

# Ethnic Discrimination in Criminal Sentencing in China\*

[This version: November 5, 2019]

Yue Hou<sup>†</sup> and Rory Truex<sup>‡</sup>

## Abstract

This paper presents the first analysis of ethnic discrimination in sentencing patterns in the People's Republic of China, focusing on drug cases in Yunnan province. We posit the “problem minority” hypothesis, which holds that discrimination in an authoritarian system emerges when an ethnic group becomes associated with behavior that generates social instability. On average, minority defendants in Yunnan have sentences that are about 1.5 to 8.0 months longer than Han defendants that have committed similar drug crimes. Further analysis of data from all provinces reveals that this bias is largest for groups heavily involved in the drug trade. Sentencing is particularly harsh in autonomous counties and prefectures where minorities are concentrated.

Keywords: authoritarian; ethnicity; China; Yunnan; drugs; judicial politics; courts; discrimination

Word Count: 9685

---

\*Our gratitude goes to Chris Callison-Burch, Rafaela Dancygier, Hanming Fang, Guy Grossman, Nahomi Ichino, Dorothy Kronick, Chengdong Liang, Benjamin Liebman, Nicholas Sambanis, Jeffrey Staton, Rachel Stern, Pablo Querubin, Yuhua Wang, and Yiqing Xu for helpful feedback at various stages of the project. Several talented graduate and undergraduate students provided excellent research assistance. All analysis and replication files will be made available on the authors' personal websites upon publication of the manuscript. Any remaining errors are our own.

<sup>†</sup>Assistant Professor of Political Science, University of Pennsylvania, yuehou@sas.upenn.edu

<sup>‡</sup>Assistant Professor of Politics and International Affairs, Princeton University, rtruex@princeton.edu.

## 1 Introduction

Across many developed democracies, the scales of justice are tipped against members of ethnic and religious minority groups (Alesina and La Ferrara 2014; Grossman et al. 2015; Mustard 2001; Rehavi and Starr 2014; Steffensmeier, Painter-Davis and Ulmer 2017; Ulmer and Johnson 2004). African American defendants in the United States receive sentences that are about ten percent longer than white defendants (Rehavi and Starr 2014). Racial discrimination proves largest in areas where the black population is most concentrated (Alesina and La Ferrara 2014; Ulmer and Johnson 2004), consistent with the broader idea of “group threat” theory (Ulmer and Johnson 2004). Criminal justice systems are institutions of social control, and racial biases in sentencing arise when the majority group feels threatened, either physically or economically, by a minority population (King and Wheelock 2007). This leads to stereotyping and prejudice among jury members, prosecutors, and judges, who tend to be members of the dominant group (Abrams, Bertrand and Mullainathan 2012; Anwar, Bayer and Hjalmarsson 2012; Lim, Silveira and Synder 2016, Rehavi and Starr 2014; Starr and Rehavi 2013).

Our paper extends this line of research with the first analysis of ethnic discrimination in criminal sentencing outcomes in the People’s Republic of China (Liebman 2015; Lu and Miethe 2002). Under what conditions will we observe ethnic discrimination in criminal sentencing in an authoritarian system? Our theory takes the dictator’s stability imperative as its starting point—authoritarian regimes have an overarching interest in maintaining social order (Gandhi and Przeworski 2007; Gandhi 2008; Svobik 2012). The day-to-day business of stability maintenance is delegated to lower level officials (Hassan 2017; Liebman 2014; Minzner 2011; Wang and Minzner 2015; Yang 2017), who have substantial discretion in making and implementing policy (Lieberthal 1992). In the criminal justice system, judges use the law to promote “social harmony” through the sentencing process (Liebman 2014, 2015; Li 2015; Minzner 2011).

We term our overarching intuition the “problem minority” hypothesis, which builds on the ideas of group threat theory (King and Wheelock 2007; Ulmer and Johnson 2004). In contrast to “model minorities”—those that have largely assimilated and are viewed in good standing by

---

the ethnic majority– “problem minorities” are involved in some behavior that has the ability to disturb public order (Porter and Washington 1993). The group itself becomes associated with a distinct social malady, and members may experience stigma and discrimination, even if they are not personally involved (Blair, Judd and Chapleau 2004; Rehavi and Starr 2014).

The implication of this view is that in general, we expect to observe ethnic discrimination in authoritarian courts, but that discrimination should be heterogenous across groups, issues, and space. When minorities are concentrated, and when they are disproportionately involved in an illicit activity that threatens stability, we expect the discriminatory logic to take hold. Harsher sentences will be meted out to members of the “problem minority” for those types of crimes. When these conditions are not met, discrimination should be nonexistent or less severe.

The paper tests this idea through an analysis of criminal sentencing patterns in drug cases in Yunnan, a province in southwestern China of about 50 million people. Because Yunnan borders Myanmar, Laos, and Vietnam, it is considered the epicenter of China’s drug problem, as large quantities of heroin, methamphetamine, and opium flow across the region (Yang 1993; Zhang and Li 2010). Some minority groups in Yunnan have kin populations in neighboring countries, and ethnic ties facilitate drug trafficking and the emergence of cartels (Personal Interview B002; Personal Interview B008). We consider this a “most likely” case for the theory (Levy 2008)– a well-defined social stability problem, associated with minority populations, and offenses that are amenable to quantification and comparison. If we do not observe ethnic bias in Yunnan on sentences for drug offenses, the theory is likely wrong.

We attempt to measure discrimination through a conditioning approach, in line with many of the foundational studies in this literature (Rehavi and Starr 2014; Mustard 2001; Ulmer and Johnson 2004). Because ethnicity is not randomly assigned, it is impossible to definitively assert that being an ethnic minority has a causal effect on sentencing outcomes (Sen and Wasow 2016). We adopt the more modest aim of a.) first identifying whether there is indeed an association between ethnicity and sentence severity and b.) seeing whether that association still exists after accounting for a rich set of covariates that should affect sentences according to China’s Criminal

---

Law. Any “unexplained variation” is suggestive of ethnic discrimination in the system, though it is always possible that some unobserved confounder or sample selection issue could explain a discrepancy across ethnic groups.

Our analysis of the roughly 10,000 drug cases in the data from Yunnan (2014-2018) reveals a degree of ethnic discrimination. Minority defendants, on average, have sentences that are about 1.5 to 8 months longer than Han (the majority group) defendants that have committed similar crimes. The median sentence for all defendants in the Yunnan data is 48 months. Though the point estimates vary slightly, this general result holds across a range of conditioning strategies and covariate sets, including the incorporation of court fixed effects.

The data offers insights into possible mechanisms underlying this relationship. The courts in Yunnan appear to vary in their severity, and ethnic minorities are more likely to be tried in harsher courts by virtue of the geography of their crimes (Ulmer and Johnson 2004). In the Chinese system, certain administrative units (provinces, prefectures, and counties) with a concentration of ethnic minorities are termed “autonomous” and have special constitutional obligations to protect minority culture. The province of Yunnan is not considered an “autonomous region” in the Chinese system, but it does have 29 autonomous counties and 8 autonomous prefectures assigned to 18 different minority groups. Our analysis shows that courts in autonomous localities are actually systematically harsher—defendants tried in these areas tend to receive sentences that are about 3 months longer than those tried in regular courts. This affects both Han and minority defendants, but the fact that minorities are more likely to live and work in autonomous counties/prefectures means that they are differentially subjected to these more punitive courts.

We also find that ethnic minorities are less likely to have legal representation, which may affect the quality of their defense (Lu and Miethe 2002). This suggests that the sentencing discrepancies may be in part driven by economic inequality, as minorities in Yunnan generally come from less developed areas and have less access to legal resources (Personal Interview B009). But it is important to emphasize that defense lawyers in China rarely take an adversarial ap-

---

proach in the courtroom or attempt to prove the innocence of their clients (Liebman 2015; Lu and Miethe 2002). Instead, their role is to facilitate justice on behalf of the state, and it is empirically unclear whether their presence (or a more combative approach) improves outcomes for their clients (Liebman 2015; Lu and Miethe 2002). This is an area that can benefit from additional research.

We test the generalizability of our Yunnan findings across the rest of the country by applying our parsing algorithm to all publicly available drug cases in China, 167,091 cases in total. We have less confidence in this data, but nevertheless it allows for additional testing of the problem minority framework, as we can disaggregate the effects by province and group. This analysis reveals variation in discrimination across groups and jurisdictions (Ulmer and Johnson 2004). Groups that are more heavily involved in the drug trade face differentially harsher sentences in the criminal justice system— the number of cases per capita or drugs confiscated per capita strongly predict a groups' bias coefficient. Conversely, members of minority groups that are not highly involved in the drug trade do not appear to have harsher sentences than their Han counterparts that committed comparable crimes, with some exceptions. This suggests that the discriminatory logic is not simply about minority status per se, but about the reputation of one's group as it relates to the broader drug issue.

## 2 Theory

Under what conditions do authoritarian regimes discriminate against ethnic minorities in the criminal justice system? Majority groups in government generally allow some minorities to assimilate while suppressing others (Cederman, Wimmer and Min 2010). Concentrated minority groups are seen as particularly threatening because they are more likely to mobilize and engage in civil conflict (Horowitz 1985; King and Wheelock 2007; Morelli and Rohner 2015).

Our argument builds on the idea of societal threat and requires some assumptions about the way governance works in places like China and other decentralized authoritarian systems (King and Wheelock 2007; Ulmer and Johnson 2004). First, we assume that the overarching goal of

---

the regime is to stay in power and maintain control over society. This assumption motivates many of the workhorse models of authoritarian politics (Gandhi and Przeworski 2007; Gehlbach, Sonin and Svolik 2016; Svolik 2012). This is not to say that officials in authoritarian systems do not have policy preferences or ideological commitments, but simply that there is a stability imperative that underpins much of the political calculus.

The second assumption is that the everyday business of maintaining stability is delegated to lower level officials, who may have some discretion in how they make and implement policy (Hassan 2017). This “fragmented authoritarianism” is a core feature of the Chinese political system, and it is responsible for local variation in policymaking across a range of issues (Lieberthal 1992). With respect to the court system, judges may work in tandem with local officials to administer justice in a way they perceive will best foster social stability in that jurisdiction (Personal Interview B001; Personal Interview B003; Personal Interview B008). The law itself is an institution of control (Liebman 2015; Minzner 2011; Ng and He 2017).

We expect to observe discrimination in criminal courts against minorities in general, but heterogeneity is a central part of the argument. Our core intuition is that ethnic discrimination in authoritarian systems arises when certain groups come to be viewed as threats to social order. Some minority groups may be disproportionately involved in activities the government comes to internalize as destabilizing or unruly—separatism, violence, terrorism, drug trafficking, and so forth. Members of these “problem groups” will experience harsher sentences.

Our core hypothesis is as follows:

H<sub>1</sub>: Compared to ethnic majority group members that have committed equivalent crimes, ethnic minority defendants will receive harsher sentences when the a.) the crime is associated with the group itself and b.) the crime is part of a larger threat to social stability.

We also expect to observe geographic variation in discrimination:

H<sub>2</sub>: Discrimination in criminal sentencing against ethnic minorities will occur in areas where a.) minorities are concentrated; b.) the prevalence of the crime is acute; and c.) minorities are differentially involved in the crime.

---

In areas where minorities are sparse, or where the crime/issue is not particularly relevant, we would not expect the “problem minority” logic to dominate the mindset of local officials. Discrimination is only likely to occur in localities that have the intersection of a large minority population involved in a destabilizing activity.

### 3 Background

Our analysis focuses on criminal cases involving drug trafficking, transport, smuggling, and possession in China’s Yunnan province. We focus on this case because it offers a clear test of the theory, and because the quality of the Yunnan drug data is relatively good. Below we provide some brief context on the nature of criminal justice, the drug problem, and ethnicity in contemporary China.

#### 3.1 China’s Criminal Justice System

The Chinese legal system has four levels: the Supreme People’s Court (SPC), provincial high courts, intermediate courts (municipal-level), and basic courts (county and district-level). The SPC promulgates judicial interpretations and exercises “bureaucratic, ministry-like” authority over the lower judiciary (deLisle 2014, p. 226).

Courts and other legal institutions operate explicitly under the leadership of the Chinese Communist Party (CCP). The CCP selects and promotes judges and court personnel, and local Party and government officials influence local courts through the CCP’s Political-Legal Committee (政法委) (Ng and He 2017, p. 19). Chinese courts work alongside other government agencies in maintaining social stability (Ng and He 2017; Minzner 2011; Personal Interview B006).

The role of defense attorneys in China’s legal system remains limited (Personal Interview B001). Defendants have the right to counsel, though many do not seek representation. According to Lu and Miethe (2002), defense lawyers in China have “dual responsibilities of not only having to protect the rights of the accused, but also to help the state to seek the truth” (p. 269). In

---

2018, 99.994% of defendants tried for criminal offenses were found guilty.<sup>1</sup>

### 3.2 Ethnicity in China

China is a multi-ethnic state dominated by a single ethnic group—the Han—who represent 92 percent of the total population. The state officially recognizes 55 ethnic minority groups, with a total population of about 110 million.

The CCP has followed the Soviet model of providing limited concessions to minority groups in the form of regional autonomy and policy privileges. The Chinese political system grants “autonomous status” to provinces, prefectures and counties with large minority populations. The CCP still has political authority over these territories, but ethnic groups are granted a variety of rights, first enshrined in the 1954 PRC Constitution and further elaborated in the Law on Regional Ethnic Autonomy.<sup>2</sup> As of 2005, more than 71 percent of China’s ethnic minority population lived within one of the more-than-1300 autonomous units (Cornell 2002; Leibold 2013).

Beyond the system of regional ethnic autonomy, the CCP provides a series of benefits to all ethnic minority citizens. These policies include preferential access to employment, higher education, and public offices; exception from family planning restrictions; and special tax breaks (Distelhorst and Hou 2014; Leibold 2013). In recent years, preferential ethnic policy has faced increasing resentment from the Han population, and Chinese society appears to be witnessing the rise of “Han chauvinism” (Leibold 2010). Many Han Chinese elites consider their own group to be more economically and culturally advanced, and some ascribe to a paternalistic mission to “help” impoverished minority groups to embrace modernity (Han and Mylonas 2014; Personal Interview B002).

---

<sup>1</sup>See data from the SPC work report and the Chinese law year book, cited in <http://dy.163.com/v2/article/detail/EA8BH6340514CC4Q.html>, last accessed Oct. 7, 2019.

<sup>2</sup>The Law on Regional Ethnic Autonomy stipulates that courts and procuratorates in autonomous regions are required to hire staff from local ethnic groups. Courts are mandated to use the language commonly used in the locality in the prosecution and trial of cases, and cases should be assigned to individuals familiar with the languages of minority ethnicities. Courts are also required to hire translators for minority defendants and plaintiffs when necessary. These provisions aim to promote ethnic diversity by safeguarding the rights of minority groups and preserving their languages and customs.



### 3.3 China's Drug Problem

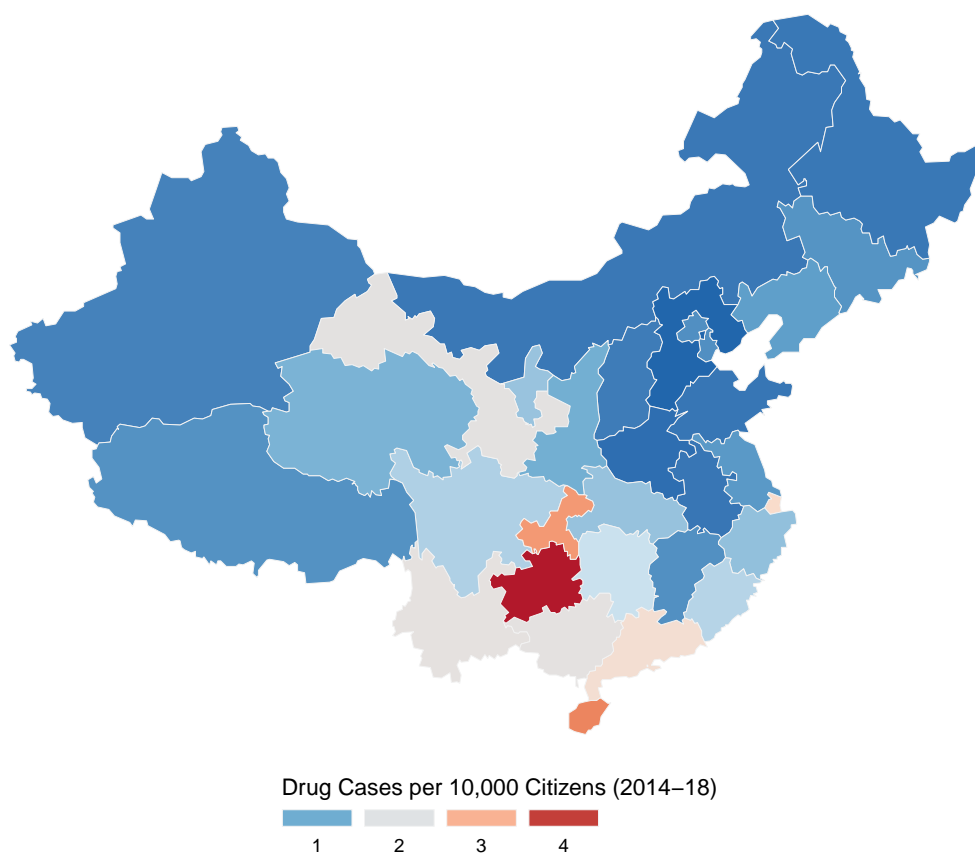
Figure 1 shows the number of heroin and methamphetamine cases per capita across China using the court data scraped and analyzed for this project. China's recreational drug market is growing substantially and is now an estimated \$82 billion (Levin 2015; Yang 1993). Drug-related offenses constituted 10.54% of all criminal cases in 2016, and their number is growing four times faster than other criminal cases (SPC 2017). Drug crimes are concentrated in China's southwestern provinces with large minority populations.

The sentencing guidelines for drug-related offenses in China are clear but allow for some judicial discretion. Articles 347 and 348 of China's Criminal Law map different drug quantities to sentencing outcomes (see text in Supporting Information). Defendants found with more than 50 grams of heroin or methamphetamine "are to be punished by 15 years of fixed-term imprisonment, life imprisonment or death sentence." Intermediate quantities (10-50 grams) yield a minimum sentence of seven years fixed term imprisonment, and smaller quantities (less than 10 grams) are to be sentenced to no more than seven years. Sentences can be made more or less severe depending on the circumstances of the case— whether the defendant was involved in international drug trafficking, lead or was involved in a criminal group, used arms or violence, or encouraged minors to commit offenses, among other factors.<sup>3</sup>

---

<sup>3</sup>Besides the Criminal Law, judges and lawyers also frequently consult with several other relevant Supreme People's Court documents. These include the Wuhan Conference Report (2015), the Dalian Conference Report (2008) and judicial interpretations of the Criminal Law issued in 2000, 2007, 2012, 2014 and 2016. These documents are not official laws, but they are referred to by judges and other legal workers (Personal Interview A001).

Figure 1: Severity of Drug Problem Across China



Note: Figure shows the number of drug cases per 10,000 citizens by province from 2014–2018. Dark blue provinces indicate areas where the drug problem is less severe; dark red provinces indicate areas where the problem is more severe. Figures constructed from author's dataset and include only methamphetamine and heroin cases. Population figures are projected for 2017 from census data.

#### 4 Data and Research Design

Many studies have found that court decisions are affected by the race and ethnicity of defendants.<sup>4</sup> The initial goal in conducting this project was to extend this literature to a new context. Our study is one of the first to consider criminal sentencing in China empirically ([Liebman 2015](#); [Lu and Miethe 2002](#)), and the first to assess the degree of ethnic discrimination in the system.

The CCP is increasingly transparent in many areas of governance ([Distelhorst and Hou 2017](#);

<sup>4</sup>African Americans are more likely to be charged with a serious offense ([Rehavi and Starr 2014](#)) and have higher incarceration rates ([Abrams, Bertrand and Mullainathan 2012](#)). The racial composition of juries also seems to affect court decisions. In felony trials in Florida, [Anwar, Bayer and Hjalmarsson \(2012\)](#) find that all-white jury pools convict black defendants significantly more often than white defendants.

---

Lorentzen, Landry and Yasuda 2013), and in some respects the granularity of the data that is now available exceeds what scholars have used in studies of sentencing in Western democracies (Liebman et al. 2018). Beginning in 2013, all courts in China became responsible for uploading judicial decisions onto a central website, “China Judgments Online” (中国裁判文书网), administered by the Supreme People’s Court. Current SPC President Zhou Qiang has emphasized “practicing the law under the sun” (阳光司法) and the promotion of “smart courts” (智慧法院) and computer-assisted judging. This push for transparency aims to raise the status of courts within the political system and with the public (Liebman et al. 2018).

As of April 2019, over 62 million court documents had been uploaded on the website, including over 7.5 million criminal proceedings. These cases are written in a standardized format across provinces and emphasize outcomes and case facts (Liebman 2015). Though the online case files should not be taken as representative of the full population of court cases in China, they nevertheless present a unprecedented opportunity to assess patterns in sentencing outcomes in the largest judicial system in the world, and an authoritarian case of particular importance.

For this paper, we scraped all first-instance court drug cases (n=14,853) judged by basic and intermediate courts in Yunnan province available on the China Judgments Online website as of October 2018. We removed cases if they had no identifying information about the crime type, if the case document had multiple defendants, if the ruling year was prior to 2014, or if the defendant did not possess either heroin or methamphetamine. We also restricted our analysis to cases involving trafficking, smuggling, transporting, or possession. This left us with 10,082 cases for the analysis. We developed an algorithm using rule-based regular expressions and machine learning to parse these cases into variable-based data. Finally, we scraped and parsed the full set of drug cases for the entire country (n=167,091) to use for generalizability tests.

Our empirical strategy is to use a conditioning approach to assess whether ethnicity is robustly associated with punishment severity once other case-based, legal, and demographic influences are accounted for (Mustard 2001; Rehavi and Starr 2014; Starr and Rehavi 2013; Steffensmeier, Painter-Davis and Ulmer 2017; Ulmer and Johnson 2004). If it is, this would be

suggestive of ethnic bias in the application of criminal justice in China.

The primary outcome variable is sentencing severity. If a defendant received a fixed prison sentence, we coded the length of the sentence (*pun.fixed.length*) in months. To conduct the analysis, it is necessary to have different punishments on the same scale, and we did so by converting all other punishments to months using an existing standard from the literature (Yin and Li 2009): a life sentence equals 264 months, a suspended death sentence equals 288 months, and a death sentence equals to 360 months. The variable *pun.severity* records the newly coded punishment severity after all conversions are made. The mean sentence in the Yunnan data was 91 months, and the median sentence was 48 months. For a secondary outcome variable, we also analyzed whether defendants received either life imprisonment, death, or a suspended death sentence (*pun.lifedeath*). This punishment was given to 7.6% of the defendants in the Yunnan data.

We coded a large number of characteristics of the defendant, including the name, ethnicity (*def.ethnicity*), sex (*def.female*), occupation (*def.occupation*), age (*def.age*) and whether a female defendant is pregnant (*def.pregnant*).<sup>5</sup> We also identify whether the defendant had previous convictions (*def.recid*), whether he was viewed as cooperative and remorseful (*good.attitude*), and whether he plead not guilty (*plead.notguilty*). In terms of the drug crime, we identify the type of the crime (*crime.type*) (e.g., trafficking, smuggling, transport and possession), drug name (*crime.drug*) (heroin, methamphetamine, marijuana, cocaine, and other types), and how much drug is identified in grams (*drug.quantity*). The location of the crime is also recorded (*crime.location*).<sup>6</sup>

These variables map to provisions in China's Criminal Law, allowing us to account for legal factors that might be associated with ethnicity and confound the inference. Most importantly, we have a relatively precise, objective measure of the severity of the crime— the quantity of the

---

<sup>5</sup>The ethnicity of the defendant is usually indicated directly in the ruling document. If not, we use the length of the name and common minority last names to identify whether a defendant is an ethnic minority (*def.minority*).

<sup>6</sup>We also collect information on the court proceedings, including the date of the trial (*ruling.date*), name of the court (*court.name*), the names of the judges (*judge.name*), and whether the defendant had a lawyer (*lawyer*).

---

drug itself (Rehavi and Starr 2014).<sup>7</sup> The Supporting Information provides summary statistics for these variables and more information on the parsing process.

## 5 Results

### 5.1 Core Estimates

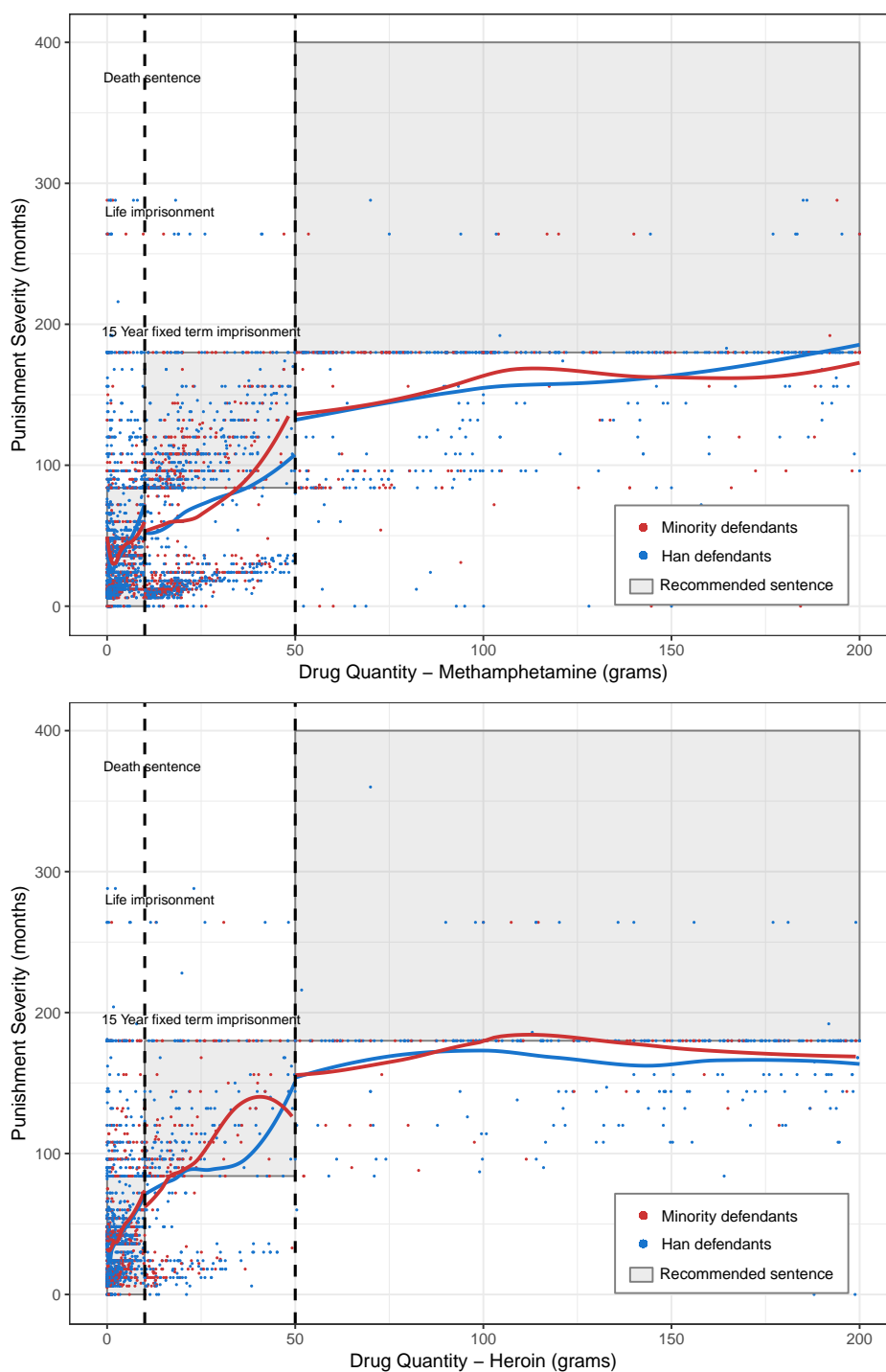
Figure 2 shows the raw data for all cases in the Yunnan data (n=10,082), comparing sentencing outcomes in months against the drug quantity in grams recorded in the court case press release. The grey boxes show the recommended sentence range for the drug quantity according to Articles 347 and 348 of China's Criminal Law. The red points indicate defendants identified as an ethnic minority and the blue points indicate Han defendants. The top panel shows methamphetamine cases and the bottom panel shows heroin. The red and blue loess lines are estimated within drug quantity bins (0-10 grams, 10-50 grams, 50+ grams) to better capture the data generating process.

The gap between the red and blue lines indicates that minority defendants tend to fare worse than Han defendants that had equivalent quantities of methamphetamine/heroin. This presents some preliminary evidence in favor of ethnic bias in criminal sentencing.

---

<sup>7</sup>As Rehavi and Starr (2014) discuss, the prosecutor's decision to pursue different criminal charges can also be racially motivated, so studies that lack information on the actual offense can underestimate bias.

Figure 2: Drug Quantities, Ethnicity, and Sentencing Outcomes in Yunnan



Note: Figure shows the drug amount of methamphetamine and heroin (grams) against the sentencing outcome (months) for 10,082 drug possession, transport, and trafficking cases from 2014–2018. The blue points and line correspond to defendants of the Han majority; red indicates defendants of an ethnic minority group. The shaded grey boxes show the recommended sentencing levels for the drug quantity according to China’s Criminal Law. Points below those boxes indicate leniency; points above that line indicate severity.

Further analysis is necessary to account for the influence of other factors that might affect sentencing outcomes. Table 1 reports the estimated effect of *def.ethnicity* on *pun.severity* and *pun.lifedeath* across a range of covariate sets. Model 1 includes just the various drug quantity indicators; Model 2 adds other attributes of the crime (*crime.act*, *international*, *minors*, *def.recid*); Model 3 adds additional defendant demographics (*def.mental*, *def.female*, *def.age*); Model 4 adds polynomial splines for all of the drug quantity variables; Model 5 adds fixed effects for the ruling year and court; and Model 6 includes additional defense attributes (*def.pleadnotguilty*, *def.goodattitude*). All models are estimated with OLS using robust standard errors.

Note that the final few variables could plausibly be considered either confounding or mediating variables (Sen and Wasow 2016). For example, it is possible that minorities are sorted into systematically harsher courts, or that they are less likely to be deemed to have a “good attitude” (Lu and Miethe 2002). The causal effect of minority status might flow through these variables, which means they should not be included in the regression. But they are also important background factors to consider and should impact sentencing according to the Criminal Law. Our approach is to simply assess whether their inclusion affects the substantive conclusions of the analysis. For the purposes of discussion, we will refer to M4 as the “baseline model” and M6 as the “full model.”

The coefficient estimates are relatively stable and suggest a degree of ethnic discrimination in Yunnan. The minimalist specification M1 produces an estimate of -4.8 months for the *pun.severity* outcome measure, but this is likely because the analysis does not incorporate other relevant case attributes or capture the functional form of the drug quantity variables. Specifications M2 through M6 yield estimates ranging from about 1.5 to 8.0 months.

Minority defendants are also more likely to be sentenced to death or life imprisonment. The coefficient estimates here range from 1 to 2 percentage points, which is substantively important given that the baseline rate of this sentence is only 7.6%.<sup>8</sup>

---

<sup>8</sup>The Supporting Information includes additional robustness checks. Table A2 presents the same specifications excluding defendants that were outliers on the drug quantity variables.

Table 1: Effect of Minority Status on Sentencing (Yunnan)

#	Covariates	Outcome	
		<i>pun.severity</i>	<i>pun.lifedeath</i>
M1.	<i>meth.amt</i> + <i>heroin.amt</i> + <i>cocaine.amt</i> + <i>marijuana.amt</i> + <i>other.amt</i>	-4.844 (1.742)	0.008 (0.005)
M2.	M1. + <i>crime.act</i> + <i>international</i> + <i>minors</i> + <i>def.recid</i>	7.474*** (1.099)	0.025*** (0.005)
M3.	M2. + <i>def.mental</i> + <i>def.female</i> + <i>def.age</i>	7.960*** (1.220)	0.025*** (0.006)
M4.	M3. + <i>amt</i> cubic splines	2.843*** (0.841)	0.015*** (0.005)
M5.	M4. + <i>ruling.year</i> (fixed effects) + <i>court.name</i> (fixed effects)	1.585* (0.899)	0.011* (0.006)
M6.	M5. + <i>def.pleadnotguilty</i> + <i>def.goodattitude</i>	1.579* (0.896)	0.011* (0.006)

Note: Table shows coefficient estimates from regressions of *pun.severity* (months) and *pun.lifedeath* on *def.minority* across six different covariate sets. M4 is the “baseline specification” referred to throughout the paper. Data come from Yunnan filtered dataset, which includes all heroin/methamphetamine cases in the [wenshu.court.gov.cn](http://wenshu.court.gov.cn) website from 2014-2018. Robust standard errors shown in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## 5.2 Mechanism Analysis

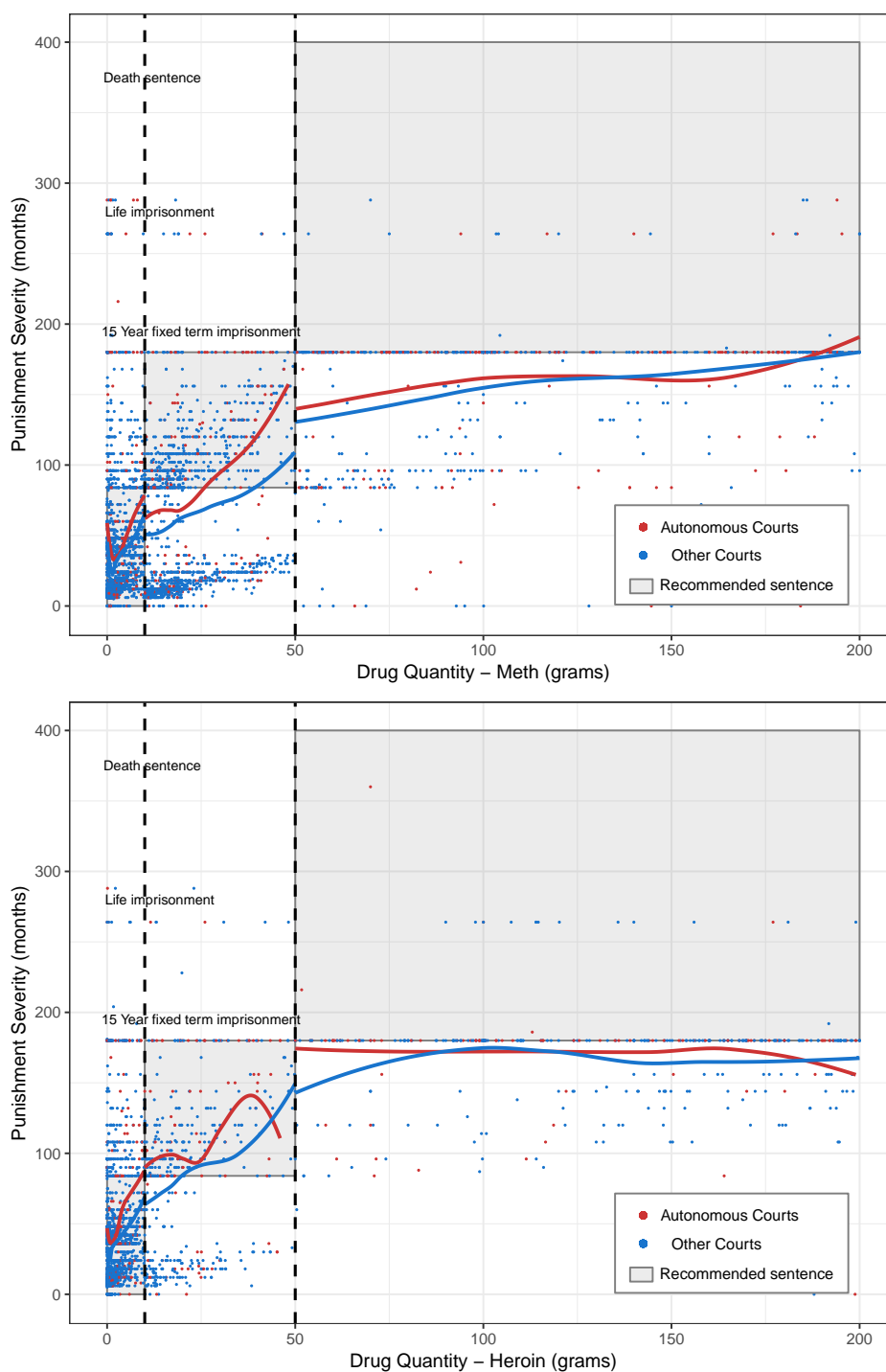
Here we consider two possible additional mechanisms relating to ethnic discrimination in China’s criminal justice system: court-level severity and legal representation.

### *Court Level Severity*

In our estimates in Table 1, the introduction of court fixed effects reduces the point estimate for the sentence by amount one month (2.843 in M4 vs. 1.585 in M5). This suggests some sort of mediating relationship might be at work— the effect of ethnicity on sentencing may flow in part through court assignment.



Figure 3: Drug Quantities, Autonomous Courts, and Sentencing Outcomes in Yunnan



Note: Figure shows the drug amount of methamphetamine and heroin (grams) against the sentencing outcome (months) for 10,082 drug possession, transport, and trafficking cases from 2014-2018. The blue points and line correspond to defendants tried in regular courts; red indicates defendants tried in courts in autonomous prefectures or counties. The shaded grey boxes show the recommended sentencing levels for the drug quantity according to China's Criminal Law. Points below those boxes indicate leniency; points above that line indicate severity.

One possibility is that courts located in autonomous areas are different from those located in other areas. By virtue of their geographic concentration, minority defendants are disproportionately likely to be tried in courts in autonomous counties and prefectures. In our Yunnan data, 27.9% of minority defendants (762/2733) are tried in autonomous courts, compared to 13% of Han defendants (956/7349).

The autonomous status marker confers a degree of autonomy in local government affairs, and governments in these areas have special obligations to protect the heritage and language of the group. Autonomous courts have more resources to hire minority judges, translators, and other staffers. These arrangements should theoretically benefit minority defendants. However, many of Yunnan's autonomous areas are heavily affected by the drug problem, and some of them (i.e., Dali Bai autonomous prefecture; Dehong Dai autonomous prefecture; and Lancang Lagu autonomous county) have issued their own regulations regarding drug crimes. These regulations emphasize harsh punishments for drug offenses, even when the quantities involved are small.

The data suggests that courts in autonomous areas are indeed systematically harsher. Figure 3 shows scatterplot relating drug quantity and sentence severity, this time breaking out cases tried in autonomous versus regular courts. There is a visible sentencing gap between courts in autonomous localities and regular areas.

Table 2 assesses this mechanism with a regression analysis. When we introduce an indicator variable (*court.autonomous*) into our baseline model M4, and we find that individuals tried in such courts tend to have sentences that are 3.0 to 3.2 months longer ( $p=0.000$ ). The effect of the defendant's ethnicity also remains significant, in the territory of about 2.7 months. The third column introduces an interaction effect between *def.minority* and *court.autonomous*. There is a small negative but insignificant coefficient on the interaction term. Autonomous courts are systematically harsher, and minority defendants receive harsher sentences in general, but it does not appear that this bias is worse in autonomous courts than normal courts.

Table 2: Effect of Autonomous Courts on Sentencing (Yunnan)

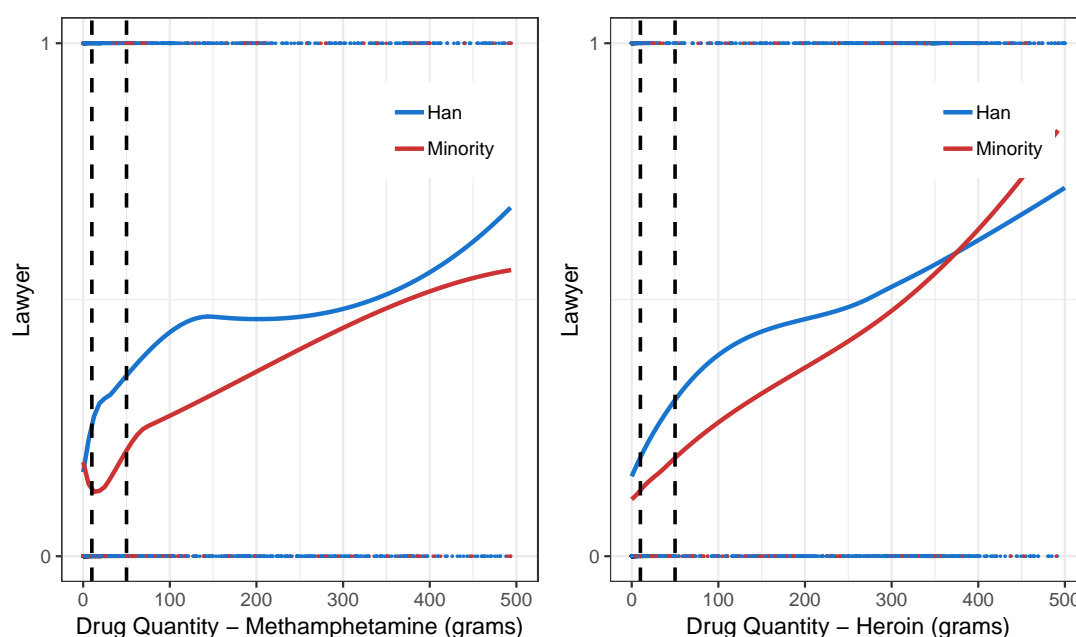
	(1)	(2)	(3)
<i>def.minority</i>	3.121*** (0.802)	2.721*** (0.813)	2.845*** (0.918)
<i>court.autonomous</i>		2.978*** (0.975)	3.204** (1.245)
<i>def.minority:court.autonomous</i>			-0.557 (1.908)
<i>crime.actsmuggle</i>	89.762*** (5.633)	89.309*** (5.632)	89.241*** (5.637)
<i>crime.acttrafficking</i>	62.631*** (1.410)	62.417*** (1.411)	62.431*** (1.412)
<i>crime.acttransport</i>	97.162*** (1.732)	97.012*** (1.731)	96.988*** (1.734)
<i>minors</i>	26.824** (11.310)	26.648** (11.305)	26.713** (11.308)
<i>def.recid</i>	12.860*** (0.986)	12.863*** (0.986)	12.862*** (0.986)
<i>def.mental</i>	-6.528 (8.561)	-6.563 (8.556)	-6.592 (8.557)
<i>def.female</i>	-1.696 (1.070)	-1.646 (1.069)	-1.647 (1.070)
<i>def.age</i>	0.179*** (0.036)	0.176*** (0.036)	0.176*** (0.036)
<i>court.intermediate</i>	40.595*** (1.523)	40.361*** (1.524)	40.373*** (1.524)
Observations	8,006	8,006	8,006
Adjusted R <sup>2</sup>	0.857	0.857	0.857
<i>amt</i> cubic splines	Yes	Yes	Yes

Note: Table shows coefficient estimates from regressions of *pun.severity* (months) on *def.minority* and *court.autonomous*. The covariates represent the “baseline specification” from M4 referred to throughout the paper. Coefficient estimates on the polynomial splines for the drug quantity variables are excluded. Data come from Yunnan filtered dataset, which includes all heroin/methamphetamine cases in the [wenshu.court.gov.cn](http://wenshu.court.gov.cn) website from 2014-2018. Robust standard errors shown in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

### Legal Representation

The case files in Yunnan include the name of the lawyer representing the defendant at trial. When this field is populated, we have a clear indicator that the person had at least some formal representation during the trial process (*lawyer*). Figure 4 shows this variable against the drug quantity of the case, separated out by the ethnicity of the defendant. The loess lines are the probability that a defendant of that ethnicity will have a lawyer at a given level of drug quantity.

Figure 4: Variation in Legal Representation by Ethnicity (Yunnan)



Note: Figure shows the drug amount of methamphetamine and heroin (grams) against the whether the defendant had legal representation for 10,082 drug possession, transport, manufacturing and trafficking cases from 2014-2018. The blue points and line correspond to defendants of the Han majority; red indicates defendants of an ethnic minority group.

We find clear evidence that ethnic minority defendants are less likely to have legal representation. The presence of a lawyer increases with the drug quantity, but there exists a substantial representation gap between minority and Han defendants. In our Yunnan data, 33.6% of Han defendants (2475/7349) have a lawyer, compared with only 27.3% of minority defendants (747/2733). This may have downstream effects for how defendants plea, whether they choose to cooperate with authorities, and whether they are able to present exonerating evidence during trial (Lu and Miethe 2002; Personal Interview B009).

Note that this mechanism is not necessarily indicative of policy or institutional discrimination directed at minorities. The gap might arise simply because of differences in income— ethnic minorities in Yunnan tend to live in relatively impoverished rural areas, and thus might not have access to proper legal representation or the means to pay for that representation. Our interview evidence suggests that defense lawyers for drug defendants can cost between 10,000 and 60,000 yuan (1,400 to 8,400 USD) for a defendant (Personal Interview A002; Personal Interview B003). Defendants who can not afford a good lawyer either end up with no legal representation or use a court-assigned lawyer who lacks the monetary incentive to prepare a good defense (Personal Interview A001; Personal Interview A002). Some minority defendants might prefer to have close friends or relatives represent them because of language barriers.<sup>9</sup>

From a policy perspective, improving the existing defense system, increasing funding for under-resourced courts, and promoting representation for more defendants may serve to alleviate some of the biases in sentencing outcomes we observe.

### 5.3 Generalizability and Additional Tests

We applied our parsing algorithm to the text from the relevant drug case files for the entire country— 166,770 observations in total. We have less confidence in the full country data. There is some variation in reporting standards across provinces, and as a result, our processed dataset for other provinces has higher item-level missingness. The analysis here should be considered less definitive, but it nevertheless allows us to assess the theory with geographic and group based tests, and to see whether ethnic discrimination in China is a Yunnan specific phenomenon.

Figures A3 and Table A3 in the Supporting Information show the estimates of ethnicity on sentencing outcomes for the pooled country dataset. The coefficient for the core analysis of *pun.severity* hovers around 1.5 to 8.5 months, in line with what we observed in Yunnan.

This analysis masks important variation in sentencing patterns, which are not uniform across

---

<sup>9</sup>This gap could also be a reflection of differences in court funding: courts located in poorer areas might have fewer resources to help defendants find legal representation. In our sample, 17.5% of defendants in courts located in autonomous counties have a lawyer, while the number is 38.4% for courts located outside of autonomous counties.

---

provinces and groups. Theoretically, we expect discrimination in criminal sentencing in drug cases to arise under a certain set of circumstances: an ethnic minority becomes associated with the drug problem, and that problem is particularly severe. The Yunnan case perfectly fits this description. What other provinces in China might be candidates? And what groups might be most affected?

Table 3 presents the rank order of ethnic groups by the degree to which they are afflicted by the drug problem, according to our dataset. Column 3 shows each groups' share of the population according to the 2010 census, and Column 5 shows the groups' share of the total drugs confiscated (2014-2018) according to our full country data. The ratio column simply divides the drug share by the population share—ratios greater than 1 indicate that the group was found with more drugs than their share of the population. For example, the Yi minority comprises only 0.654% of the population (8,714,393 people), but had 4.56% of the total drugs confiscated (367,443 grams of methamphetamine/heroin), yielding a ratio of 6.97.

The groups most affected by the drug problem are primarily located in Yunnan province, most notably the Dai and Hani people, but also smaller groups like the Jingpo, Blang, Lahu, and Wa (Personal Interview B002). The Dongxiang minority group, a Mongolic people concentrated in an autonomous prefecture in Gansu province, had drug quantities over 24 times their proportion of the population. The minority-drug nexus is less prominent but still present in other provinces— notably Sichuan (Lisu and Yi peoples), Guizhou (Yi), Guangxi (Yao, Gin, Mulao, Zhuang), and Ningxia (Hui).

Table 3: Ethnic Groups and the Drug Problem (select groups)

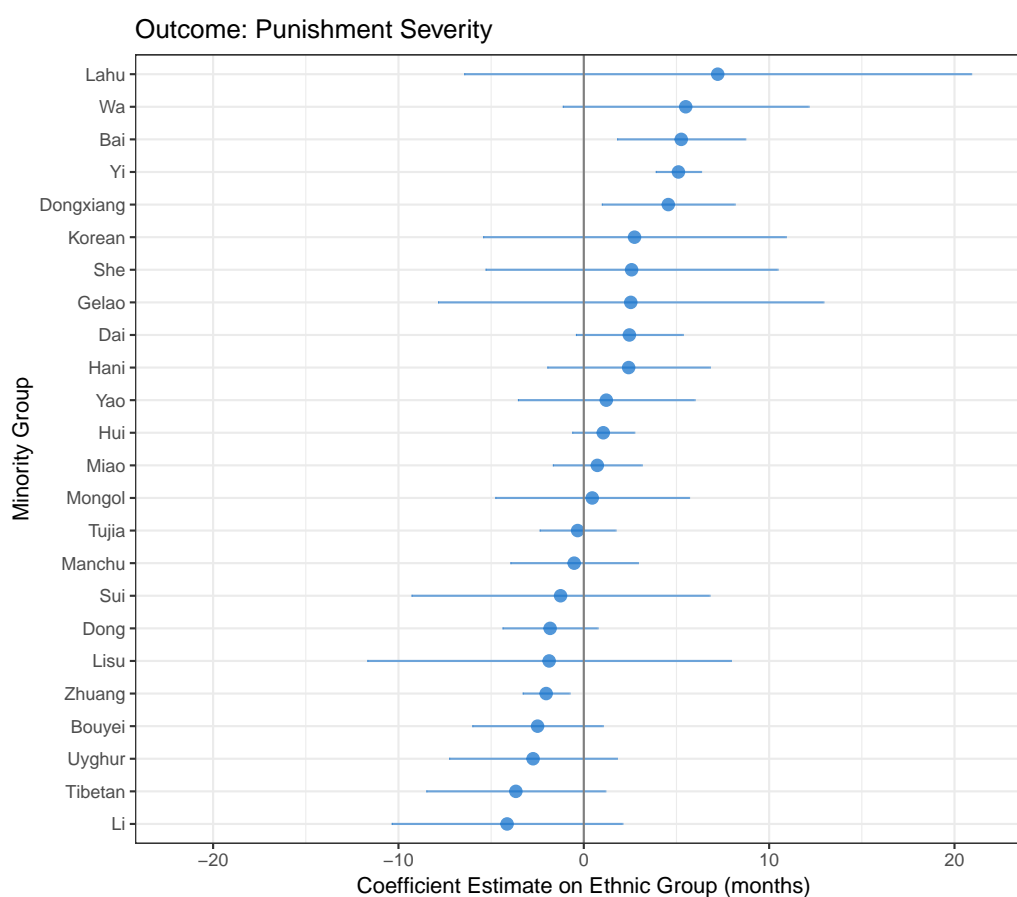
Group	Population		Drugs		Ratio (5)/(3)	Location
	Total	%	Total	%		
	(2)	(3)	(4)	(5)		
1. Blang	119,639	0.009	48,665	0.60	67.26	Yunnan
2. Dai	1,261,311	0.095	445,749	5.53	58.44	Yunnan
3. Jingpo	147,828	0.011	41,623	0.52	46.56	Yunnan
4. Dongxiang	621,500	0.047	92,827	1.15	24.70	Gansu
5. Lahu	485,966	0.036	57,213	0.71	19.47	Yunnan
6. Wa	429,709	0.032	29,483	0.37	11.35	Yunnan
7. Hani	1,660,932	0.125	105,744	1.31	10.53	Yunnan
8. De'ang	20,556	0.002	1,144	0.01	9.21	Yunnan
9. Yi	8,714,393	0.654	367,443	4.56	6.97	Sichuan, Yunnan, Guizhou
10. Gin	28,199	0.002	1,059	0.01	6.21	Guangxi
11. Salar	130,607	0.010	4,146	0.05	5.25	Qinghai, Gansu
12. Lisu	702,839	0.053	13,752	0.17	3.24	Yunnan, Sichuan
13. Jino	23,143	0.002	400	0.00	2.87	Yunnan
14. Nakhi	326,295	0.024	4,825	0.06	2.45	Yunnan
15. Hezhen	5,354	0.000	70	0.00	2.19	Yunnan
16. Bai	1,933,510	0.145	24,936	0.31	2.13	Yunnan, Guizhou, Hunan
17. Mulao	216,257	0.016	2,478	0.03	1.90	Guangxi
18. Zhuang	16,926,381	1.271	192,211	2.39%	1.88	Guangxi, Yunnan, Guizhou
19. Yao	2,796,003	0.210	25,466	0.32	1.51	Hunan, Guangdong, Guangxi
20. Miao	9,426,007	0.708	79,233	0.98	1.39	Hunan, Yunnan, Sichuan
24. Han	1,220,844,520	91.643	6,299,900	78.20	0.85	All provinces
29. Tibetan	6,282,187	0.472	18,942	0.24	0.50	Tibet
40. Uyghur	10,069,346	0.756	7,293	0.09	0.12	Xinjiang

Note: Table shows top twenty ethnic groups in China most afflicted by drug problem, as well as other ethnic groups of interest (Han, Tibetans, Uyghurs). Population totals are drawn from the 2010 census. Drug figures are aggregated from individual court case data from the SPC website (2014-2018). Groups with a ratio of 1 have drug quantities in proportion to the population. Ratios greater than 1 indicate ethnic groups with a greater share of drugs than share of the population.

We can replicate the same analysis as above, this time substituting in indicators for the 55

different ethnic groups in place of the aggregated minority variable.<sup>10</sup> Han Chinese represent the excluded category. The coefficient estimate on each group can serve as a measurement of ethnic bias, with coefficients greater than zero indicating that minorities in that group fare systematically worse compared to Han defendants that have committed comparable crimes. Figure 5 presents these disaggregated estimates, employing the baseline specification (M4). The figure excludes estimates where there were less than 25 cases of defendants from that ethnic group or where the group itself has less than 400,000 members.

Figure 5: Estimates of Ethnic Bias Across Minority Groups



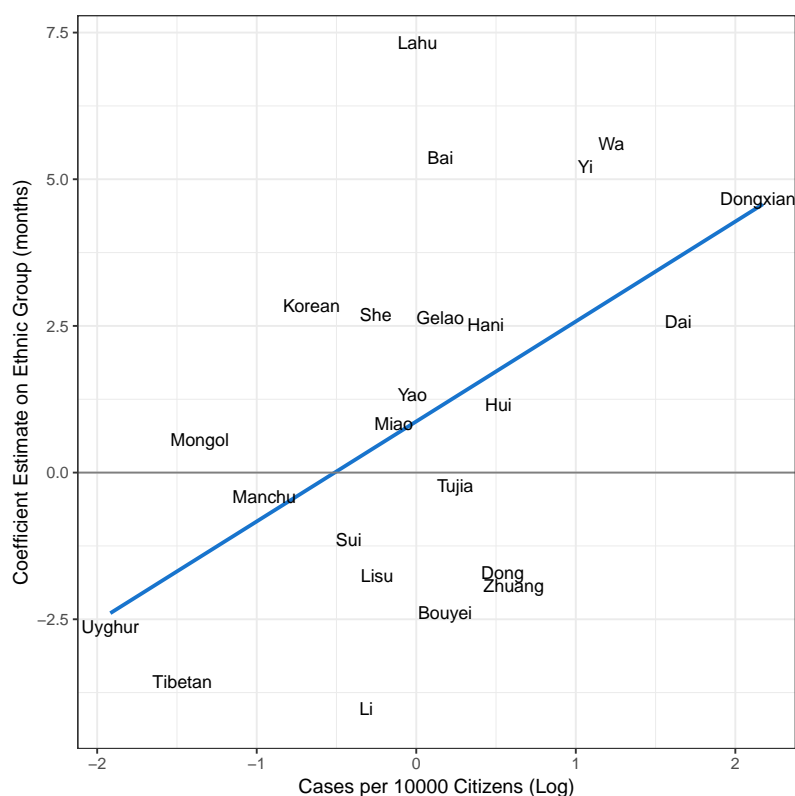
Note: Table shows coefficient estimates from regressions of *pun.severity* (months) on dummy variables for each minority group. The excluded category is Han. Point estimates to the right of zero indicate the minority group is on average treated more harshly the Han defendants. The estimates employ the full specification (M4) which includes all demographic and crime level characteristics in the dataset as well as polynomial splines for each drug quantity variable. Segments are 95% confidence intervals.

<sup>10</sup>This analysis was not feasible in the Yunnan analysis given the low number of observations.



Many of the coefficient estimates lack precision, reflecting the small number of cases for each group. Most of the estimates are positive (14/24) but only three groups show definitive evidence of discrimination– the Bai, Yi, and Dongxiang peoples. One group– the Zhuang people– appears to benefit from leniency. Figures A4 and A5 in the Supporting Information assess the robustness of these estimates.

Figure 6: Testing the Problem Minority Hypothesis



Note: Figure shows coefficient estimates from regressions of *pun.severity* (months) on dummy variables for each minority group against a measure of group involvement in the drug trade– the log of the total drug cases (2014-2018) per 10,000 citizens for each group. The zero line on the y-axis corresponds to sentencing patterns equivalent to Han Chinese (Han is the excluded category).

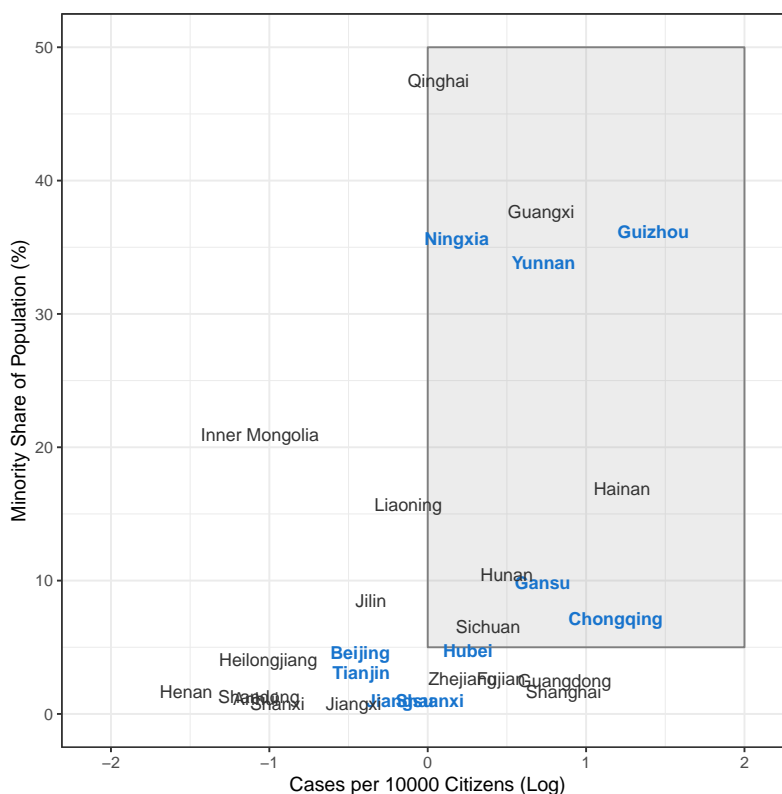
These estimates, though noisy, are consistent with the “problem minority” framework posed above. Figure 6 shows that defendants from groups that are more heavily involved in the drug trade are more likely to receive harsher sentences. The figure plots the coefficient estimate on the group– a proxy for discrimination– against the group’s involvement in the drug trade

(the log case rate per 10,000 citizens). We see a strong positive relationship between the drug problem and punishment severity. The more a group is involved in the drug trade, the harsher the sentences defendants from that group receive.

This logic also produces important geographic variation in ethnic bias in criminal sentencing. In the Supporting Information, Figures A6 and A7 shows coefficient estimates on *def.minority* and associated p-values for each province, exploring robustness across the six covariate sets. Points in blue indicate coefficients that reject the one-sided null hypothesis of no effect at the .05 level, points in grey show estimates that are insignificant. The provinces highlighted in blue are those where we are fairly confident that there is discrimination in the sentencing process against ethnic minorities involved in drug crimes. This list includes: Beijing, Chongqing, Gansu, Guizhou, Hubei, Jiangsu, Ningxia, Shaanxi, Tianjin, and Yunnan. This is not to say that discrimination is not present in other provinces; we simply did not detect it with our data and methods.

Figure 7 arrays the different Chinese provinces based on the concentration of minorities and the severity of the drug problem. The shaded grey box shows provinces with large minority populations that are highly afflicted by the drug problem. The provinces highlighted with bold blue text are those where there is relatively robust evidence of discrimination. We observe that discrimination does tend to cluster in the areas predicted by the theory—Ningxia, Yunnan, Guizhou, Gansu, and Chongqing all meet the conditions of Hypothesis 2. But we were surprised to find weak or no evidence of discrimination in certain provinces that meet our conditions—notably Guangxi and Sichuan—and strong evidence of discrimination in provinces that do not have large minority populations involved in the drug trade. It is interesting that three of China's four provincial-level municipalities also show evidence of discrimination—Beijing, Chongqing, and Tianjin. This suggests the “problem minority” dynamic captures some but not all of the story.

Figure 7: Exploring Variation Across Provinces



Note: Figure shows the minority share of population against the the log of the total drug cases (2014-2018) per 10,000 citizens for each province. The grey shaded area indicates provinces with a high concentration of drugs and minority groups. The provinces named with bold blue text are those with robust evidence of discrimination. See Figures A6 and A7 in the Supporting Information for full province estimates.

## 6 Limitations and Alternative Perspectives

It is important to be transparent about several limitations of the analysis. First, we have only been able to assess ethnic bias in the sentencing process, for individuals that have formally entered into the criminal justice system and been charged with a crime. It is possible that biases (in either direction) are present in which individuals get stopped by police, how they are treated by police, what evidence is collected, what type of crime they are charged with, what legal options are presented to them, and the quality of representation they are given (Personal

---

Interview B002). There is empirical evidence of racial discrimination in the United States in these stages, and discrimination at one stage may affect whether discrimination manifests itself at another (Anwar and Fang 2006; Fryer 2019; Know, Lowe and Mummolo 2019; Rehavi and Starr 2014).

Unfortunately a “whole of process” approach, whereby bias is mapped at each stage and selection is fully modeled, is currently not possible in the Chinese case because of data constraints. Thus when we observe that anti-minority bias is stronger for some groups than others, this conclusion should be read and interpreted narrowly, as about bias in the criminal sentencing process. We are unable to assert anything about biases elsewhere in the system.

Second, the data we analyze is not a census of all drug cases in China, but rather a large convenience sample with unknown biases (Liebman et al. 2018). Lower level courts are legally required to post case files to the central government website, but they also have discretion to withhold cases for various reasons. Other scholars have estimated that the courts seem to upload about 40 to 50% of cases (Liebman et al. 2018; Liu et al. 2019), though this proportion varies substantially across jurisdictions. If we exclude settlement cases, where disclosure is not required by the SPC, the case publishing rate in 2015 could have been as high as 75.77% (Tang and Liu 2019, p. 23). We have reason to believe that drug cases should have a relatively high proportion of cases online (Personal Interview B003), as a.) the judicial proceedings are generally straightforward b.) the issue itself is not excluded from the broader transparency mandate and c.) mediation is less common for criminal disputes without a victim. But it is currently impossible to get the full sample of all case files, and it is difficult to know definitively how many cases are missing from our sample.

As with many studies of the Chinese political system and other authoritarian regimes, our approach has been to do the best we can given the limited data that is available, rather than refrain from studying an important topic entirely. We hope readers are sympathetic to this mindset. Future research can replicate our findings if data quality further improves.

Third, as with any observational research design, it is possible that our findings are driven

by some omitted confounding variable. Assessments of the causal effect of race or ethnicity are particularly fraught, as such identities are immutable and are causally prior to effectively all other variables (Sen and Wasow 2016). This makes standard conditioning strategies problematic (Rehavi and Starr 2014). It is possible that minority defendants are committing crimes that are slightly different in ways that are not captured by our coding scheme, which in turn drives the discrepancy in sentencing.

In our defense, our analysis takes into account all factors relevant to sentencing as identified in China's Criminal Law and other relevant SPC documents, and we have a very granular, objective measure of the crime in our drug quantity variable. So if there are any reasons why judges in China are systematically giving minorities harsher sentences for equivalent crimes, they are not justified in the law. The findings also hold up to particularly conservative specifications that introduce court fixed effects.

## 7 Conclusion

One motivation for this project came from the ongoing repression of the Uyghur minority in China's Xinjiang Uyghur Autonomous region. Here, Uyghurs have been characterized as prone to terrorism, extremism and separatism, and the government has engaged in mass forced reeducation in an attempt to promote stability in the region.

Using drug cases from Yunnan province, we estimate that minority defendants, on average, receive sentences that are about 1.5 to 8.0 months longer than Han defendants who have committed similar drug-related offenses. The magnitude of our finding might seem small in light of the events in Xinjiang, but our argument is that not all minorities are treated the same under authoritarian rule, and that discrimination is not uniformly severe. Discrimination is targeted against "problem minorities" most heavily involved in the drug trade, and more severe punishments are meted out in autonomous areas where minorities are concentrated. These findings are consistent with an authoritarian stability maintenance logic, whereby the judicial system becomes an institution through which local governments exercise control (Liebman 2014; Minzner

2011; Wang and Minzner 2015; Yang 2017).

The next step in this research agenda is to further test the generaliability of the “problem minority” hypothesis, for other issues in China, and for minorities in other authoritarian systems. Some of the great wrongs in human history have come when an authoritarian regime dominated by a single ethnic group comes to view an ethnic minority as threatening. By understanding the political conditions through which discrimination arises, we can contribute to its elimination.

---

## References

- Abrams, David S., Marianne Bertrand and Sendhil Mullainathan. 2012. "Do Judges Vary in Their Treatment of Race." *The Journal of Legal Studies* (2):347–383.
- Alesina, Alberto and Eliana La Ferrara. 2014. "A Test of Racial Bias in Capital Sentencing." *American Economic Review* 104(11):3397–3433.
- Anwar, Shamena and Hanming Fang. 2006. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review* 47(2):127–151.
- Anwar, Shamena, Patrick Bayer and Randi Hjalmarsen. 2012. "The Impact of Jury Race in Criminal Trials." *The Quarterly Journal of Economics* (2):1017–1055.
- Blair, Irene V, Charles M Judd and Kristine M Chapleau. 2004. "The Influence of Afrocentric Facial Features in Criminal Sentencing." *Psychological Science* 15(10):674–679.
- Cederman, Lars-Erik, Andreas Wimmer and Brian Min. 2010. "Why Do Ethnic Groups Rebel?: New Data and Analysis." *World Politics* (1):87–119.
- Cornell, Svante E. 2002. "Autonomy as a Source of Conflict: Caucasian Conflicts in Theoretical Perspective." *World Politics* 54(2):245–276.
- deLisle, Jacques. 2014. China's Legal System. In *Politics in China: An Introduction*, ed. William A. Joseph. Oxford University Press pp. 224–53.
- Distelhorst, Greg and Yue Hou. 2014. "Ingroup Bias in Official Behavior: A National Field Experiment in China." *Quarterly Journal of Political Science* p. 203–230.
- Distelhorst, Greg and Yue Hou. 2017. "Constituency Service Under Nondemocratic Rule: Evidence from China." *Journal of Politics* (3):1024–1040.
- Fryer, Roland G. 2019. "An Empirical Analysis of Racial Differences in Police Use of Force." *Journal of Political Economy* 127(3):1210–1261.

- 
- Gandhi, Jennifer. 2008. *Political Institutions Under Dictatorship*. Cambridge University Press.
- Gandhi, Jennifer and Adam Przeworski. 2007. "Authoritarian Institutions and the Survival of Autocrats." *Comparative Political Studies* 40(11):1279–1301.
- Gehlbach, Scott, Konstantin Sonin and Milan W Svobik. 2016. "Formal Models of Nondemocratic Politics." *Annual Review of Political Science* 19:565–584.
- Grossman, Guy, Oren Gazal-Ayal, Samuel D. Pimentel and Jeremy M. Weinstein. 2015. "Descriptive Representation and Judicial Outcomes in Multiethnic Societies." *American Journal of Political Science* (1):44–69.
- Han, Enze and Harris Mylonas. 2014. "Interstate Relations, Perceptions, and Power Balance: Explaining China's Policies Toward Ethnic Groups, 1949–1965." *Security Studies* p. 148–181.
- Hassan, Mai. 2017. "The Strategic Shuffle: Ethnic Geography, the Internal Security Apparatus, and Elections in Kenya." *American Journal of Political Science* 61(2):382–395.
- Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley CA: University of California Press.
- King, Ryan D and Darren Wheelock. 2007. "Group Threat and Social Control: Race, Perceptions of Minorities and the Desire to Punish." *Social Forces* 85(3):1255–1280.
- Know, Dean, Will Lowe and Jonathan Mummolo. 2019. "The Bias is Built In: How Administrative Records Mask Racially Biased Policing." *Working Paper* pp. 1–66.
- Leibold, James. 2010. "More Than a Category: Han Supremacism on the Chinese Internet." *The China Quarterly* 203:539–559.
- Leibold, James. 2013. "Ethnic Policy in China: Is Reform Inevitable?" *Policy Studies* 68:1–65.
- Levin, Dan. 2015. "Despite a Crackdown, Use of Illegal Drugs in China Continues Unabated." *The New York Times* .



- 
- Levy, Jack S. 2008. "Case Studies: Types, Designs, and Logics of Inference." *Conflict Management and Peace Science* 25(1):1–18.
- Li, Enshen. 2015. "Towards the Lenient Justice? A Rise of 'Harmonious' Penalty in Contemporary China." *Asian Journal of Criminology* 10(4):307–323.
- Lieberthal, Kenneth G. 1992. "Introduction: the 'Fragmented Authoritarianism' Model and its Limitations." *Bureaucracy, Politics, and Decision Making in post-Mao China* pp. 1–30.
- Liebman, Benjamin L. 2014. "Legal Reform: China's Law-Stability Paradox." *Daedalus* (2):96–109.
- Liebman, Benjamin L. 2015. "Leniency in Chinese Criminal Law? Everyday Justice in Henan." *Berkeley Journal of International Law* (33).
- Liebman, Benjamin L., Margaret E. Roberts, Rachel E. Stern and Alice Z. Wang. 2018. "Mass Digitization of Chinese Court Decisions: How to Use Text as Data in the Field of Chinese Law." *21st Century China Center Research Paper No.2017-01* .
- Lim, Claire S.H., Bernardo S. Silveira and James M. Synder. 2016. "Do Judges' Characteristics Matter? Ethnicity, Gender, and Partisanship in Texas State Trial Courts." *American Law and Economic Review* (2):302–357.
- Liu, John Zhuang, T.J. Wong, Yang Yi and Tianyu Zhang. 2019. "Transparency in an Autocracy: China's "Missing Cases" in Judicial Opinion Disclosure." *Working Paper* pp. 1–52.
- Lorentzen, Peter, Pierre Landry and John Yasuda. 2013. "Undermining Authoritarian Innovation: the Power of China's Industrial Giants." *The Journal of Politics* 76(1):182–194.
- Lu, Hong and Terance D. Miethe. 2002. "Legal Representation and Criminal Processing in China." *The British Journal of Criminology* 42(2):267–280.
- Minzner, Carl F. 2011. "China's Turn Against Law." *The American Journal of Comparative Law* (4):935–984.

- 
- Morelli, Massimo and Dominic Rohner. 2015. "Resource Concentration and Civil Wars." *Journal of Development Economics* p. 32–47.
- Mustard, David B. 2001. "Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the US Federal Courts." *The Journal of Law and Economics* 44(1):285–314.
- Ng, Kwai Hang and Xin He. 2017. *Embedded Courts: Judicial Decision-Making in China*. Cambridge University Press.
- Porter, Judith R and Robert E Washington. 1993. "Minority Identity and Self-esteem." *Annual Review of Sociology* 19(1):139–161.
- Rehavi, M. Marit and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* (6):1320–1354.
- Sen, Maya and Omar Wasow. 2016. "Race as a Bundle of Sticks: Designs that Estimate Effects of Seemingly Immutable Characteristics." *Annual Review of Political Science* 19:499–522.
- SPC. 2017. "People's Court Drug Control Work White Paper (renmin fayuan jindu gongzuo baipishu)." *Supreme People's Court*.
- Starr, Sonja B and M Marit Rehavi. 2013. "Mandatory Sentencing and Racial Disparity: Assessing the Role of Prosecutors and the Effects of Booker." *Yale Law Journal* 123:2.
- Steffensmeier, Darrell, Noah Painter-Davis and Jeffery Ulmer. 2017. "Intersectionality of Race, Ethnicity, Gender, and Age on Criminal Punishment." *Sociological Perspectives* 60(4):810–833.
- Svolik, Milan W. 2012. *The Politics of Authoritarian Rule*. Cambridge University Press.
- Tang, Yingmao and John Zhuang Liu. 2019. "Mass Publicity of Chinese Court Decisions: Market-Driven or Authoritarian Transparency?" *China Review* 19(2):15–40.
- Ulmer, Jeffery T and Brian Johnson. 2004. "Sentencing in Context: A Multilevel Analysis." *Criminology* 42(1):137–178.

- 
- Wang, Yuhua and Carl Minzner. 2015. "The Rise of China's Security State." *The China Quarterly* (222):339–359.
- Yang, Dali L. 1993. "Illegal Drugs, Policy Change, and State Power: The Case of Contemporary China." *Journal of Contemporary China* (4):14–34.
- Yang, Dali L. 2017. "China's Troubled Quest for Order: Leadership, Organization and the Contradictions of the Stability Maintenance Regime." *Journal of Contemporary China* 26(103):35–53.
- Yin, Mingcan and Xiaoming Li. 2009. "Empirical Analysis of Homicide (Guyi Sharenzui Shizheng Yanjiu)." *Chinese Criminal Science (Zhongguo Xingshifa Zazhi)* 18(1).
- Zhang, Kun and Yanming Li. 2010. "Comparing Foreign Drug Infiltration in Yunnan and Xinjiang (Yunnan he Xinjiang jinwai dupin shentoude weihaixing bijiao)." *Journal of Yunnan Police Officer Academy (Yunnan jinguan xueyuan xuebao)* 6(6).

---

## **A Supporting Information**

The Supporting Information includes the following materials:

### **A.1 Data Collection and Data Parsing**

Table A1: Variable Definitions and Descriptive Statistics

Data Parsing Procedure

### **A.2 Missingness and Data Quality**

Figure A2: Drug Cases in China Judgments Online

### **A.3 Robustness Checks and Additional Analyses**

Table A2: Effect of Minority Status on Sentencing (Yunnan) - Outliers Excluded

Figure A3: Drug Quantities, Ethnicity, and Sentencing Outcomes (All Provinces)

Table A3: Estimates of Effect of Minority Status on Sentencing (All Provinces)

Figure A4: Robustness of Minority Group Estimates (1/2)

Figure A5: Robustness of Minority Group Estimates (2/2)

### **A.4 China's Criminal Law**

Criminal Law of the People's Republic of China: Section 7: Crimes of Smuggling, Trafficking in, Transporting and Manufacturing Narcotic Drugs

## A.1 Data Collection and Parsing

Table A1: Variable Definitions and Descriptive Statistics

Variable	Description	Mean
<i>pun.severity</i>	fixed imprisonment length (months)	36.4
<i>pun.lifedeath</i>	sentenced to death/life imprisonment (indicator)	0.013
<i>def.minority</i>	defendant minority (indicator)	0.095
<i>def.age</i>	defendant age (years)	36.9
<i>def.female</i>	defendant female (indicator)	0.15
<i>def.mental</i>	defendant mentally handicapped (indicator)	0.00
<i>def.recid</i>	defendant recidivist (indicator)	0.32
<i>drug.pooled.quantity</i>	quantity of heroin + meth (grams)	48.9
<i>international</i>	crime transnational in nature (indicator)	0.00
<i>minors</i>	crime involves minors (indicator)	0.00
<i>pleadnotguilty</i>	defendant pleads innocent (indicator)	0.00
<i>goodattitude</i>	defendant deemed cooperative (indicator)	0.95
<i>lawyer</i>	defendant had attorney present (indicator)	0.19
<i>crime.act</i>	possession, smuggling, trafficking, transport	NA
<i>court.name</i>	court (fixed effects)	NA
<i>ruling.year</i>	year of case ruling (fixed effects)	NA

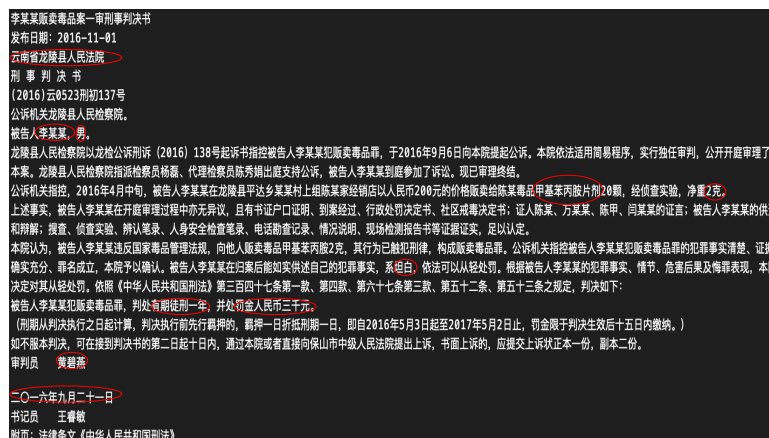
Note: Table shows definitions and descriptive statistics for all variables used in core regression analysis. Data come from full filtered dataset, which includes all heroin/methamphetamine cases in the [wenshu.court.gov.cn](http://wenshu.court.gov.cn) website from 2014-2018 for all provinces.

## Data Parsing Procedure

We processed the raw and unstructured text files of court ruling in the following steps. First, we sampled 1,449 drug cases country wide as our training set. Based on the Article 347 of the Criminal Law regarding drug crimes and our reading of these cases, we labeled court rulings manually into variable-based annotated data. We then developed a parsing algorithm of rule-based regular expressions to explore patterns in these court documents, extract relevant information, and label our variables automatically. When the text was not structured as expected, we used named entity recognition (NER)—an extraction process that takes a string of text as input and identifies relevant variables (e.g., defendant name, birthday, and sex).<sup>11</sup>

Next, we sampled 300 cases from two provinces and used them as our validation set. We trained the algorithm until most of the binary variables reached an accuracy rate of 95% as measured by an F1 score for both the training set and the validation set. We then applied the algorithm to the rest of the Yunnan drug cases and the countrywide dataset. Figure A1 is an example of how the parsing algorithm extracts relevant information from an unstructured court document.

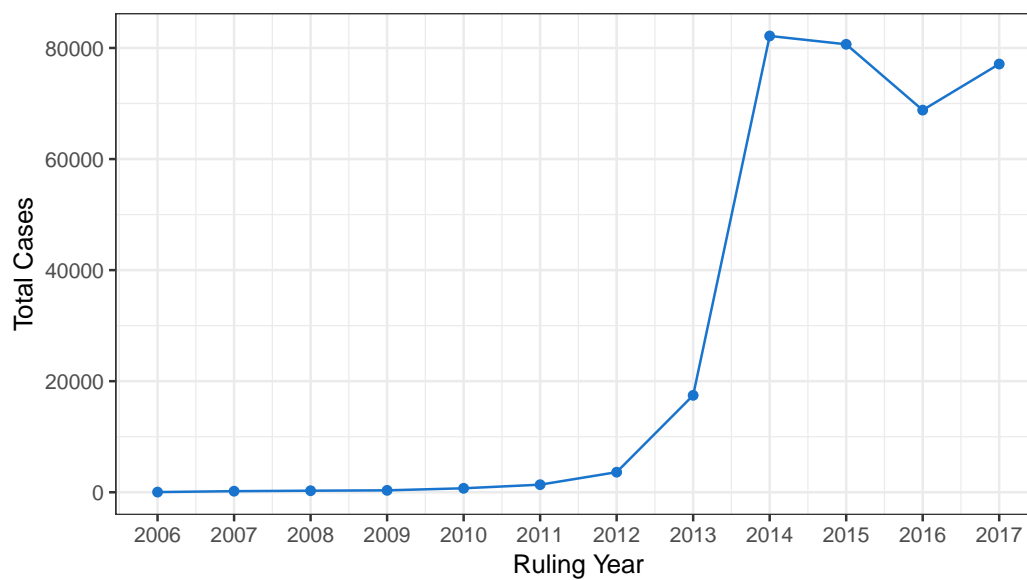
Figure A1: Parsing Example



<sup>11</sup>For more on NER, see the Stanford Natural Language Processing Group webpage at <https://nlp.stanford.edu/software/CRF-NER.html>.

## A.2 Missingness and Data Quality

Figure A2: Drug Cases in China Judgments Online



Note: Figure shows the total number of drug cases in the China Judgements Online database. The case coverage appears to improve dramatically in 2014; the analysis in the paper is therefore limited to cases from 2014-2018.

### A.3 Robustness Checks and Additional Analyses

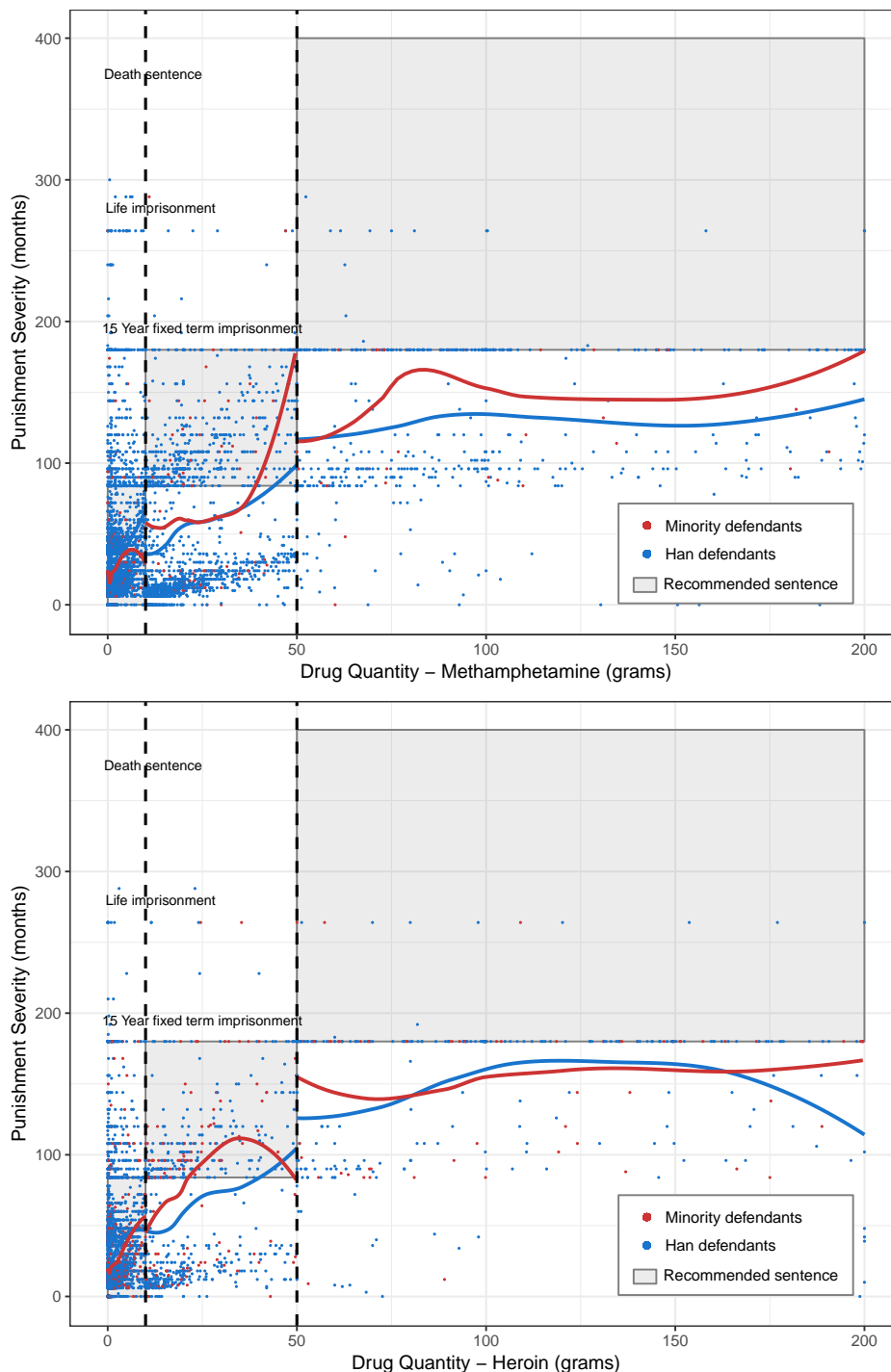
Table A2: Effect of Minority Status on Sentencing (Yunnan) - Outliers Excluded

#	Covariates	Outcome	
		<i>pun.severity</i>	<i>pun.lifedeath</i>
M1.	<i>meth.amt</i> + <i>heroin.amt</i> + <i>cocaine.amt</i> + <i>marijuana.amt</i> + <i>other.amt</i>	-6.645 (1.519)	0.002 (0.004)
M2.	M1. + <i>crime.act</i> + <i>international</i> + <i>minors</i> + <i>def.recid</i>	5.367*** (0.978)	0.011*** (0.004)
M3.	M2. + <i>def.mental</i> + <i>def.female</i> + <i>def.age</i>	5.976*** (1.085)	0.0127** (0.005)
M4.	M3. + <i>amt</i> cubic splines	3.064*** (0.785)	0.016*** (0.005)
M5.	M4. + <i>ruling.year</i> (fixed effects) + <i>court.name</i> (fixed effects)	1.485** (0.861)	0.011** (0.005)
M6.	M5. + <i>def.pleadnotguilty</i> + <i>def.goodattitude</i>	1.482* (0.858)	0.011** (0.005)

Note: Table shows coefficient estimates from regressions of *pun.severity* (months) and *pun.lifedeath* and on *def.minority* across six different covariate sets. M4 is the “baseline specification” referred to throughout the paper. Data come from Yunnan filtered dataset, which includes all heroin/methamphetamine cases in the [wenshu.court.gov.cn](http://wenshu.court.gov.cn) website from 2014-2018. Analysis excludes defendants that were outliers on the drug quantity variable, 86 individuals in total. Robust standard errors shown in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



Figure A3: Drug Quantities, Ethnicity, and Sentencing Outcomes (All Provinces)



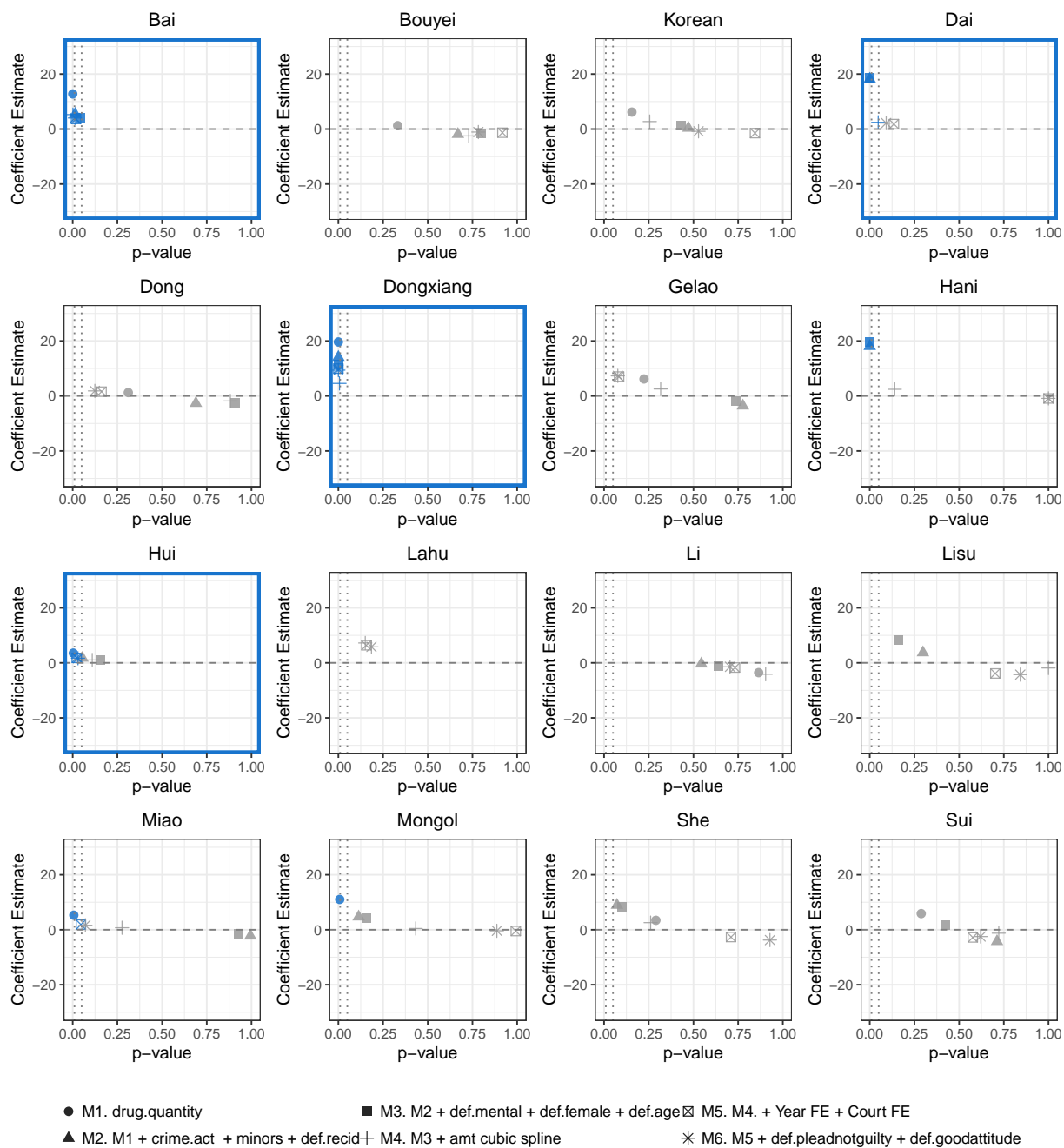
Note: Figure shows the drug amount of methamphetamine and heroin (grams) against the sentencing outcome (months) for 167,091 drug possession, transport, and trafficking cases from 2014-2018. The blue points and line correspond to defendants of the Han majority; red indicates defendants of an ethnic minority group. The shaded grey boxes show the recommended sentencing levels for the drug quantity according to China's Criminal Law. Points below those boxes indicate leniency; points above that line indicate severity. Figure shows a 5% sample of all cases for ease of visualization.

Table A3: Effect of Minority Status on Sentencing (All Provinces)

#	Covariates	Outcome	
		<i>pun.severity</i>	<i>pun.lifedeath</i>
M1.	<i>meth.amt</i> + <i>heroin.amt</i> + <i>cocaine.amt</i> + <i>marijuana.amt</i> + <i>other.amt</i>	8.452*** (0.515)	0.0148*** (0.001)
M2.	M1. + <i>crime.act</i> + <i>international</i> + <i>minors</i> + <i>def.recid</i>	2.240*** (0.363)	0.007*** (0.001)
M3.	M2. + <i>def.mental</i> + <i>def.female</i> + <i>def.age</i>	2.537*** (0.423)	0.007*** (0.001)
M4.	M3. + <i>amt</i> cubic splines	1.167*** (0.305)	0.006*** (0.001)
M5.	M4. + <i>ruling.year</i> (fixed effects) + <i>court.name</i> (fixed effects)	1.706*** (0.366)	0.004*** (0.001)
M6.	M5. + <i>def.pleadnotguilty</i> + <i>def.goodattitude</i>	1.599*** (0.362)	0.004*** (0.001)

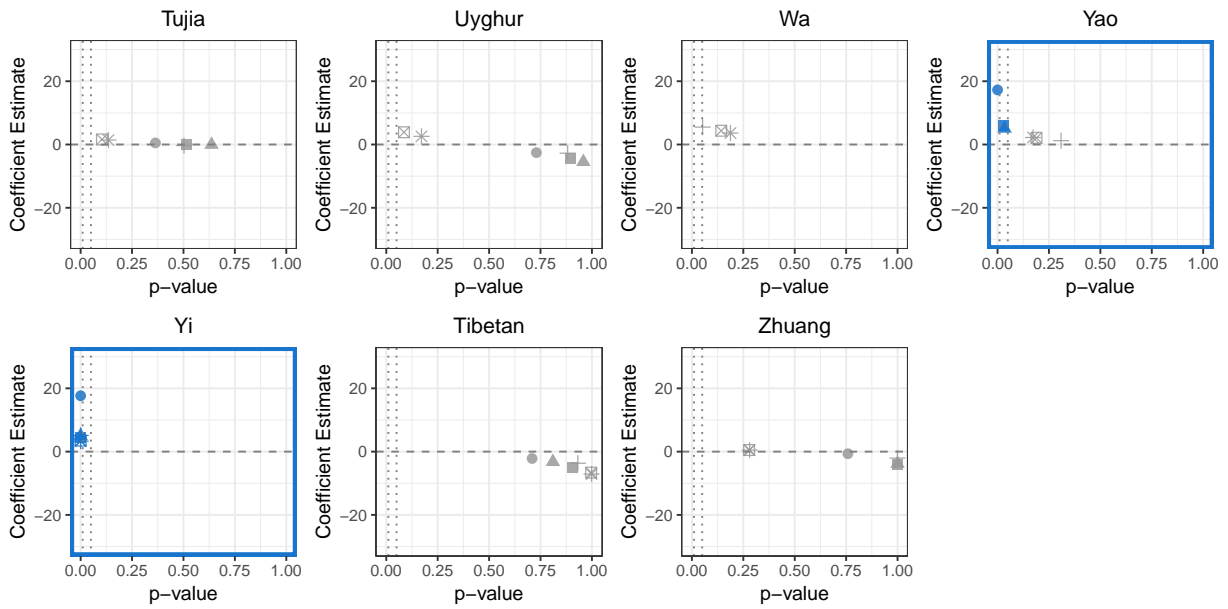
Note: Table shows coefficient estimates from regressions of *pun.severity* (months) and *pun.lifedeath* on *def.minority* across six different covariate sets. M4 is the “baseline specification” referred to throughout the paper. Data come from full filtered dataset, which includes all heroin/methamphetamine cases in the [wenshu.court.gov.cn](http://wenshu.court.gov.cn) website from 2014-2018. Robust standard errors shown in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Figure A4: Robustness of Minority Group Estimates (1/2)



Note: Figure shows coefficient estimates and p-values from a one sided hypothesis of no effect from regressions of *pun.severity* (months) on dummy variables for each minority group across six different covariate sets. Each facet shows the results for a different minority. M4 is the “baseline specification” referred to throughout the paper. Data come from filtered dataset, which includes all heroin/methamphetamine cases in the wenshu.court.gov.cn website from 2014-2018. Estimates shown in blue represents where  $p < 0.05$ , and groups highlighted in blue are those where there is robust evidence of discrimination.

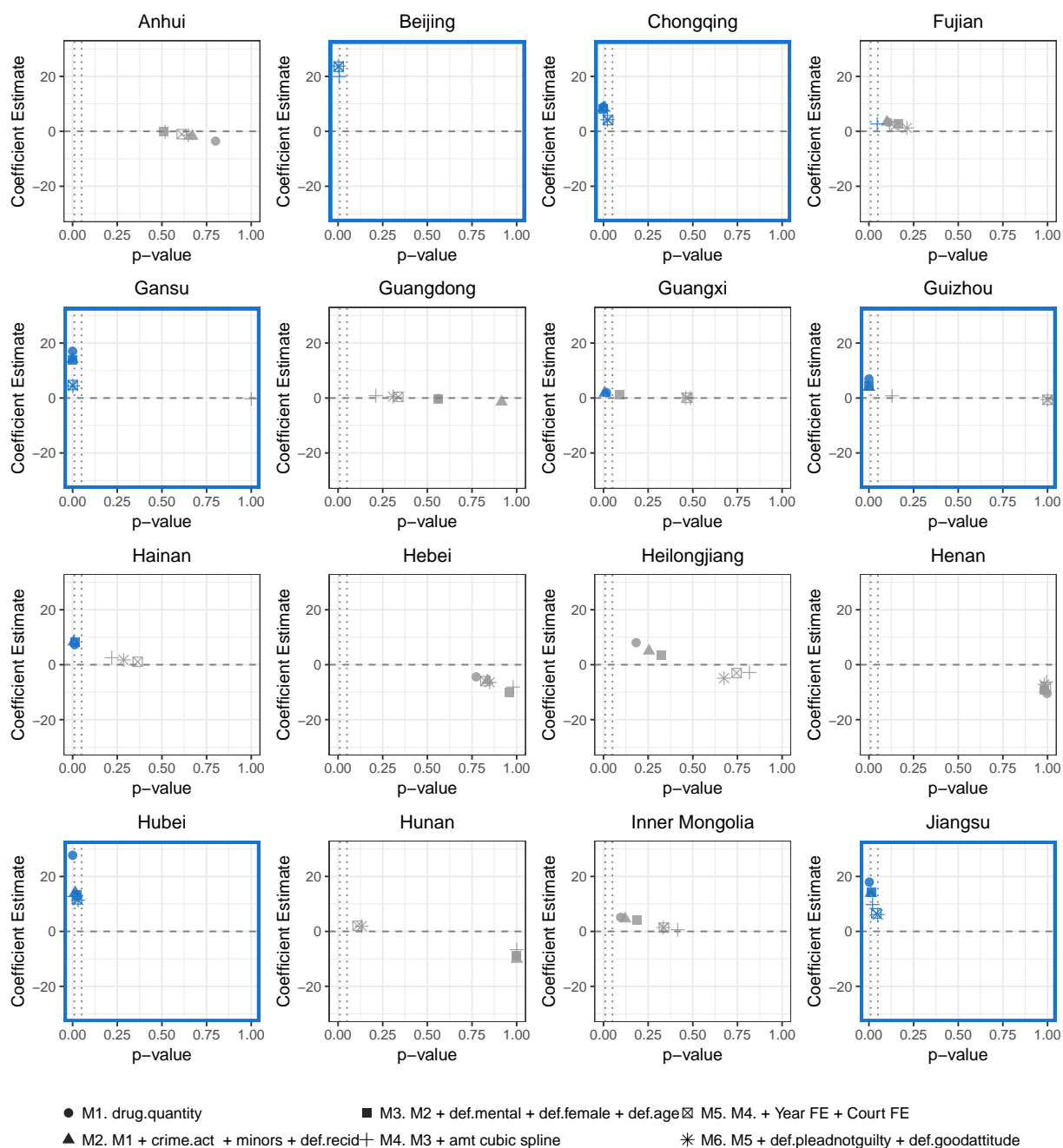
Figure A5: Robustness of Minority Group Estimates (2/2)



- M1. drug.quantity
- M3. M2 + def.mental + def.female + def.age
- ⊠ M5. M4. + Year FE + Court FE
- ▲ M2. M1 + crime.act + minors + def.recid
- ⊕ M4. M3 + amt cubic spline
- \* M6. M5 + def.pleadnotguilty + def.goodattitude

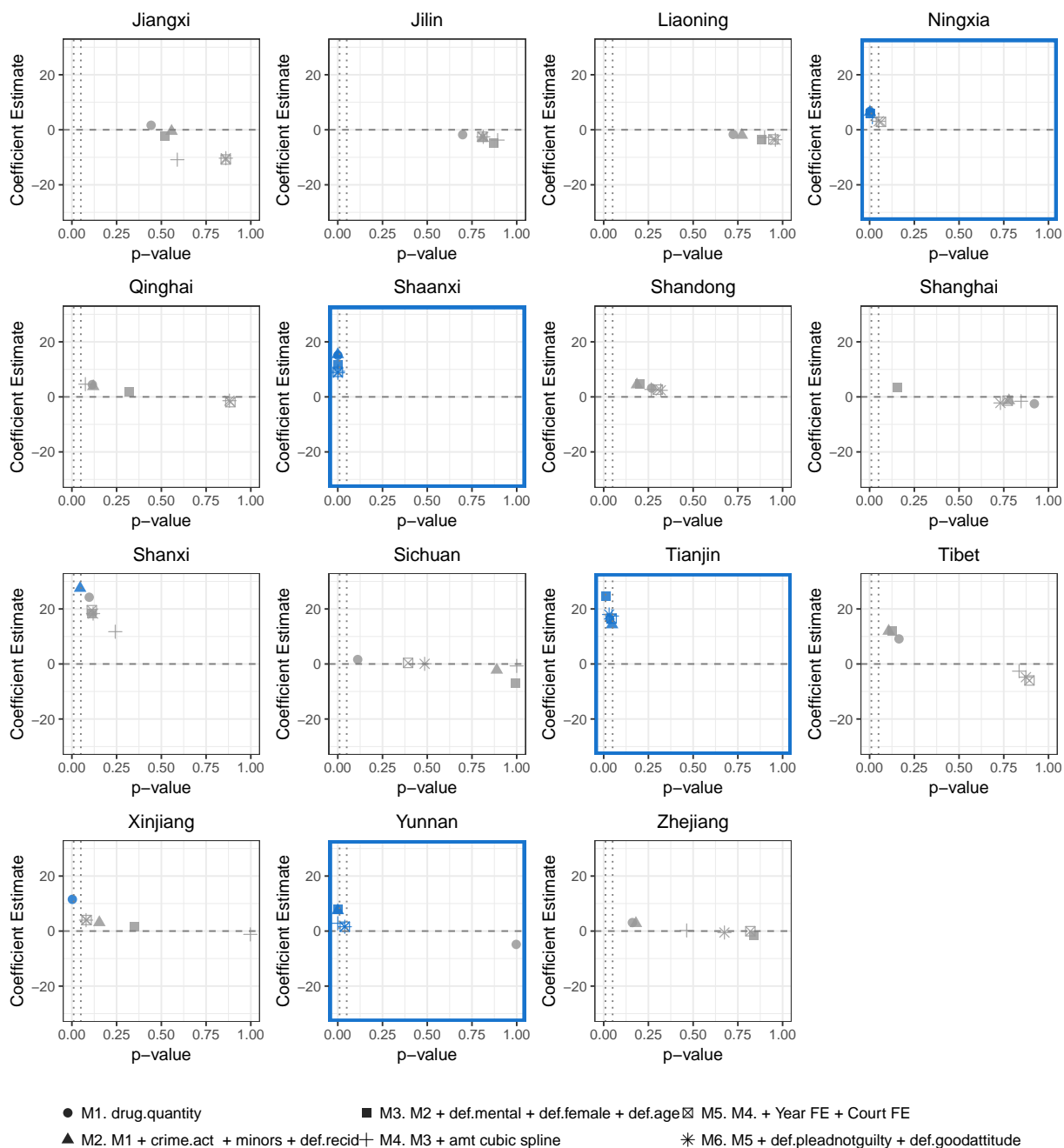
Note: Figure shows coefficient estimates and p-values from a one sided hypothesis of no effect from regressions of *pun.severity* (months) on dummy variables for each minority group across six different covariate sets. Each facet shows the results for a different minority. M4 is the ‘baseline specification’ referred to throughout the paper. Data come from filtered dataset, which includes all heroin/methamphetamine cases in the wenshu.court.gov.cn website from 2014-2018. Estimates shown in blue represents where  $p < 0.05$ , and groups highlighted in blue are those where there is robust evidence of discrimination.

Figure A6: Exploring Variation Across Provinces (1/2)



Note: Figure shows coefficient estimates and p-values from a one sided hypothesis of no effect from regressions of *pun.severity* (months) on *def.minority* across six different covariate sets. Each facet shows the results for a different province. M4 is the “baseline specification” referred to throughout the paper. Data come from filtered dataset, which includes all heroin/methamphetamine cases in the wenshu.court.gov.cn website from 2014-2018. Estimates shown in blue represents where  $p < 0.05$ , and provinces highlighted in blue are those where there is robust evidence of discrimination against minorities.

Figure A7: Exploring Variation Across Provinces (2/2)



Note: Figure shows coefficient estimates and p-values from a one sided hypothesis of no effect from regressions of *pun.severity* (months) on *def.minority* across six different covariate sets. Each facet shows the results for a different province. M4 is the “baseline specification” referred to throughout the paper. Data come from filtered dataset, which includes all heroin/methamphetamine cases in the wenshu.court.gov.cn website from 2014-2018. Estimates shown in blue represents where  $p < 0.05$ , and provinces highlighted in blue are those where there is robust evidence of discrimination against minorities.

## A.4 China's Criminal Law

### Section 7 Crimes of Smuggling, Trafficking in, Transporting and Manufacturing Narcotic Drugs

Article 347: Whoever smuggles, traffics in, transports or manufactures narcotic drugs, regardless of the quantity involved, shall be investigated for criminal responsibility and given criminal punishment.

Whoever smuggles, traffics in, transports or manufactures narcotic drugs and falls under any of the following categories, shall be sentenced to fixed-term imprisonment of 15 years, life imprisonment or death and also to confiscation of property:

1. persons who smuggle, traffic in, transport or manufacture opium of not less than 1,000 grams, heroin or methamphetamine of not less than 50 grams or other narcotic drugs of large quantities;
2. ringleaders of gangs engaged in smuggling, trafficking in, transporting or manufacturing narcotic drugs;
3. persons who shield with arms the smuggling, trafficking in, transporting or manufacturing of narcotic drugs;
4. persons who violently resist inspection, detention or arrest to a serious extent; or
5. persons involved in organized international drug trafficking.

Whoever smuggles, traffics in, transports or manufactures opium of not less than 200 grams but less than 1,000 grams, or heroin or methamphetamine of not less than 10 grams but less than 50 grams or any other narcotic drugs of relatively large quantities shall be sentenced to fixed-term imprisonment of not less than seven years and shall also be fined.

Whoever smuggles, traffics in, transports or manufactures opium of less than 200 grams, or heroin or methamphetamine of less than 10 grams or any other narcotic drugs of small quantities shall be sentenced to fixed-term imprisonment of not more than three years, criminal detention or public surveillance and shall also be fined; if the circumstances are serious, he shall be sentenced to fixed-term imprisonment of not less than three years but not more than seven years and shall also be fined.

Where a unit commits any crime mentioned in the preceding three paragraphs, it shall be fined, and the persons who are directly in charge and the other persons who are directly responsible for the offense shall be punished in accordance with the provisions of the preceding three paragraphs respectively.

Whoever makes use of minors or aids and abets them to smuggle, traffic in, transport or manufacture narcotic drugs or sells narcotic drugs to minors shall be given a heavier punishment.

With respect to persons who have repeatedly smuggled, trafficked in, transported or manufactured narcotic drugs and have not been dealt with, the quantity of narcotic drugs thus involved shall be computed cumulatively.

Article 348: Whoever illegally possesses opium of not less than 1,000 grams, or heroin or methamphetamine of not less than 50 grams, or any other narcotic drugs of large quantities shall be

sentenced to fixed-term imprisonment of not less than seven years or life imprisonment and shall also be fined; whoever illegally possesses opium of not less than 200 grams but less than 1,000 grams, or heroin or methamphetamine of not less than 10 grams but less than 50 grams or any other narcotic drugs of relatively large quantities shall be sentenced to fixed term imprisonment of not more than three years, criminal detention or public surveillance and shall also be fined; if the circumstances are serious, he shall be sentenced to fixed-term imprisonment of not less than three years but not more than seven years and shall also be fined.