xiao Yang’s time in office, one may even argue that the criminal justice policy has evolved into “swift and lenient.”

\textbf{Conclusions and Implications}

Although China’s criminal justice process can still be described as a conveyor belt that turns suspects into convicts, its emphasis has changed. Previously, the police were at the centre of criminal proceedings. With the rise of plea leniency, the procuratorate, by determining most of the final outcomes, has taken over. It is the only apparatus that has tentacles in all investigatory, prosecutorial and judicial phases. Prosecutors become an arbitrator before the judge. In this process, “the police take notes, the prosecutors make recommendations, and the judges review the recommendations.”\textsuperscript{129} Needless to say, making recommendations is prioritized when nearly 100 per cent of them are accepted. Indeed, prosecutors can also recommend sentences for cases outside the plea leniency system.\textsuperscript{130} The courts have been further marginalized. Trial centredness, which has never been substantiated, has been superseded by prosecution centredness; the sequential process has just been condensed. In addition to cooking, prosecutors now also submit “Michelin” star reviews. Judges have lost the opportunity to eat; they are now the “editors,” if not the “readers,” of the reviews.\textsuperscript{131}

The ideology underlying China’s criminal justice system had long been to locate the “objective truth,” with officials “expected to bring their knowledge to bear in reaching the unique factual truth.”\textsuperscript{132} Under plea leniency, this has become “consented truth.” As long as the procuratorate and the defendant reach a consensus, the process is deemed to be over. This ideological change will affect how China defines a “wrongful conviction,” since the consensus may cover up miscarriages of justice. As prosecution centredness replaces investigation centredness, the role of the confession is only strengthened: the roots of the miscarriage of justice have yet been eradicated.

The perceived triumph of plea leniency in breaking the narrow hold of minor crimes in an expedited procedure and conquering the whole penal territory within a few years shows that of the two functions of criminal justice, China cares more about crime control than rights protection.\textsuperscript{133} While both fairness and efficiency are listed as the justice system’s paramount goals, efficiency has been favoured and speed is prioritized. The regime’s traditional emphasis on crime control and legitimacy leaves little room for procedural justice and rights protection; the presumption of innocence remains a principle on paper only.

It seems that the regime has found a way to both efficiently process criminal cases and maintain its legitimacy. With plea leniency taking hold, wrongful convictions are harder to detect and petitions and appeals are fewer. Rights protection has been sacrificed in exchange for leniency. Citizens are punished, with little resistance. Plea leniency has turned China’s criminal justice system into a

\textsuperscript{127} Liebman 2015.
\textsuperscript{128} NPC 2018, Art. 288.
\textsuperscript{129} Yan 2020.
\textsuperscript{130} NPC 2018, Art. 282.
\textsuperscript{131} Lu 2021, 85.
\textsuperscript{132} Mou 2020, 13.
\textsuperscript{133} Packer 1964.
ruthlessly efficient machine and the judicial process into a mere administrative process. Moreover, this administrative process has no judicial remedy. The plea leniency system has brought a sea change in the criminal justice system.

Acknowledgements. I am grateful for the insightful, detailed and considerate comments of two anonymous reviewers. Yuechan Song provided superb research assistance.

Competing interests. None.

References


Chen, Ruihua. 2017b. “Renzui renfa congkuan zhidu de ruogan zhengyi wenti” (Several issues on the system of leniency on admission of guilt and acceptance of punishment). Zhongguo faxue 1, 35–52.


Xin HE professor in the Law Faculty of the University of Hong Kong. He studies China’s legal systems from a socio-legal perspective. His monograph, Divorce in China: Institutional Constraints and Gendered Outcomes, was published by New York University Press in 2021.