Does China Achieve the Goal of Treating Like Cases Alike? Evidence from the Judicial Documents of China[1]

Author：Mingyang Chen
(CUPL)

Abstract

Treating like cases alike is a chain that starts with the facts of the case, followed by reasoning, and then the verdict, which is a common pursuit among legal professionals. This study focuses on the variable of reasoning, examining whether the effect of consistent judgment for similar cases exists in practice. By employing NLP, the study uses DTM, Cosine Similarity, and Euclidean distance methods to analysis 523 judicial documents from China. Three quantitative models are then constructed to explore the relationship between these three elements. The study finds that: 1. The higher the case similarity, the higher the reasoning similarity. 2. There is no significant correlation between reasoning and sentences differences. 3. There is also no significant correlation between case similarity and sentences differences. Overall, Chinese judicial practice has not yet achieved the effect of consistent judgment for similar cases. That is a result of professional training as a judge and extralegal factors. This paper puts forward a new method to estimate biased judgements.

[1] Contact me at cmy0211@yeah.net to get the programs and data of this paper.

INTRODUCTION

"Treating like cases alike" is always a dream of legal professionals. It makes the law stable and predictable, and plays an important role in justice. It is a fundamental requirement of formal justice in the field of judgement.[1] In July 2020, the Supreme Court of China implemented the "Guiding Opinions on Strengthening Case Retrieval to Unify the Application of Law" (see at figure C1)[2]. This theory and policy used to raise a hot topic in both China and abroad. However, although there are many discussions on it and scholars achieve an agreement on the importance of this theory, we still do not know whether "treating like cases alike" is achieved and to what extend China achieves the goal. This essay is going to use text analysis in order to answer these questions: 1. Is treating like cases alike only a dream up in the air or has already been realistic in China? 2. To what degree China's judgement achieve the goal of treating like cases alike? In the past, some researchers tried to answer the 2 questions by traditional legal study method, but the conclusion is not reliable due to their biased sampling. In order to get a stable conclusion, I will use hundreds of judicial documents in China, and use basic text analysis as a tool to investigate these questions.[3] This paper only conducts descriptive research to show how China's judges' rough review. but it is still meaningful because we can now answer the question that Chinese people care about.

LITERATURE REVIEW

Theory: Treating like cases alike

Hart believes that the justice itself could be various, but the most important rule for justice is every 2 individuals should be treated equally under the same justice. This is the origin of the principle called "treating like cases alike". Hart insists that "treating like cases alike " is an administration of judgement, not the justice itself. As long as the Judge do right verdicts, then like cases would be treated alike.[4] We can find in his earlier essay in 1985 that Hart has already declared this opinion. "This (treating like cases alike) is justice in the administration of the law, not justice of the law…implement just this aspect of law and which are designed to ensure that rules are applied only to what are genuinely cases of the rule or at least to minimize the risks of in-equalities in this sense."[5] Timothy J. Capurso also holds a view by analyzing judicial decision theory that if Judge X will decide a certain case one way, while Judge Y would decide that same case in another way, it may not be advantageous.[6] Later scholars developed a series of reasons to support the theory including "Unity of Legal Interpretation of Rules and Order", "Prevention of Arbitrariness", "Predictability of Punishment", "Requirement of Legal Stability", "Publicity and Authority Assurance of Judgments", "Protection of Expectant Interests" and etc.[7] However, Andrei Marmor points out treating like cases

[1] Haibo Sun. In What Sense Does Case Retrieval Assist in Achieving Consistent Judgments for Similar Cases? Tsinghua Law Review. Vol. 15, Iss. 1, pp. 79-97 (2021).

[2] This is a policy promoted by China's Supreme Court. The Chinese name for this policy is "关于统一法律适用加强类案检索的指导意见". I will show the policy website in appendix C.

[3] Welbers, Kasper and Van Atteveldt, Wouter and Benoit, Kenneth. Text analysis in R. Communication Methods and Measures, Vol. 11, No. 4, pp. 245-265 (2017) ; Kosuke Imai. Quantitative Social Science: An Introduction. Princeton University Press, 2017.

[4] H. L. A. Hart, The Concept of Law, Oxford: Clarendon Press, 1996, pp. 155-161.

[5] H. L. A. Hart, Positivism and the Separation of Law and Morals. Harvard Law Review, Vol. 71, No. 4, pp. 593-629 (1958).

[6] Timothy J. Capurso. How Judges Judge: Theories on Judicial Decision Making. University of Baltimore Law Forum, Vol. 29, No. 01, Article 2 (1998).

[7] Kenneth I. Winston. On Treating Like Cases Alike, The Independent Review, Vol. 4, No. 1, pp. 107-118 (1999) ; Andrei Marmor. Should Like Cases Be Treated Alike, Legal Theory, Vol. 11, Iss. 1, pp. 27-38 (2005) ; Jinghui Chen. Same Case, Same Judgment: A Legal Duty or a Moral Requirement? China Legal Science, Iss. 3, pp. 46-61 (2013).

alike is not any conclusive answer. If for the same reasons above, to betray the principle of treating like cases alike is allowed.[8] But in most situation, it can raise a warning flag about possible problems when they are treated differently, or as demanding an explanation for differences in treatment, the principle is sound, and important. As a result, although treating like cases alike has some disadvantages, it is a common practice, and is called a "weak claim" according to Chen.[9] The international society has already received this theory, e.g., the International Association of Refugee Law Judge in 1990s.[10]

Treating like cases alike differs in different countries. As we want to build a model of China's situation, so we need to have a review of China's researches in order to build a theory frame. The theory is considered as a deduce of equality and used to raise a discussion in China, especially the concept and the importance.[11] Among those who disagree the theory, Zhou is a stand-out. He argues treating like cases alike is a fictional legal myth, because it starts with applying the law, while pursues the same result. If the Judge can apply the law correctly, then we do not have to limit their verdict by this theory.[12] Sun then published an article to support the opposite. Judges can sometimes apply the law correctly but get a wrong conclusion, so the theory would play a role when this situation occurs. Treating like cases alike is not only a myth, but a tool to ensure the predacity of judgements, the protection for expectations and a limit to discretion.[13] Then Zhou admitted the value of treating like cases alike, believing we need to use it as a tool and should never forget that different cases sometimes account as well.[14] Then Lei gave an end to this topic by clearing the concept and believe that treating like cases alike should be a goal of China's judgement process.[15]

But what we really want by saying "treating like cases alike"? Christopher Sherri and Avani Mehta Sood's essay could answer that. An inconsistent verdict should face reasonable doubts so they need to be open access in order to prevent wrongful verdict. That is to say, a special verdict should write reason fully to deal with the doubts.[16] Chinese scholars have similar view. They believe treating like cases alike is a combination of reason writing and the verdict conclusion.[17] In conclusion, there should be same reason for same sentences.[18]

Related policy: China's official's attitude

[8] Andrei Marmor. Should Like Cases Be Treated Alike, Legal Theory, Vol. 11, Iss. 1, pp. 27-38 (2005).
[9] Jinghui Chen. Same Case, Same Judgment: A Legal Duty or a Moral Requirement? China Legal Science, Iss. 3, pp. 46-61 (2013).
[10] Hugo Storey. Consistency in Refugee Decision-Making: A Judical Perspective, Refugee Survey Quarterly, Vol. 32, Iss. 4, pp. 112-125 (2013).
[11] Jianjun Bai. The Constitutional Significance and Empirical Research of Consistency in Judgment for Similar Cases. China Legal Science, Vol. 32, No. 3, pp. 131-140 (2003).
[12] Shaohua Zhou. Same Case, Same Judgment: A Fictional Myth of the Rule of Law. Law Science Magazine, Vol. 32, No. 11, pp. 131-140 (2015) ; Shaohua Zhou. A Rational Examination of 'Same Case, Same Judgment' in Criminal Cases. Legal Science, Vol. 37, No. 3, pp. 3-15 (2020) .
[13] Haibo Sun. The 'Same Case, Same Judgment' Principle: Not a Fictional Myth of the Rule of Law. Jurists Review, Vol. 5, pp. 141-157 (2019) ; Haibo Sun. Should Similar Cases Be Treated Similarly? . Legal System and Society Development, Vol. 25, No. 3 (2019) ; Qi Zhang. On the Similar Treatment of Similar Cases. Global Legal Review, Vol. 36, No. 1, pp. 21-34 (2014) .
[14] Shaohua Zhou. Differentiated Judgments in Criminal Cases and Their Rationality. China Legal Science, Vol. 4, pp. 145-164 (2019).
[15] Lei Lei. How to Understand 'Same Case, Same Judgment'?—Misconceptions and Clarifications. Political Science and Law Review, Vol. 5, pp. 28-38 (2020).
[16] Christopher Sherrin. Inconsistent Verdicts and the Possibility of Innocence: A Comment on R v RV, Wrongful Conviction Law Review, Vol. 2, No. 1, pp. 78-81 (2021) ; Avani Mehta Sood. What's So Special About General Verdicts? Questioning the Preferred Verdict Format in American Criminal Jury Trials, Theoretical Inquiries in Law, Vol. 22, No. 2, pp. 55-84 (2021).
[17] Shude Liu. The principle of 'Same Case, Same Judgment' in the Context of Criminal Justice. China Legal Science, Vol. 1, pp. 68-76 (2011).
[18] Andrei Marmor. Should Like Cases Be Treated Alike. Legal Theory, Vol. 11, Iss. 1, pp. 27-38 (2005).

The official of China also holds same view we reviewed before. In 2011, there is a policy called "Guiding Case System" (see at figure C2) [19] aiming to promote the goal of treating like cases alike.[20] Chen comments that this rule can be a stimulation for China's judgement's justice. What is more, recently, AI and big data is a shock to the judgement system. In order to apply those in the court, we need to accept the theory. So, the Supreme Court in China promote a new policy (see at figure C3) [21], Opinions on Speeding up the Development of Intelligent Courts by the Supreme People's Court, aiming to use machine learning and deep learning technologies to meet the precise demands of case handlers for legal, case, and professional knowledge, and to promote the standardization of treating like cases alike and sentencing by Judges. By applying those tools, it can provide new solutions for difficult cases, and also unify the standards of judicial adjudication to prevent unfair judicial decisions and achieve the goal of treating like cases alike. 3 years later, the Supreme Court of China again attach great importance to the "Guiding Opinions on Strengthening Case Retrieval to Unify the Application of Law". This time by saying "The case retrieval report must be objective, comprehensive, and precise, encompassing the following elements: …It should also detail the critical points of judgment from analogous cases and the contentious issues at the heart of the pending case. Furthermore, the report should include an analysis explaining the application of such case findings, whether they are to be considered as precedents or merely as references", the Supreme Court emphasizes that Judges should also focus on reason writing. We can confidently say treating like cases alike is such a frame after the review of prior researches and the policy in China: the similar cases share similar reason writing and then have similar verdict conclusion.

So, the overall pattern of the judiciary should be: most judgments conform to the theory of treating like cases alike and only a small number of outlier cases do not follow this pattern. These outlier cases can be extracted and examined individually for their rationality. If China's judicial judgments conform to this pattern, then we can confidently say that China's judicial judgments have essentially achieved the general normative goal of treating like cases alike.

Theory frame and hypnosis

According to the theory treating like cases alike, as well as factors influencing reason writing and sentencing, I divide treating like cases alike into a 3-part chain and a 2-part factors subset. The chain is case—reason—sentences. This is a rough frame as I discussed before. the following figure 1.1 illustrate this.

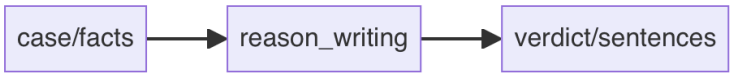


Figure 1.1

[19] This is a policy promoted by China's Supreme Court. The Chinese name for this policy is "最高人民法院关于案例指导工作的规定". I will show the policy website in my appendix C.

[20] See Xingliang Chen. China'S Guiding Case System : A Study on The Mechanisms Of Rule Formation. Peking University Law Journal, Vol 1, No 2, pp. 215-258 (2015); Aiwu Bao. A Watched Flower Never Blooms — A Study of Guiding Cases System and "Treating Like Cases Alike". Doctoral Thesis, City University of Hong Kong, 2020.

[21] This is a policy promoted by China's Supreme Court. The English name for this policy is "The Supreme People's Court's Opinions on Accelerating the Construction of Smart Courts". The Chinese name for this policy is "最高人民法院关于案例指导工作的规定". I will show the policy website in my appendix C.

The 2-part factors mean the factors affect the chain. I get this frame from the concept "law in action". Law in action means law is not only about the text, but also something beyond that. Actions of Judges and other officers make up what is called law, meaning there are more things outside the text.[22] So, we need to figure out what is in text and how can my analysis work. I will then take Peczenik's theory to illustrate. According to Peczenik, the source of law could be divided into 3 kinds: Must-Sources, Should-Sources and May-Sources. all-things-considered principal claim that the later 2 factors should be taken into consideration when judging so it is reasonable that Judges use Should-Sources and May-Sources in their reason writing.[23] Besides, there are extralegal factors as well which are not observed in documents. So, I divide those factors by whether it could be observed in a document. One part is the factors that could be observed in judicial documents, e.g., forgiveness from victims and plead guilty. Some of that can be quantitative like in Wu's research.[24] Those factors are written in China's criminal law. In Peczenik's words, they are so-called "binding factors". In order to estimate these factors, Wu uses the traditional quantitative method picking up factors in documents and assigning a dummy value to those factors. The tool of text analysis also works here but not so precise. While some cannot be quantitative, e.g., moral rules, or according to Peczenik, calling them should- or may- sources of law. Judges will use these to get a conclusion and that makes sense according to Peczenik. Chen and Cheng also have discussed this issue by researching on cases in China and recommended judges to open these to the public.[25] For example, some Judges would consider the social impact of the case as a factor on the sentences and even write this into documents. These factors can only be estimated by text analysis method. However, the 2 types of factors can both be estimated by text analysis, because we can only focus on the similarity of texts, instead transferring all factors into numbers by the traditional method. I will discuss this later when talking about NLP. The other part is factors outside the documents, e.g., genders and races. Later I will show some researches on these. These factors can be tested by other methods like experiment method and etc.[26] The second theory frame is showed in figure 1.2.

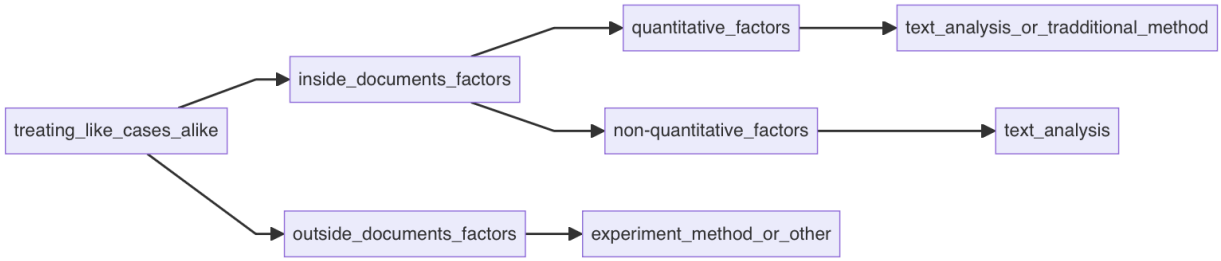


Figure 1.2

[22] Karl N. Llewellyn. The Bramble Bush: on our law and its study. Quid Pro Books, 2012, p.10.
[23] Aleksander Peczenik. On Law and Reason. Springer, 2008, pp. 62-63, 261-263.
[24] Yuhao Wu. The Boundaries of Sentencing Discretion: Collective Experience, Individual Decision-Making, and Bias Identification. Chinese Journal of Law Studies, Vol. 43, No. 6, pp. 109-129 (2021).
[25] Liang Chen, Jinhua Cheng. How to Import Moral Judgment into Judicial Adjudication. Exploration and Free Views, No. 08, pp. 59-72, 178 (2023).
[26] Daniel M. Schneider. Empirical Research on Judicial Reasoning: StatutoryEmpirical Research on Judicial Reasoning: Statutory Interpretation in Federal Tax Cases, New Mexico Law Review, Vol. 31, Iss. 2, pp. 325-258 (2001) ; John Zhuang Liu, Xueyao Li. Legal Techniques for Rationalizing Biased Judicial Decisions: Evidence from Experiments with Real Judge. Journal of Empirical Legal Studies, Vol. 16, Iss. 3, pp. 630–670 (2019).

Prior Empirical studies have investigated substantial factors thar could have an effect on Judges' verdict. Races, genders, ages and education level could be factors influencing judgement, and lots researches, involving from the psychology field and law field, has done a lot on those factors.[27] Because it is impossible to estimate factors outside documents by only using documents as a sample, I will not do research on these factors in this article. As for factors inside documents, according to Bai's study, there is still a considerable scope where the application of the law has not achieved the same judgment for the same case and the same penalty for the same amount, but rather different judgments for the same case and different penalties for the same amount.[28] Bai and Wu even developed a model for verdict based on so-called "Judges' experiences". Through Wu's models he tries to find which judge's verdict is off-path, meaning it may be an incorrect one.[29] But let us see their conclusions from the first theory frame first, which is the 3-part chain. We can find scholars focus more on the verdict conclusion, or the sentences in the field of criminal law only. Both outside and inside documents factors we mentioned before is used to build a correlation with the sentences. The reason writing part is usually ignored. Almost no scholars can find the correlation between cases and reason writing. Then let us look at these researches from the second theory frame's view. Traditional quantitative method can only take a portion of factors into consideration. That means the result is not stable enough. Applying text analysis method is a better choice.

I want to mention that the cases and reason writing are all texts, so NLP would be an excellent tool. If cases' text similarity is higher, then according to my theory frame, the text similarity of reason writing would be higher as well. Besides, the conclusion part of a judicial documents would be more similar too.

Applying NLP into judicial documents

Legal documents emphasize "legal terminology", meaning they use same words and sentences to convey same meanings. While applying NLP to analysis texts, the biggest problems would be the diversity of meaning expression, emotional connotations, and the sentences structures. They contribute to a phenomenon that similar texts would express totally different meanings. But because the "legal terminology", especially in China, a member of civil law family, the words Judge use in the texts are stable, meaning it would convey same meanings if they use same words or sentences. Take China's judicial documents as an example, if the judge use the words "*zhuguan'exing*" ("主观恶性", meaning "subjective malice" in Chinese), then what he wants to convey is what he writes. Briefly speaking, Judges have already tidied the judicial texts for us researchers so we can use NLP to get a roughly correct outcome. So as a result, we can image: assume we have an AC case and AR

[27] Jennifer Mayer Cox, Naomi E. S. Goldstein, John Dolores, Amanda D. Zelechoski & Sharon Messenheimer. The Impact of Juveniles' Ages and Levels of Psychosocial Maturity on Judges' Opinions About Adjudicative Competence. Law and Human Behavior, pp. 1-9 (2010) ; Christina L. Boyd, Lee Epstein, Andrew D. Martin. Untangling the Causal Effects of Sex on Judging. American Journal of Political Science, Vol. 54, Iss. 2, pp. 389-411 (2010) ; Shai Danziger, Jonathan Levav, and Liora Avnaim-Pesso. Extraneous factors in judicial decisions. Proceedings of the National Academy of Sciences of the United States of America. Vol. 108, No. 17, pp. 6889-6892 (2011) ; Allison P. Harris, Maya Sen. Bias and Judging. Bias and Judging. Annual Review of Political Science, Vol. 22, pp. 241-259 (2019) ; Shuai Wei, Moulin Xiong. Judges' Gender and Sentencing in China: An Empirical Inquiry. Feminist Criminology, Vol. 15, Iss. 2, pp. 217-250 (2020) ; Erik Voeten. Gender and judging: evidence from the European Court of human rights. Journal of European Public Policy, Vol. 28, Iss. 9, pp. 1453-1473 (2020).

[28] Jianjun Bai. The Constitutional Significance and Empirical Study of the Principle of Consistency in Judgment for Similar Cases. China Legal Science, Vol. 3, pp. 131-140 (2003).

[29] Jianjun Bai. Research on Sentencing Prediction Based on the Collective Experience of Judges. Chinese Journal of Law Studies, Vol. 38, No. 6, pp. 140-154 (2016) ; Yuhao Wu. The Boundaries of Sentencing Discretion: Collective Experience, Individual Decision-Making, and Bias Identification. Chinese Journal of Law Studies, Vol. 43, No. 6, pp. 109-129 (2021).

reason writing, a BC case and a BR reason writing, a CC case and a CR reason writing. If AC is more similar to BC, then the BR would be more similar to AR than CR.

This method is reliable because in CS field many experts apply this in order to train AI to predict the sentences. Kosuke Ima develop a question of whether constitution laws in the world are more similar with America constitution law by text analysis based on David S. Law and etc. and Zachary Elkins' research.[30] Haoxi Zhong then points out legal documents are text data of high quality. They use China's judicial documents to develop a model for sentences prediction.[31] Olga Alejandra Alcántara Francia and etc. have a review on text mining in the field of law study, commenting that this tool is a practical one to predict the sentences.[32] Similar to my research, Benjamin M. Chen and his cooperators use BLAST to detect the influence of the Chinese guiding cases. [33]These researches prove that NLP is a reliable tool in legal study.

What is more, NLP possesses an advantage that traditional method lacks. That is, it can take all factors inside documents into consideration just as I mentioned before. But why? All factors picked up by traditional method is based on key words. The documents are tokenized and we ask our computer to find key words like "*zhuguan'exing*" I mentioned before, and give it a dummy value. NLP is similar but it will give all words a value. The value depends on the method used. So, by giving all words a value and estimate 2 texts' similarity, we can take all factors from the documents into analysis. Although there will be some noises, we have to accept it and in my method the noise would be minimized. I will discuss it later.

METHODOLOGY

Data

This research uses 523 judicial documents in China. I collect those documents randomly from the China Judgment Online (CJO). All documents are published in 2019 and 2020. Due to my technology limitation I do not have enough samples, but they are enough to meet the purpose of my research.

1. Sample: China's Judicial Documents

I choose the robbery as my sample. If there is a joint crime, I will extract the defendant only. If it is a multiple defendants' trial, I will not take it into my sample. That would make cases easier to be analyzed. The reasons for choosing robbery are: 1. Robbery is a high-incidence crime and seldom use summary procedure making the contain of documents is rich. Compared with other crime, e.g., larceny and reckless driving, the texts of robbery are longer and more suitable to be analyzed. 2. Robbery according to China's criminal law, have many aggravating circumstances e.g., robbery resulting in death or armed robbery. What is more in China's law, robbery also has a special form known as "transformative robbery," which is stipulated in Article 269 of the criminal law. It states

[30] David S. Law, Mila Versteeg. The Declining Influence of the United States Constitution, New York University Law Review, Vol. 87, No. 3 (2012) ; Zachary Elkins, Tom Ginsburg, James Melton. Comments on Law and Versteeg's The Declining Influence of the United States Constitution, New York University Law Review, Vol. 87, No. 6, pp. 2088-2101 (2012) ; Kosuke Imai. Quantitative Social Science: An Introduction. Princeton University Press, 2017, pp. 236-238.

[31] Haoxi Zhong, Guo Zhipeng, Cunchao Tu, Chaojun Xiao, Zhiyuan Liu, Maosong Sun. Legal Judgment Prediction via Topological Learning; Zikun Hu, Xiang Li, Cunchao Tu, Zhiyuan Liu, Maosong Sun. Few-Shot Charge Prediction with Discriminative Legal Attributes. COLING 2018.

[32] Olga Alejandra Alcántara Francia, Miguel Nunez-del-Prado, Hugo Alatrista-Salas. Survey of Text Mining Techniques Applied to Judicial Decisions Prediction. Applied Sciences, Vol. 12, No. 20 (2022).

[33] Benjamin M. Chen, Zhiyu Li, David Cai, Elliott Ash. Detecting the influence of the Chinese guiding cases: a text reuse approach. Artificial Intelligence and Law (2023).

that “a person who commits theft, fraud, or snatch theft, and who uses violence or threatens violence on the spot to conceal stolen goods, resist arrest, or destroy evidence” will be legally transformed into the crime of robbery by a technical construction of the law. In the case of snatch theft, there is also a special provision known as "snatch theft by vehicle," which is considered a form of robbery. Additionally, robbery is a compound act that includes two elements: "the use of violence or threat of violence" and "the intention to acquire property," which leads to many issues regarding accomplices and attempts. For example, if a person attacks another violently and then decides to take their belongings, whether it is considered robbery or intentional injury coupled with theft. Another example is when a person violently robs and knocks out the victim, and a third party joins in to take the property; this raises questions about the determination of complicity. 3. Robbery is a type of property crime that can easily be confused with other property crimes, such as extortion and blackmail. For instance, in the case document we collected, "Chen Xiaoyong on robbery and extortion", the judge reasoned by distinguishing between "using violent and coercive means" and "using threatening means" to determine which of the defendant's multiple actions constituted robbery and which constituted extortion. In summary, robbery provides rich material for the Judge' reasoning in legal documents. Therefore, selecting robbery as the material for this study can provide a stronger diversity of reasoning-related data, which can better reflect the issues.

Besides, all documents are cases of first instance. China has a two-instance trial system and the second instance trial can correct the first instance trial. Most first instance trials are done by basic court in China. So, if we mix all cases, the outcome is not reliable. But there are problems only using cases of first instance, because some of them maybe incorrect. However, from the data of Law Yearbook of China we can confidently say it is not a matter. The rate of case reversal in China has been around 10% to 15% for a long time. In the latest yearbook data, the overall rate for criminal cases (calculated by subtracting the number of concluded first-instance cases from the number of concluded second-instance cases) is only 11.97%. These also include situations where the reasoning does not change significantly.[34] Therefore, even if there are some errors in the reasoning of the first instance, it will not severely affect the overall structure of the sample. So only using cases of first instance is reliable.

2. Structure of Chinese Judicial Documents

Every document from our sample could be divided into 4 parts: “Facts Accused by the Authority”, “Facts Found by the Court”, “The Court Believes” and “The Judgment is as Follow”. (see appendix E) The second part is the fact found by court based on evidence collected. This maybe sometimes differs from the facts from the authority. But we have to know that Judges write reasons and conclusions based on the facts from the court instead of that from the authority. So, I will pick up the second part as the case(fact) text. The third part is the reason writing part and the last part is the sentences part. Some special documents do not follow this structure, so I have to read them by myself and pick up each part.

What I need to add about is the sentence part. Sentences in China is combined of 2 parts, main sentences and additional sentences. Main sentences are usually a fixed-term imprisonment, and additional sentences is usually deprivation of political rights or fine or others. Some scholars give a formula to calculate the severity of punishment in China, but they are not reliable however because

[34] Law Yearbook of China, 1997—2007 ; Law Yearbook of China, 2021.

most of them do not open how they get this formula.[35] So I decide to use the main sentences only. I will prove the rationality by the PCA. The severity of sentences is in fact a multi-dimension data so we can use PCA to reduce dimensionality. In all formulas scholars give, and just as scholars all believe, that the additional sentences in China only takes a very small portion. So, after PCA, we can get a principal component that is highly similar to the main sentences. As a result, we can straightly use the main sentences because it makes sense. This method is also used in many researches in China.[36] Then I use how long months the sentences are to estimate this variable.

Then I can create a csv file with 4 columns. The first column is the number of cases. The second column contains the case variable. I pick all texts from the second portion of the judicial documents and write them into the second column. The third column is the reason writing part, and the method is the same as the case part. The last column is the data of main sentences, standing for the verdict conclusion. I will use this file to analyze the theory next. All NAs are removed by complete.cases function in R during my analysis.

Method

1. Documents Term Matrix

DTM is a classical method in NLP. It is practical in text analysis and text mining. What DTM does is to vectorize the text. Here is how it functions. First, the computer needs to cut the long texts into terms. This is called tokenization. Then, DTM will give every word a value by term frequency. Other methods may only give a dummy value. But this method makes sense due to Hart's theory. Hart says that the core of the principle of "same case, same judgment" lies in identifying which factors are important and which are not. For instance, differences such as the height of children compared to adults, or variations in appearance, religious beliefs, etc., should be considered as irrelevant differences. What we need to do is to extract the more important factors to determine the similarity of cases and the uniformity of judgments.[37]

As I mentioned, DTM use the principle of term frequency to give a value to every term. Elements that are highly discussed in common by Judges are, of course, more significant. For instance, when comparing adults and children of the same height, "age" is clearly a more important factor. This relies on the legal literacy of Judges, and we trust that their literacy is sufficient to support them in extracting the crucial elements. To prevent doubts, I will show part of the results of DTM later.

One thing I need to clear, is the package, function and other settings I use as well as the reasons for that. Chinese texts cannot be easily tokenized like English. There need to be a tokenizer to help. Luckily, Jiang Wu developed a package called "chinese.misc", aiming to help researchers to do Chinese texts analysis easier. This package use tm package as an assistant and I choose jiebaR as the tokenizer, which is a popular tokenizer not only in R, but also in Python (in Python it is called

[35] Jianjun Bai. Quantitative Analysis of the Severity of Punishment. Chinese Social Sciences, No. 6, pp. 114-125, 206 (2001) ; Jianjun Bai. Quantitative Analysis of the Severity of Crime. Chinese Social Sciences, No. 6, pp. 123-133, 208 (2003) ; Jianjun Bai. Empirical Research on Sentencing Standards. Chinese Journal of Law Studies, No. 1, pp. 97-105 (2008) ; Changming Hu. An Empirical Analysis of the Impact of Social Structural Factors on Sentencing: A Case Study of Sociological Research on Theft Crimes. Legal Application, No. 3, pp. 54-59 (2011).

[36] Yuhao Wu. The Boundaries of Sentencing Discretion: Collective Experience, Individual Decision-Making, and Bias Identification. Chinese Journal of Law Studies, Vol. 43, No. 6, pp. 109-129 (2021) ; Boyang Xu, You Zhou, and Chunli Zhang. An Empirical Test of Social Bond Theory and Self-Control Theory on Sexual Offending: Based on an Analysis of a Sample of 260 Sexual Offenders in China. Crime Research, Vol. 4, pp. 50-64 (2021); You Zhou, Boyang Xu, Ivan Y. Sun, Yan Zhang, and Lennon Y. C. Chang. Examining Sexual Crime Severity in China: A General-Specific Model on Sex Offending Against Adults. Sexual Abuse, Vol. 34, No. 7, pp. 830-856 (2022).

[37] H. L. A. Hart, The Concept of Law, Oxford: Clarendon Press, 1996, pp. 159-160+162.

"jieba").[38] Stopwords and text weighting are also important. Stopwords refer to words that frequently appear but do not carry substantial meaning and thus need to be removed, such as certain modal or conjunction words. Text weighting is used to address the issue where some words appear multiple times across many texts, rendering them insignificant in differential analysis. For example, in the differential analysis of multiple constitutional texts, words like "people" and "right" appear in every document. These words are unhelpful for differential analysis and can even obscure the differences between texts due to their high proportion. Therefore, tf-idf method is commonly used to give weights to these words. In this study, I use the built-in stopwords library from the jiebaR package to remove stopwords. Regarding the issue of weighting, this paper does not employ weighting methods. This is because many highly repetitive words are actually the result of Judges' strict adherence to the principle of treating like cases alike. Words such as "illegal possession purpose," "violence," and "coercion" are almost present in every document. If text weighting were used to clean the text, these words would be removed, which would distort the results. Therefore, this study does not incorporate text weighting.

2. Cosine Similarity

Cosine similarity is a popular method calculating the similarity of texts. By tokenizing the texts, we get a DTM which could present texts in a matrix. Each words have a value, so they can be presented in a multi-dimension space as a vector. Each vector stands for a document, and if the smaller the angle, the 2 texts would be more similar. We can use the cosine value to present the size of angle. Let me illustrates this by 3 short texts in my appendix E. I need to illustrate in 3 Chinese sentences including "I am a lawyer", "I am a criminal law lawyer" and "I am not a criminal law lawyer". Then tokenize sentences by jiebaR and give them a value. I use the dummy value to illustrate. The later vectors following each sentence are the DTM results.

Next, we can use the formula to calculate the cosine similarity between each sentence. I use No.1 as the benchmark. The result is 0.8944272 between No.1 and No.2, 0.7745967 between No.1 and No.3. It is obvious that the second text is more similar to the first text. The result is shown in table D1. By applying this method on long texts that I collect, we can get 2 sets of data, one is the similarity of cases, the other is similarity of reason writing. And if China achieves the goal of treating like cases alike, then there will be a significant correlation between them.

3. Sentences Similarity

I use Euclidean Distance to calculate the similarity between each 2 cases. This set of data cannot only be simply subtracted. We need to get a positive number because there will be a situation—case A has a 0.8 similarity to case C, and case B has a 0.8 similarity to case C as well. But case A is a more serious one with a terrible "*zhuguan'exing*", while case B is with a less "*zhuguan'exing*", and case C is a normal case. All differences between cases are from "*zhuguan'exing*". Then case A has a result of 14 months, case B 10 months and case C 12 months. Using case C as a benchmark, it is better to use Euclidean Distance to calculate this variable, or we will get an unexpected result that cases share the same similarity have different sentences similarity. Below is the formula of Euclidean Distance. Then we can create 3 sets of data in R by methods I illustrate. Using that 3 sets of data, we can build 3 linear models to test the hypnosis I put forward. In the next part of this paper, I will show the result of my analysis. And, OLS is used to build linear models.

[38] For example, Sheng Bi and etc. use jieba in Python to separate words of Chinese legal documents. See Sheng Bi, Zafar Ali, Meng Wang, Tianxing Wu, Guilin Qi. Learning heterogeneous graph embedding for Chinese legal document similarity. Knowledge-based systems, Vol. 250 (2022).

FINDINGS

Descriptive result

First, I will show part of the tokenization result in order to prove the reliability of that method. Besides, I will also use the traditional method to ensure the validation. Table A1 shows the result of tokenization. The table content is in Chinese, so I will explain the content later. The Variance column is the result of tokenization. The Percent column stands for the term frequency. We can see the first 5 terms are all statutory circumstances in criminal law. The later 3 terms are names of people or places. So, the jiebaR is well trained to be a good tokenizer. Then I use R to calculate the 3 sets of texts and numbers to get data of similarity. I will show them in the following part of my paper.

The table A2 shows the case similarity. The normality test indicates that the data is well-distributed according to QQ plot and the histogram of this set of data. All QQ plots and histograms could be found in my appendix A. We can find an interesting result that the data of reason similarity is higher than that of case. That means the text of reason writing has a high density of egal terminology. The result meets to our expectation for legal terminology.

One problem is that the normality test is not good so we have to deal with that problem. We need to normalize the data of sentences. Box-Cox is a popular method but this set of data contains a 0 value, so Box-Cox is not suitable. We have to seek other methods' help. Sqrt is not a good tool too because after sqrt the data is not normalized that well. The last method is to log the data. We use log1p function in R to deal with 0 value. This is also a practical method in empirical legal study.[39] Then we test the normality of this new data and find the result is a good one.

Linear Regression

Then I use the 3 sets of data to do a Pearson correlation. For all models I will delete the data of the base case. The figure A2.5 shows the result of correlation. The correlation heatmap indicates that there is a significant positive correlation between the case similarity and reason similarity. There is no significant correlation between reason similarity and sentences similarity. Moreover, there is a significant negative correlation between the case similarity and sentences similarity. The blank data is -0.06 but not significant. Then we can roughly have 3 models and I will show the result in the following content. After that I use the 3 sets of data, including case similarity, reason similarity and normalized sentences similarity, to build 3 linear regression models. For all models I delete the benchmark case.

Table A3 illustrates the model of case similarity and reason similarity. The tables show that the model coefficient is 0.279884, indicating a positive correlation between the two variables. This means that as the case similarity increases, the reason similarity also increases. The significance of the model coefficient and the overall model at the 0.01 level indicates that this positive correlation is statistically significant. Additionally, the model's explanatory power (R^2) is 7.49%, which aligns with the actual situation. This is because the data itself is influenced by factors such as names and places in the cases, which are almost never identical. Furthermore, the reasoning can be affected by extralegal factors (which will be discussed later), making the model's explanatory power quite realistic. Also, the table

[39] Xuesong Qian, Sheng Fang. Has the Reform of the Security Rights System Affected the Debt Financing of Private Enterprises? — Empirical Evidence from the Natural Experiment of China's 'Property Law', Economic Research Journal, Vol. 52, Iss. 5, pp. 146-160 (2017).

shows that reason similarity does not lead to similarity in sentences, a phenomenon that merits considerable thought. The model of reason similarity and sentences similarity shows that the coefficient is negative, indicating that the more similar the case are, the smaller the difference in sentences lengths. Thus, case similarity contributes to sentences similarity. Both the p-value of the coefficient and the overall p-value of the model are significant at the 0.01 level. But it is puzzling that while reason similarity resembles case similarity, it does not mirror the similarity in sentences lengths, yet the case similarity does. Moreover, the explanatory power (R^2) of this model is only 1.7%, which leads us to question the model's accuracy. Such results are likely due to the model being unreasonable or not robust, and it is highly possible that an outlier has caused this outcome. Therefore, further conclusions require regression diagnostics and confirmation of robustness.

Regression Diagnostics

For every linear model, we need to conduct regression diagnostics to test their stability. There may be some other model like a polynomial regression model, making our result is not reliable. Other problems will affect the model's stability as well, so regression diagnostics is a must. I use plot to conduct regression diagnostics, and the result is showed below.[40] All figures here are showed in my appendix A.

The diagnostic results of 3 models indicate that the variables in this model are essentially uniformly distributed, with no apparent trend of increase or decrease, suggesting that there is no issue with multicollinearity. The normality of the sample is good; there are no serious intra-relationship issues among the samples; and no points have a Cook's distance greater than 0.5, indicating the absence of significant outliers. In summary, my conclusions obtained in this study is reliable. Based on those 3 models, we can confidently assert that "the higher the case similarity, the higher the reason similarity. All figures here are showed in figure A3.1-5.4. The only potential problem is about the third linear regression. Although the regression diagnostics indicate some skewness in the model's samples, a small degree of skewness is common in social science experiments, and it is still considered acceptable. Of course, one could choose to reject this conclusion on the grounds that strict adherence to the prerequisites of linear regression dictates that the sample must exhibit normality. Regardless of whether the conclusion is accepted or not, it makes no difference in this study, as the robustness test for the model did not pass, leading to the rejection of the null hypothesis for this model regardless.

ROBUSTNESS

Change the Benchmark

The benchmark case, if we change that, the result may alter, and that is a good way to test robustness of results. Normally, we expect that the benchmark is a standard case with no special circumstances. And it needs to be a correct verdict. We basically have 3 methods to solve this problem: 1. Selecting Guiding Cases from the Supreme Court of China as the benchmark. However, guiding cases from the Supreme Court are usually established to address a particular legal issue, and while authoritative and mostly correct, they do not meet the requirement of "no special circumstances." 2. I choose a base case by myself, but that is not reliable as well because we cannot believe the

[40] Ford, C. 2020. Understanding Robust Standard Errors. UVA Library StatLab, at https://library.virginia.edu/data/articles/understanding-robust-standard-errors/ (last accessed February 1, 2024).

researcher's selection bias. 3. Using different cases as a benchmark to run the 3 models again. If we can get the same conclusion in all cases, then the models are robust. But to run hundreds of times is a burden for computer, and a waste of resources. So, I use sample function in R to pick 5 numbers between 1 and 523, then use those cases as the benchmark to run the models. Table A4.1 is the result.

To achieve a conclusion with 95% confidence, I employed an alpha level of 0.05, which makes the results more reliable. It can be observed that among the five samples, only the correlation between case similarity and reason similarity remains consistently significant (with both the model coefficients and overall p-values being less than 0.05). The other two models failed to yield more robust results, leading to the belief that their significance might be due to random sampling outcomes, which are not reliable. Moreover, in four out of the five samplings, the correlation between case details and reasoning reached a significance level of 0.01, implying that the model can be accurate with a confidence level of 99% in most cases. The explanatory power of the model also frequently exceeded 10%, and the instances where the explanatory power was lower could be entirely attributed to the inappropriate selection of benchmark in the sample, e.g., the case is not a that good one.

Other Text Similarity Algorithms

Except for Cosine Similarity, Jaccard Distance, Euclidean Distance and Correlation Similarity are all other prevalent method for text similarity. These results are shown in table A4.2.

1. Jaccard Distance

As I mentioned before, DTM has values standing for the term frequency. But maybe some will argue that we should treat each term equally. And Jaccard Distance is a tool for that. Jaccard Distance is to calculate the set of all terms within a text, then take the size of the intersection of these sets and divide it by the size of their union. It is also referred to as the Jaccard index, and the Jaccard distance is sometimes defined as 1 minus the Jaccard index. There is also an algorithm that uses the method of subtracting the intersection from the union. Regardless of the terminology or the specific algorithm used, the essence of the method remains the same. This article employs the method of dividing the intersection by the union minus the intersection. While this method of calculation is somewhat crude when applied to legal judgment documents, as it assigns the same value to all words, the text cosine similarity measure can also present a problem in that there may be biases in the assignment of values. For instance, in this set of texts, the phrase "People's Republic of China" is assigned a relatively high value because it is a term that every judge needs to use it when citing the law. Therefore, using Jaccard distance can help to eliminate potential biases that may arise from the assignment of values. The model coefficient for case similarity to reason similarity remains positive and significant, while the coefficient for reason similarity to sentences similarity is not significant. The conclusions drawn are consistent with those obtained using cosine similarity, indicating that the more similar the case similarity is, the more similar the reason is, but there is no relationship between the sentences' length and either of them. It is evident that the changing in method does not change the result. But we need a further test to ensure that.

2. Correlation Similarity

In the studies of legal texts, directly calculating the correlation coefficient is also a method to explore similarity. This is because, in legal texts, researchers often perform operations such as converting various systems into a list. Each legal text then has its own list, which are integrated to

form a matrix. The correlation coefficient is then calculated to represent the degree of similarity.[41] A DTM, with a similar structure, can also be subjected to the same treatment. The correlation coefficient can be directly computed using the cor function in R. In this study, the range of the correlation coefficient is from 0 to 1. Indeed, the true range of the correlation coefficient is actually from -1 to 1, but it is rare for negative values to appear between texts, unless one text contains terms that the other text lacks, and vice versa. Although applying this method, we find the third model again is significant, but it is not acceptable because in other methods we cannot get a same conclusion. But it is lucky that we find the first model is robust.

Result of Traditional Method

Some scholars may doubt this method because we cannot know the exact variables. Especially the reason-sentences model. The traditional method can take a lot of quantitative variables into consideration. If we can find a same conclusion by only using quantitative variables, then the result of my method is trustable. So, I select 28 variables using R based on China's criminal law. I select those variables from reason writing part. My variables are the same as Wu's research.[42] I will show the list of variables I select in my appendix B. Then I will use Cosine Similarity again to calculate the similarity, and conduct the linear regression again to see the result. Table A4.3 is the result. We get same conclusion. Using traditional method can get an unrobust result of significant correlation. The result does not only prove the robustness of my second model, but also have further meaning. My method takes more variables into consideration and get the same conclusion as what the traditional method gets. So, the prior conclusion is estimated in a further step then.

Logistic Regression

Some may argue that these 3 factors are not in a linear relationship. Instead, if we can divide them into groups and find a trend in groups cater to our expectation, then we can prove the theory I put forward.[43] In this situation, logistic regression is a good choice. A good way to divide cases is to use the reason writing part as the base. So, we can divide those data by low-similarity reason writing groups and high-similarity reason writing groups. But a problem is how can we divide groups? The machine learning could help in this situation.

1. K-means

K-means is a practical clustering algorithm used in machine learning and data mining. The K-means algorithm aims to partition a dataset into 'k' clusters in which each observation belongs to the cluster with the nearest mean. This results in a partitioning of the data space into 'k' Voronoi cells. K-means is a kind of unsupervised learning, so after clustering I have to label each group according to the centers. Before clustering, the reason writing data needs to be standardized. I use scale function in R to do that. Then I need to decide which k I will use to cluster. Elbow method is a good one to help. Using this method, k has better to be 6 according to elbow method. The result of elbow method is shown in figure A6.1. Then using k = 6 I get a 6-group sample. The result can be seen in figure

[41] David S. Law, Mila Versteeg. The Declining Influence of the United States Constitution, New York University Law Review, Vol. 87, No. 3 (2012); Anu Bradford, Yun-chien Chang, Adam Chilton, Nuno Garoupa. Do Legal Origins Predict Legal Substance? Journal of Law and Economics, Vol. 64, No. 2, pp. 207-231 (2021). Yun-chien Chang, Property Law: Comparative, Empirical, Economic Analyses, Cambridge: Cambridge University Press, 2023, pp. 39-45.

[42] Yuhao Wu. The Boundaries of Sentencing Discretion: Collective Experience, Individual Decision-Making, and Bias Identification. Chinese Journal of Law Studies, Vol. 43, No. 6, pp. 109-129 (2021).

[43] Jianjun Bai. The Constitutional Significance and Empirical Research of Consistency in Judgment for Similar Cases. China Legal Science, Vol. 32, No. 3, pp. 131-140 (2003).

A6.2. Every cluster stands for a level of similarity and in my sample the 6th is the lowest. Next step is to use this data to conduct logistic regression.

2. Logistic Regression

First is to conduct a regression between case and clusters. I use the 6th cluster as the benchmark. It has the lowest similarity among all clusters, so all other clusters, if as expected, should have same direction in logistic regression, meaning all ORs are positive or negative in one model. Table A4.4.2 shows the result of logistic regression. It shows that: the higher the reason similarity, the higher the case similarity. The second model shows the relationship between clusters and sentence. As you can see there is a negative OR as expected but it is not significant. I will report the clustering result in table A4.4.1 as well. Again, we get the same conclusion. Even the relationship is not linear, the result of logistic regression can get the same conclusion.

Data Dimension Reduction

A high data dimension will cause problems. In this research too many dimensions with a value 0 will case a potential problem that is those cases with many 0 values are not that similar as the data shows. I need to explain that the one who choose "*renzuirenfa*" will be more likely to "*provide a truthful fact*", so variables are in a linear correlation relationship. In this situation, PCA will be helpful to reduce this kind of problem. We get same result by employing PCA method to get the main characters of data. For the reason data, the PC1 to PC100 can explain over 80% of the data. While for the case data, PC1 to PC236 all can explain over 80% of the data. The results of PCA will be shown in figure A7.1 and figure A7.2. The new sets of data are used to build new models and get same results. Table A4.5 shows the result.

DISCUSSION

A strange thing is, why there is a significant correlation between case and reason, but 2 insignificant correlations between sentences and other 2 variables? How to explain this phenomenon? I want to use my data frame to discuss it.

Let me first explain why the similar reason and case similarity do not bring the sentences similarity. I would argue that there may be one reason for that, which is the factors outside the documents, play a more important role than we have expected. And that leads to the bias in predicting the verdicts. Factors I mentioned before, and have been proved by researchers, e.g., genders and races and even the mood of Judges would affect the conclusion. But due to the professional training as a Judge or a lawyer, they would use reason writing as a tool to hide their true reasons for verdict. As a result, reasoned elaboration may be a guarantee of judicial transparency and trustworthiness.[44] John also points out that the order of how Judges write reasons, meaning some may write reasons before a conclusion while some prefer to get a conclusion at first, would affect the sentence decision and, this could lead to the result of my paper.[45] Because some Judges will get a conclusion before they write reasons for verdict. Besides, a factor may function not only in the court but also outside the court, and Judges themselves may not write all reasons in the documents, e.g., Björn Dressel and Tomoo Inoue found that social network, both external to the courts and within can be a factor on verdict, but no one

[44] John Zhuang Liu, Xueyao Li. Legal Techniques for Rationalizing Biased Judicial Decisions: Evidence from Experiments with Real Judge. Journal of Empirical Legal Studies, Vol. 16, Iss. 3, pp. 630–670 (2019).

[45] John Zhuang Liu. Does Reason Writing Reduce Decision Bias? Experimental Evidence from Judges in China. The Journal of Legal Studies, Vol. 47, No. 1, pp. 83-118 (2018).

would write them down and of course they do not have to write all down.[46] This problem especially in China the moral attitude of Judges is discussed by many scholars. For example, some Judges may consider the verdict from a moral perspective. In order to deal with that, Chen and Cheng suggest the court should have a procedure to open these factors to public.[47] This could explain it as well. And I want to add that some provinces in China have their guidance on sentencing decision for Judges and that may lead to my conclusion. So further researches can focus on different provinces in China. All these factors also contribute to a low R^2 of models.

But why there is a correlation between case similarity and reason similarity? Judges can identify the legal requirements and apply the law correctly due to the professional training of China's National Judicial Examination, leading to the correlation of case similarity and reason similarity. As a consequence, Judges can provide similar reasoning for similar cases, but the specific reasons for the variation in sentencing are personal and immeasurable. For instance, some Judges tend to favor harsher sentences, while others prefer more lenient ones. What is more, the China's National Judicial Examination, and it is an examination focus more on how to write reasons for a case, and usually ignore how to give a sentence correctly. So, Judges in China do not receive a training on the sentencing part. I believe that may be the same in many other countries. Coupled with the reasons discussed earlier, this leads to the results obtained in this paper. So further researches can be conducted on to what extend exactly the extrajudicial factors have an impact on verdict, instead of just investigating whether which factor has an influence on verdict.[48]

One thing I want to talk about more, is the unstable correlation between reason writing and verdict conclusion. We can see that there is always an unrobust significant linear model whether I use text analysis or the traditional method. This can be explained that if we choose a high-quality benchmark, e.g., the No. 1 case in my sample, which is correctly sentenced, then we can get a good model. If the base case itself is a biased one, then the result of model would not be good. So, the discretion of the Judges, if being used incorrectly or without any limitation, then it will destroy the goal of treating like cases alike. From all above discussion, a recommendation for China's judgement is to focus more on the discretion of Judges, which is also what China's Supreme Court is trying to do though a series of policies.

CONCLUSION

This paper puts forward a method to estimate biased judgements. It answers question "Does China achieve the goal of TLCA?" My answer is no. Although we can find from China's judicial documents that the higher the case similarity is, the higher the reason similarity. But that is just a result of the professional training, ensuring that Judges possess the capability to identify the legal requirements. While factors outside documents, sometimes even being hidden by Judges on purpose, contribute to the result of model 3, showing there are no significant correlation between sentences and cases, reason writing. But at least I can confidently say China partly achieve the goal of treating like cases alike and then next policy should be to control the discretion and maybe sentencing skills need to be emphasized in the China's National Judicial Examination.

---

[46] Björn Dressel, Tomoo Inoue. Informal networks and judicial decisions: Insights from the Supreme Court of the Philippines, 1986–2015. International Political Science Review, Vol. 39, No. 05, pp. 616-633 (2018).

[47] Liang Chen, Jinhua Cheng. How to Import Moral Judgment into Judicial Adjudication. Exploration and Free Views, No. 08, pp. 59-72, 178 (2023).

[48] This question has been researched for years and has a lot of findings. Some early researches could be seen. See Reskin, Barbara, Visher, Christy. The Impact of Evidence and Extra-Legal Factors in Jurors' Decisions. Law & Society Review, Vol. 20, No. 3, pp. 423-438 (1986) ; Marilyn Chandler Ford. The Role of Extralegal Factors in Jury Verdicts, Vol. 11, No. 01, pp. 16-39 (1986).

REFERENCES

[1] Aiwu Bao. A Watched Flower Never Blooms — A Study of Guiding Cases System and “Treating Like Cases Alike”. Doctoral Thesis, City University of Hong Kong, 2020.

[2] Aleksander Peczenik. On Law and Reason. Springer, 2008.

[3] Allison P. Harris, Maya Sen. Bias and Judging. Bias and Judging. Annual Review of Political Science, Vol. 22, pp. 241-259 (2019).

[4] Andrei Marmor. Should Like Cases Be Treated Alike, Legal Theory, Vol. 11, Iss. 1, pp. 27-38 (2005)

[5] Anu Bradford, Yun-chien Chang, Adam Chilton, Nuno Garoupa. Do Legal Origins Predict Legal Substance? Journal of Law and Economics, Vol. 64, No. 2, pp. 207-231 (2021).

[6] Avani Mehta Sood. What’s So Special About General Verdicts? Questioning the Preferred Verdict Format in American Criminal Jury Trials, Theoretical Inquiries in Law, Vol. 22, No. 2, pp. 55-84 (2021).

[7] Benjamin M. Chen, Zhiyu Li, David Cai, Elliott Ash. Detecting the influence of the Chinese guiding cases: a text reuse approach. Artificial Intelligence and Law (2023).

[8] Björn Dressel, Tomoo Inoue. Informal networks and judicial decisions: Insights from the Supreme Court of the Philippines, 1986–2015. International Political Science Review, Vol. 39, No. 05, pp. 616-633 (2018).

[9] Boyang Xu, You Zhou, and Chunli Zhang. An Empirical Test of Social Bond Theory and Self-Control Theory on Sexual Offending: Based on an Analysis of a Sample of 260 Sexual Offenders in China. Crime Research, Vol. 4, pp. 50-64 (2021).

[10] Changming Hu. An Empirical Analysis of the Impact of Social Structural Factors on Sentencing: A Case Study of Sociological Research on Theft Crimes. Legal Application, No. 3, pp. 54-59 (2011).

[11] Christina L. Boyd, Lee Epstein, Andrew D. Martin. Untangling the Causal Effects of Sex on Judging. American Journal of Political Science, Vol. 54, Iss. 2, pp. 389-411 (2010).

[12] Christopher Sherrin. Inconsistent Verdicts and the Possibility of Innocence: A Comment on R v RV, Wrongful Conviction Law Review, Vol. 2, No. 1, pp. 78-81 (2021).

[13] Daniel M. Schneider. Empirical Research on Judicial Reasoning: StatutoryEmpirical Research on Judicial Reasoning: Statutory Interpretation in Federal Tax Cases, New Mexico Law Review, Vol. 31, Iss. 2, pp. 325-258 (2001).

[14] David S. Law, Mila Versteeg. The Declining Influence of the United States Constitution, New York University Law Review, Vol. 87, No. 3 (2012).

[15] Erik Voeten. Gender and judging: evidence from the European Court of human rights. Journal of European Public Policy, Vol. 28, Iss. 9, pp. 1453-1473 (2020).

[16] Ford, C. 2020. Understanding Robust Standard Errors. *UVA Library StatLab*. https://library.virginia.edu/data/articles/understanding-robust-standard-errors/ (accessed February 1, 2024).

[17] H. L. A. Hart, Positivism and the Separation of Law and Morals. Harvard Law Review, Vol. 71, No. 4, pp. 593-629 (1958).

[18] H. L. A. Hart, The Concept of Law, Oxford: Clarendon Press, 1996.

[19] Haibo Sun. In What Sense Does Case Retrieval Assist in Achieving Consistent Judgments for Similar Cases? Tsinghua Law Review. Vol. 15, Iss. 1, pp. 79-97 (2021).

[20] Haibo Sun. Should Similar Cases Be Treated Similarly? . Legal System and Society Development, Vol. 25, No. 3 (2019).

[21] Haibo Sun. The 'Same Case, Same Judgment' Principle: Not a Fictional Myth of the Rule of Law. Jurists Review, Vol. 5, pp. 141-157 (2019).

[22] Haoxi Zhong, Guo Zhipeng, Cunchao Tu, Chaojun Xiao, Zhiyuan Liu, Maosong Sun. Legal Judgment Prediction via Topological Learning.
[23] Hugo Storey. Consistency in Refugee Decision-Making: A Judical Perspective, Refugee Survey Quarterly, Vol. 32, Iss. 4, pp. 112-125 (2013).
[24] Jennifer Mayer Cox, Naomi E. S. Goldstein, John Dolores, Amanda D. Zelechoski & Sharon Messenheimer. The Impact of Juveniles' Ages and Levels of Psychosocial Maturity on Judges' Opinions About Adjudicative Competence. Law and Human Behavior, pp. 1-9 (2010) .
[25] Jianjun Bai. Empirical Research on Sentencing Standards. Chinese Journal of Law Studies, No. 1, pp. 97-105 (2008) .
[26] Jianjun Bai. Quantitative Analysis of the Severity of Crime. Chinese Social Sciences, No. 6, pp. 123-133, 208 (2003).
[27] Jianjun Bai. Quantitative Analysis of the Severity of Punishment. Chinese Social Sciences, No. 6, pp. 114-125, 206 (2001).
[28] Jianjun Bai. Research on Sentencing Prediction Based on the Collective Experience of Judges. Chinese Journal of Law Studies, Vol. 38, No. 6, pp. 140-154 (2016).
[29] Jianjun Bai. The Constitutional Significance and Empirical Research of Consistency in Judgment for Similar Cases. China Legal Science, Vol. 32, No. 3, pp. 131-140 (2003).
[30] Jinghui Chen. Same Case, Same Judgment: A Legal Duty or a Moral Requirement? China Legal Science, Iss. 3, pp. 46-61 (2013).
[31] John Zhuang Liu. Does Reason Writing Reduce Decision Bias? Experimental Evidence from Judges in China. The Journal of Legal Studies, Vol. 47, No. 1, pp. 83-118 (2018).
[32] Karl N. Llewellyn. The Bramble Bush: on our law and its study. Quid Pro Books, 2012, p.10.
[33] Kenneth I. Winston. On Treating Like Cases Alike, The Independent Review, Vol. 4, No. 1, pp. 107-118 (1999)
[34] Kosuke Imai. Quantitative Social Science: An Introduction. Princeton University Press, 2017, pp. 236-238.
[35] Law Yearbook of China, 1997—2007 ; Law Yearbook of China, 2021.
[36] Lei Lei. How to Understand 'Same Case, Same Judgment'?—Misconceptions and Clarifications. Political Science and Law Review, Vol. 5, pp. 28-38 (2020).
[37] Liang Chen, Jinhua Cheng. How to Import Moral Judgment into Judicial Adjudication. Exploration and Free Views, No. 08, pp. 59-72, 178 (2023).
[38] Marilyn Chandler Ford. The Role of Extralegal Factors in Jury Verdicts, Vol. 11, No. 01, pp. 16-39 (1986).
[39] Olga Alejandra Alcántara Francia, Miguel Nunez-del-Prado, Hugo Alatrista-Salas. Survey of Text Mining Techniques Applied to Judicial Decisions Prediction. Applied Sciences, Vol. 12, No. 20 (2022).
[40] Qi Zhang. On the Similar Treatment of Similar Cases. Global Legal Review, Vol. 36, No. 1, pp. 21-34 (2014) .
[41] Reskin, Barbara, Visher, Christy. The Impact of Evidence and Extra-Legal Factors in Jurors' Decisions. Law & Society Review, Vol. 20, No. 3, pp. 423-438 (1986)
[42] Shai Danziger, Jonathan Levav, and Liora Avnaim-Pesso. Extraneous factors in judicial decisions. Proceedings of the National Academy of Sciences of the United States of America. Vol. 108, No. 17, pp. 6889-6892 (2011)
[43] Shaohua Zhou. A Rational Examination of 'Same Case, Same Judgment' in Criminal Cases. Legal Science, Vol. 37, No. 3, pp. 3-15 (2020) .
[44] Shaohua Zhou. Differentiated Judgments in Criminal Cases and Their Rationality. China Legal Science, Vol. 4, pp. 145-164 (2019).
[45] Shaohua Zhou. Same Case, Same Judgment: A Fictional Myth of the Rule of Law. Law Science Magazine, Vol. 32, No. 11, pp. 131-140 (2015).

[46] Sheng Bi, Zafar Ali, Meng Wang, Tianxing Wu, Guilin Qi. Learning heterogeneous graph embedding for Chinese legal document similarity. Knowledge-based systems, Vol. 250 (2022).
[47] Shuai Wei, Moulin Xiong. Judges' Gender and Sentencing in China: An Empirical Inquiry. Feminist Criminology, Vol. 15, Iss. 2, pp. 217-250 (2020).
[48] Shude Liu. The principle of 'Same Case, Same Judgment' in the Context of Criminal Justice. China Legal Science, Vol. 1, pp. 68-76 (2011).
[49] Timothy J. Capurso. How Judges Judge: Theories on Judicial Decision Making. University of Baltimore Law Forum, Vol. 29, No. 01, Article 2 (1998).
[50] Welbers, Kasper and Van Atteveldt, Wouter and Benoit, Kenneth. Text analysis in R. Communication Methods and Measures, Vol. 11, No. 4, pp. 245-265 (2017).
[51] Xingliang Chen. China'S Guiding Case System : A Study on The Mechanisms Of Rule Formation. Peking University Law Journal, Vol 1, No 2, pp. 215-258 (2015).
[52] Xuesong Qian, Sheng Fang. Has the Reform of the Security Rights System Affected the Debt Financing of Private Enterprises? — Empirical Evidence from the Natural Experiment of China's 'Property Law', Economic Research Journal, Vol. 52, Iss. 5, pp. 146-160 (2017).
[53] You Zhou, Boyang Xu, Ivan Y. Sun, Yan Zhang, and Lennon Y. C. Chang. Examining Sexual Crime Severity in China: A General-Specific Model on Sex Offending Against Adults. Sexual Abuse, Vol. 34, No. 7, pp. 830-856 (2022).
[54] Yuhao Wu. The Boundaries of Sentencing Discretion: Collective Experience, Individual Decision-Making, and Bias Identification. Chinese Journal of Law Studies, Vol. 43, No. 6, pp. 109-129 (2021).
[55] Yun-chien Chang, Property Law: Comparative, Empirical, Economic Analyses, Cambridge: Cambridge University Press, 2023.
[56] Zachary Elkins, Tom Ginsburg, James Melton. Comments on Law and Versteeg's The Declining Influence of the United States Constitution, New York University Law Review, Vol. 87, No. 6, pp. 2088-2101 (2012).
[57] Zikun Hu, Xiang Li, Cunchao Tu, Zhiyuan Liu, Maosong Sun. Few-Shot Charge Prediction with Discriminative Legal Attributes. COLING 2018.

PACKAGES OF R

[1] Arel-Bundock V (2022). “modelsummary: Data and Model Summaries in R.” _Journal of Statistical Software_, *103*(1), 1-23. https://doi.org/10.18637/jss.v103.i01 .

[2] Feinerer I, Hornik K (2023). _tm: Text Mining Package_. R package version 0.7-11, https://CRAN.R-project.org/package=tm .

[3] Fox J, Weisberg S (2019). _An R Companion to Applied Regression_, Third edition. Sage, Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/ .

[4] Gohel D, Skintzos P (2024). _flextable: Functions for Tabular Reporting_. R package version 0.9.5, https://CRAN.R-project.org/package=flextable .

[5] H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

[6] Taiyun Wei and Viliam Simko (2021). R package 'corrplot': Visualization of a Correlation Matrix (Version 0.92). Available from https://github.com/taiyun/corrplot .

[7] Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

[8] Wenfeng Q, Yanyi W (2019). _jiebaR: Chinese Text Segmentation_. R package version 0.11, https://CRAN.R-project.org/package=jiebaR .

[9] Wu J (2020). _chinese.misc: Miscellaneous Tools for Chinese Text Mining and More_. R package version 0.2.3, https://CRAN.R-project.org/package=chinese.misc .

APPENDIX A

Figures

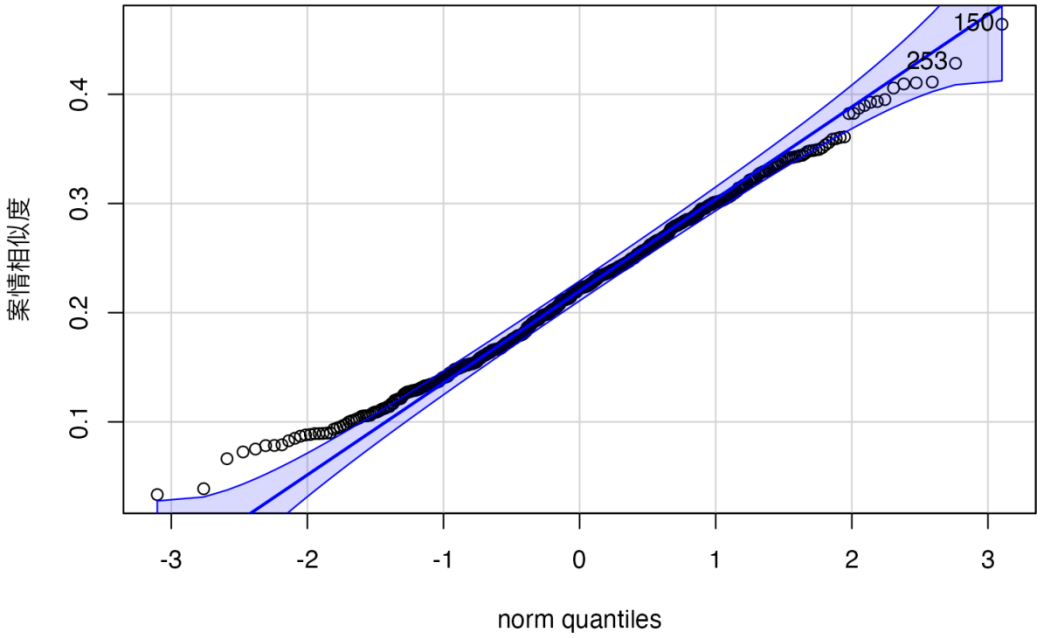


Figure A1.1: QQ plot of case similarity

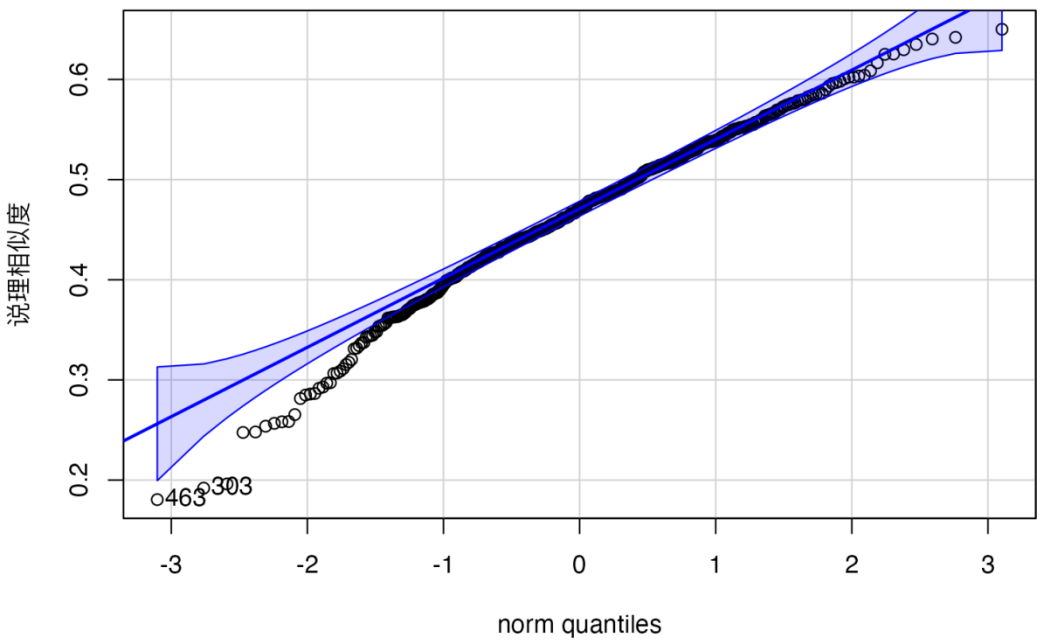


Figure A1.2：QQ plot of reason similarity

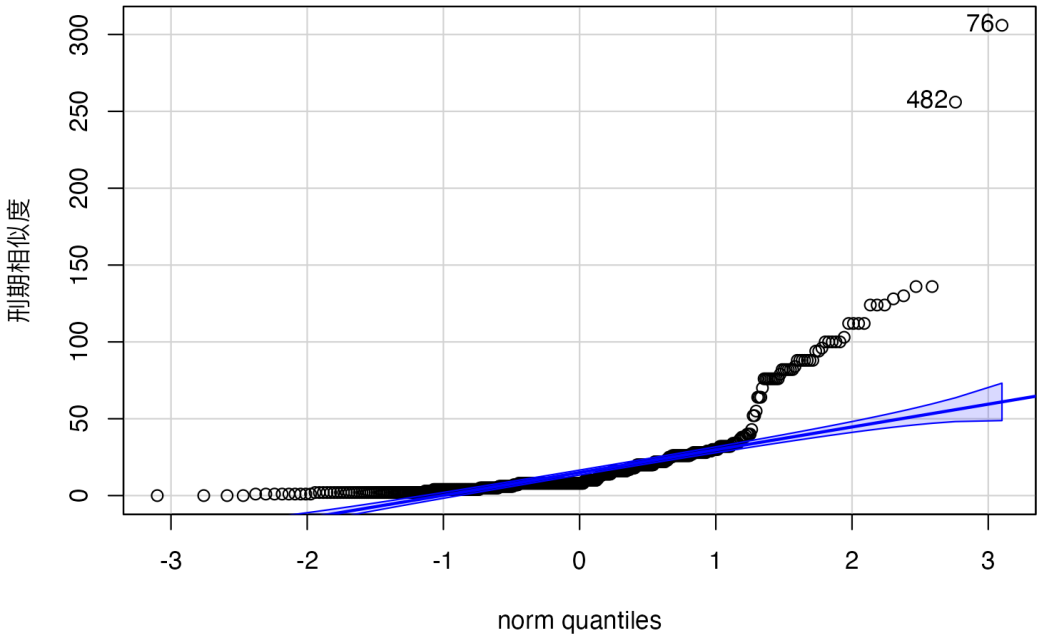


Figure A1.3: QQ plot of sentences similarity

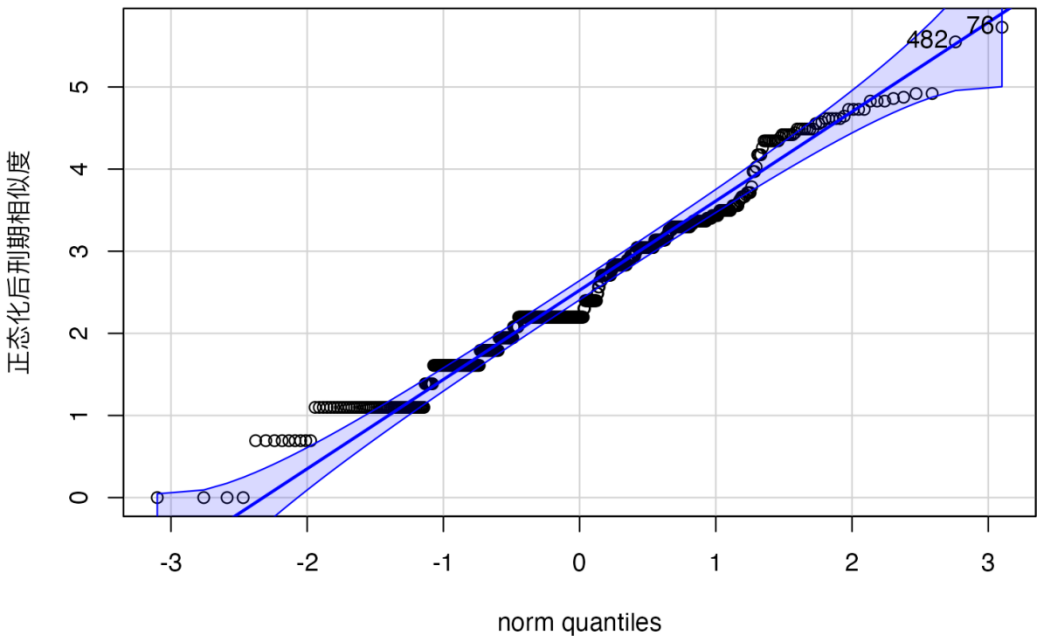


Figure A1.4: QQ plot of normalized sentences similarity

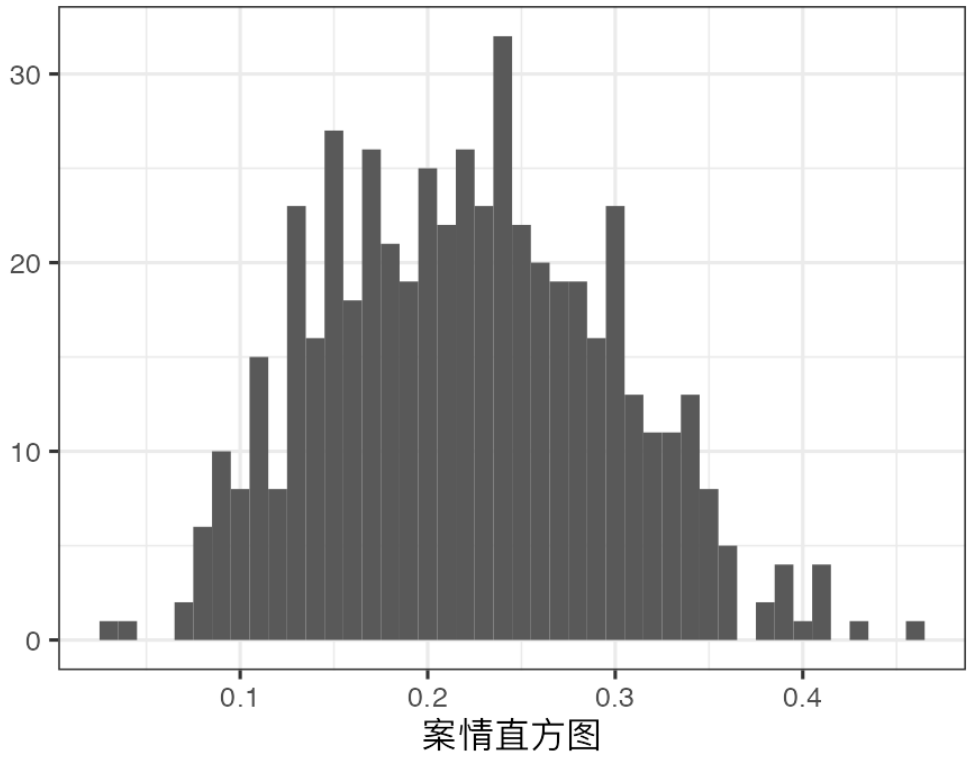


Figure A2.1: histogram of case similarity

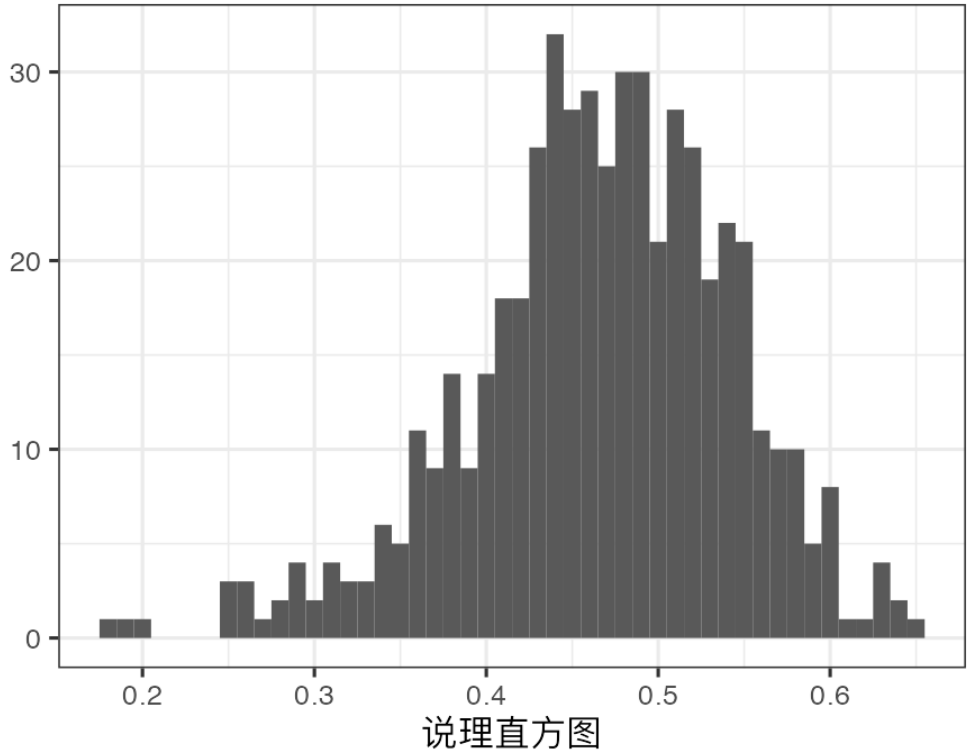


Figure A2.2: histogram of reason similarity

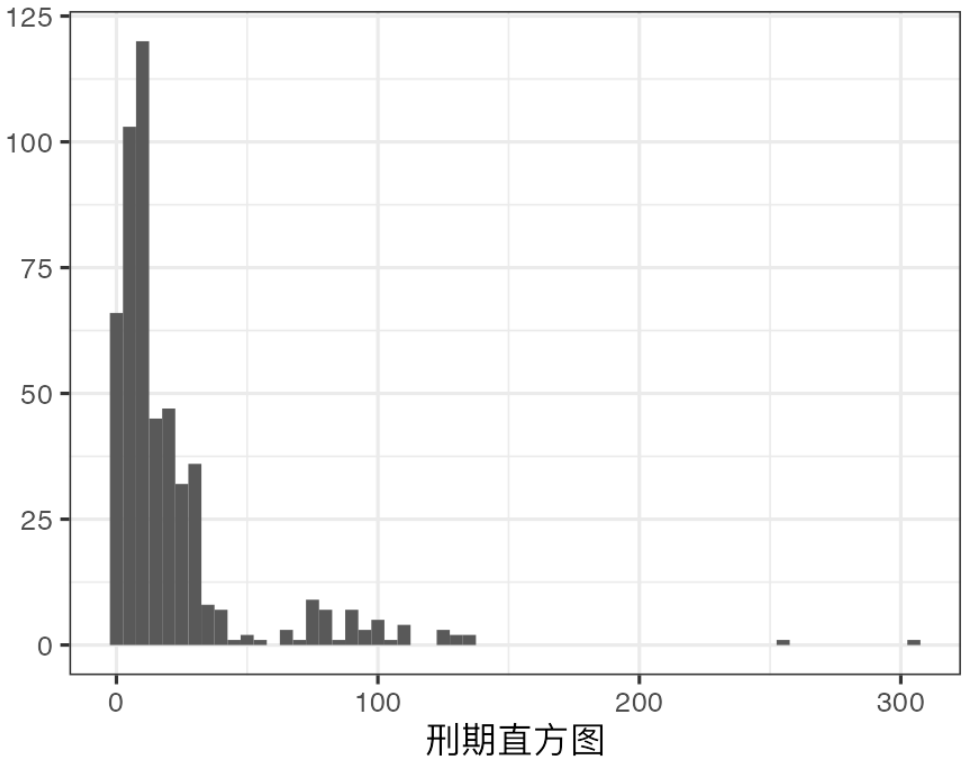


Figure A2.3: histogram of sentences similarity

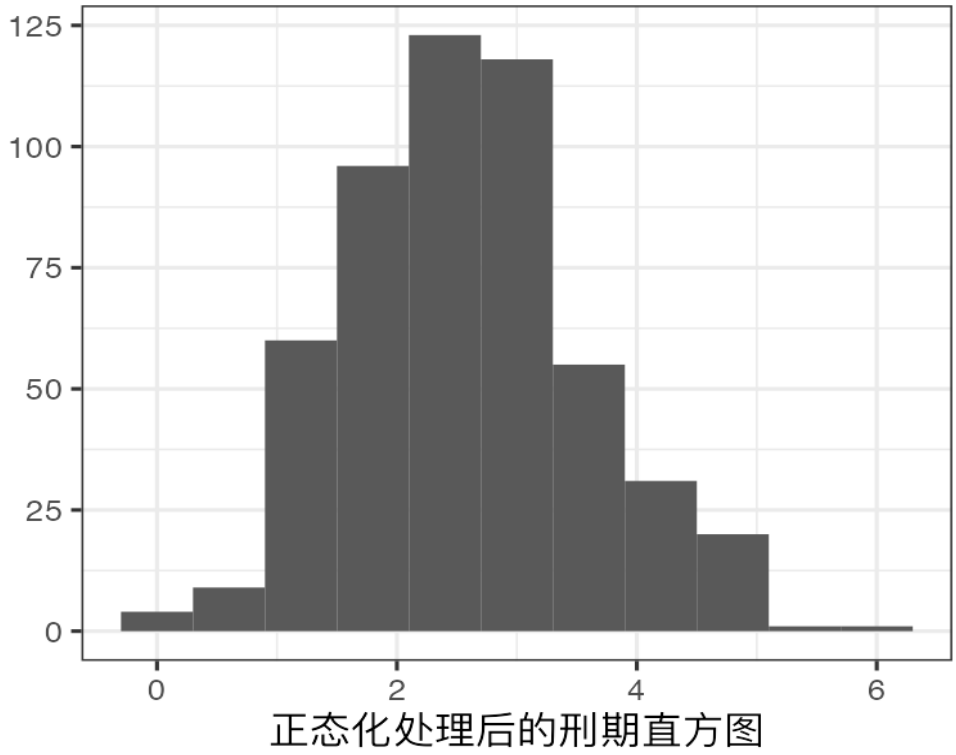


Figure A2.4: histogram of normalized sentences similarity

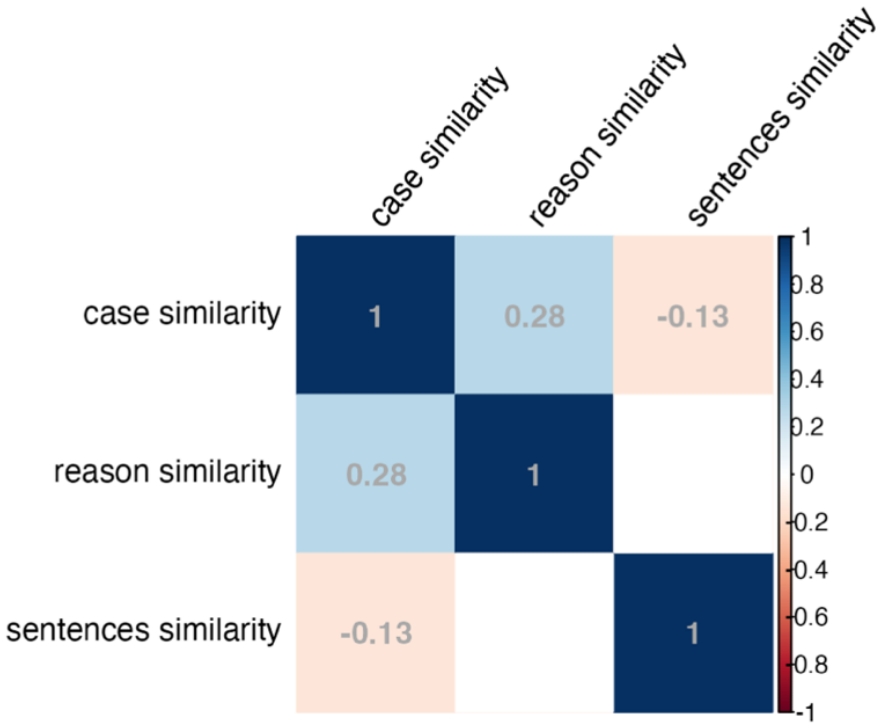


Figure A2.5: Pearson correlation heatmap

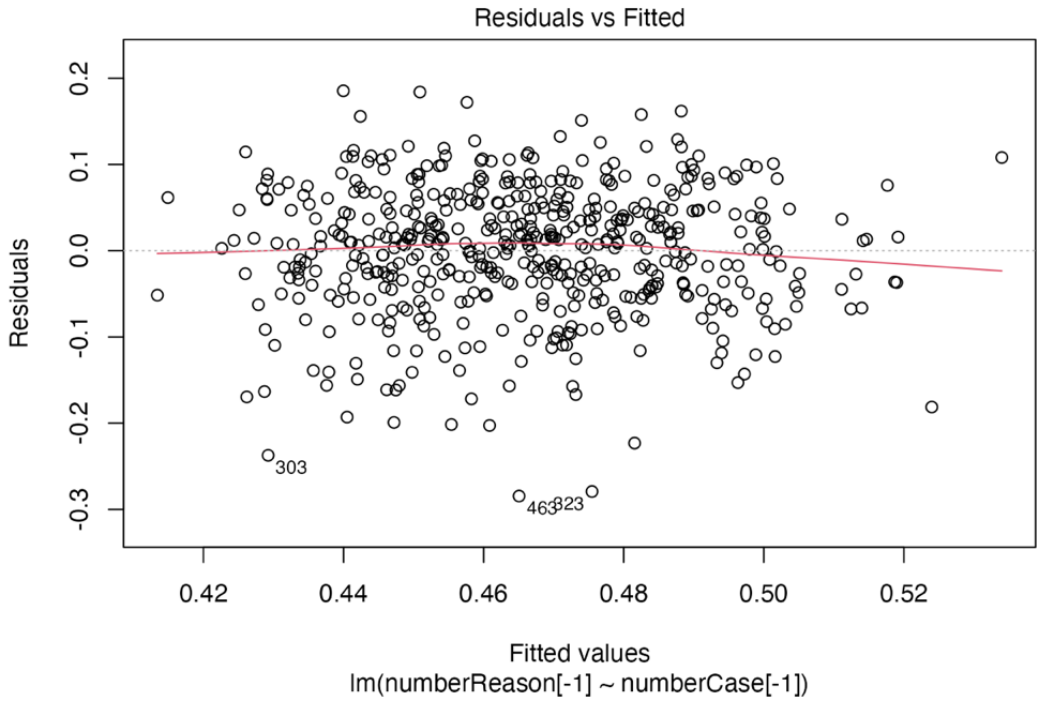


Figure A3.1: model 1 regression diagnostics 1

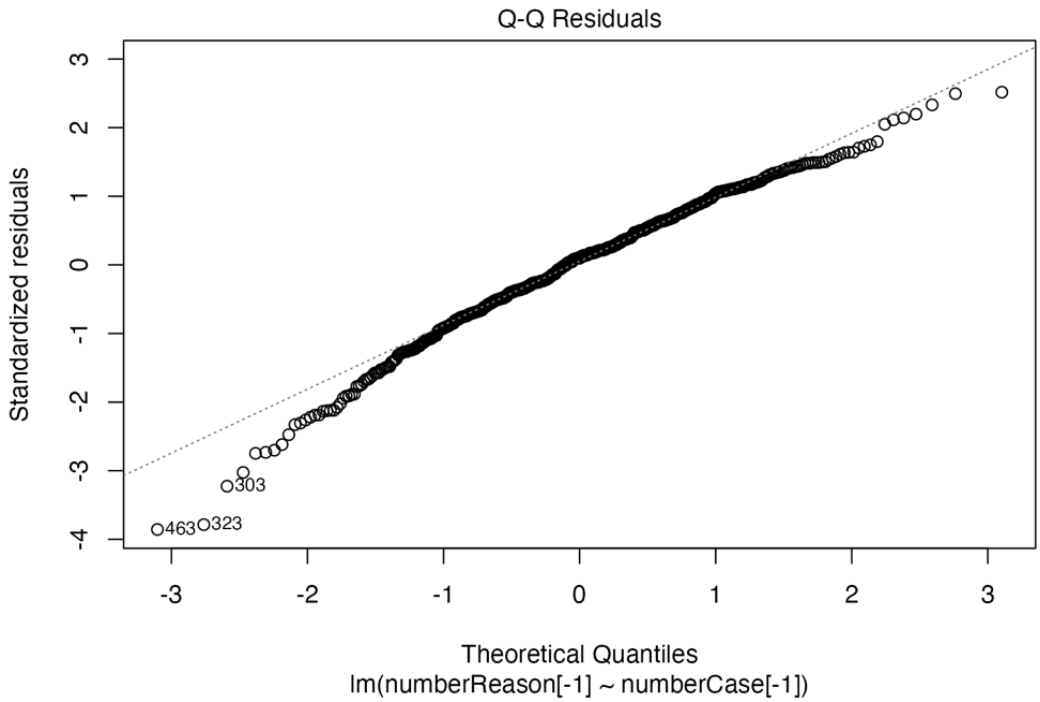


Figure A3.2: model 1 regression diagnostics 2

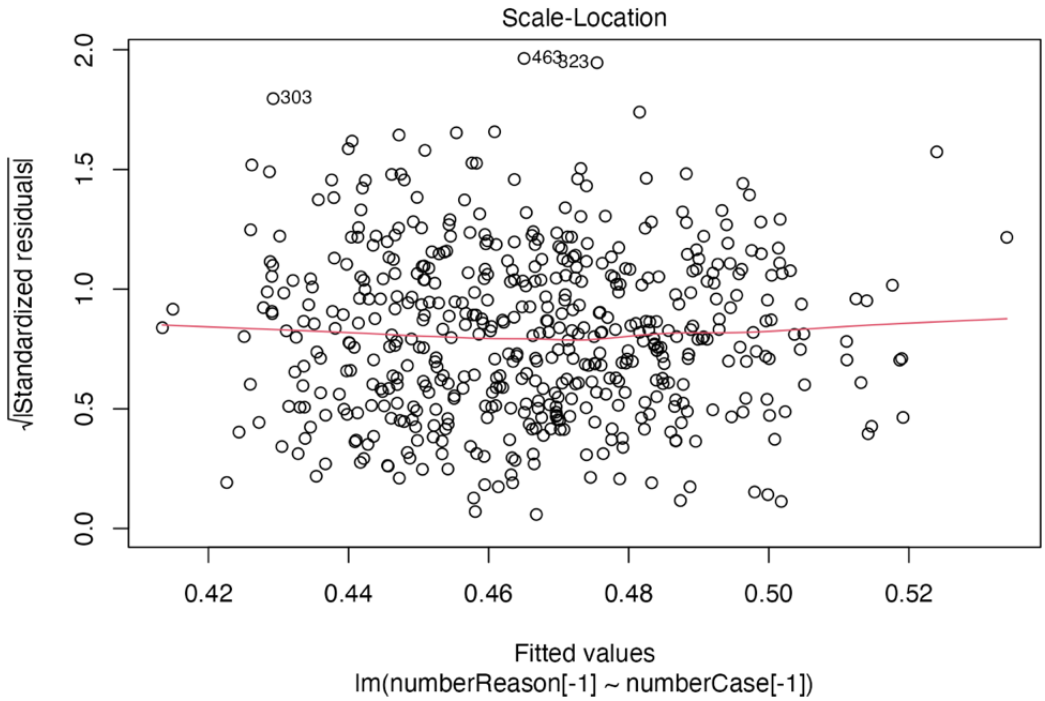


Figure A3.3: model 1 regression diagnostics 3

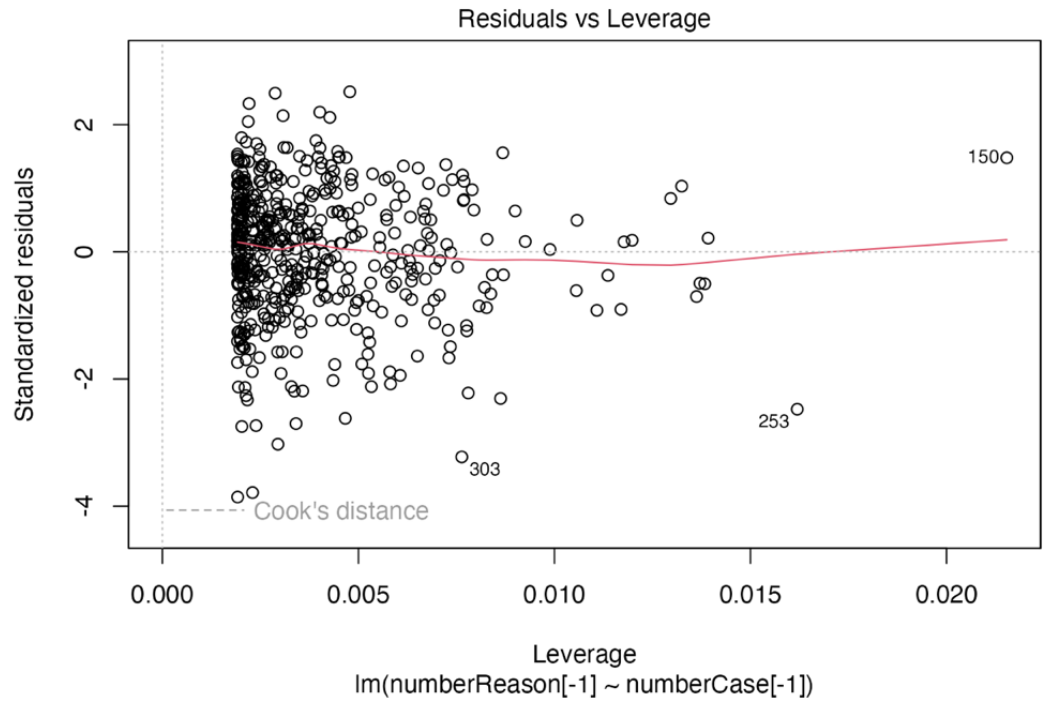


Figure A3.4: model 1 regression diagnostics 4

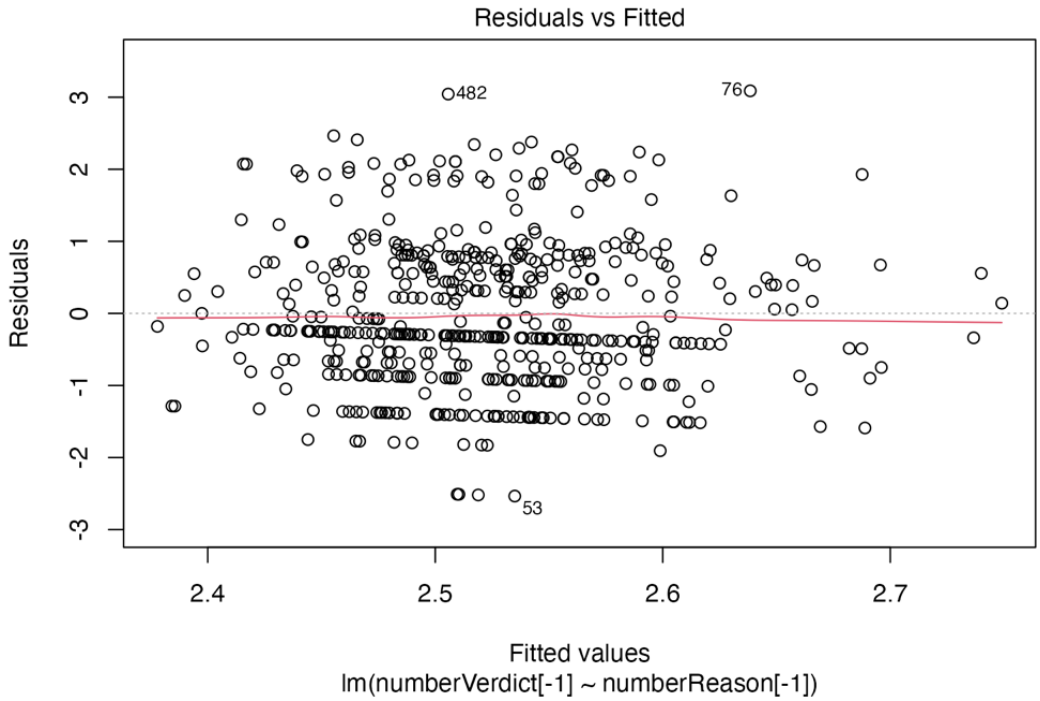


Figure A4.1: model 2 regression diagnostics 1

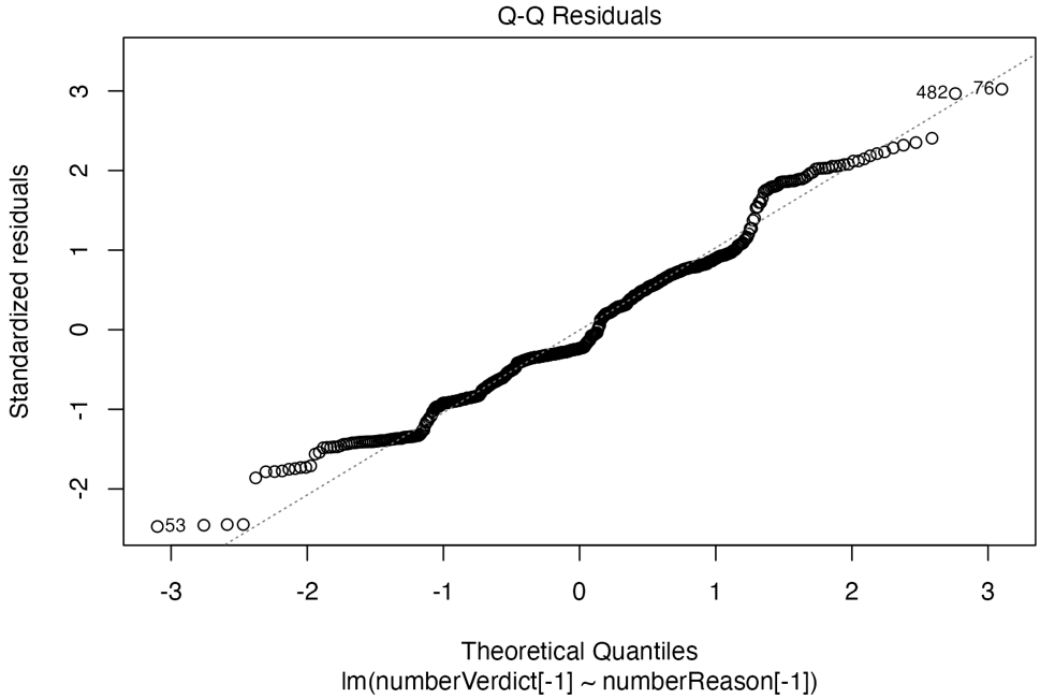


Figure A4.2: model 2 regression diagnostics 2

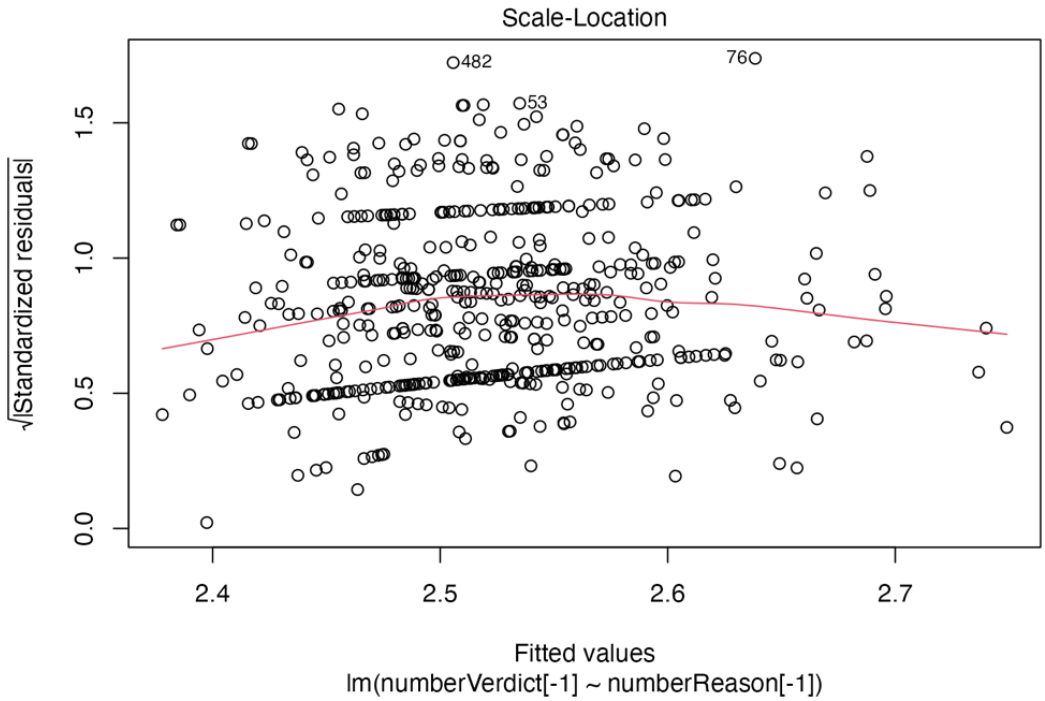


Figure A4.3: model 2 regression diagnostics 3

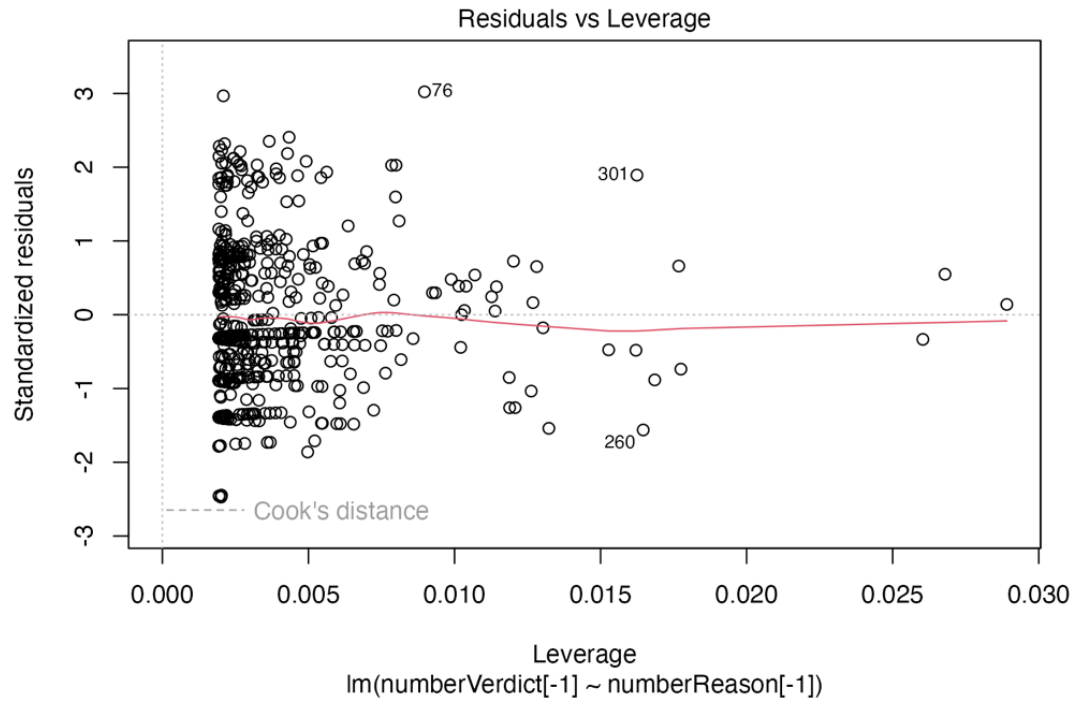


Figure A4.4: model 2 regression diagnostics 4

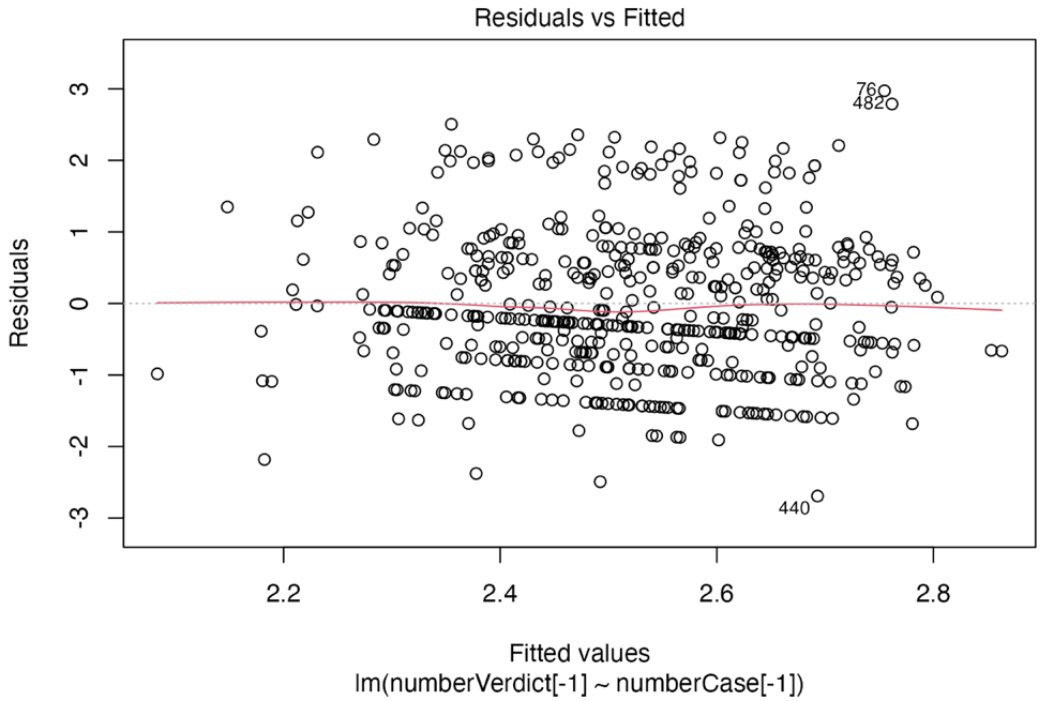


Figure A5.1: model 3 regression diagnostics 1

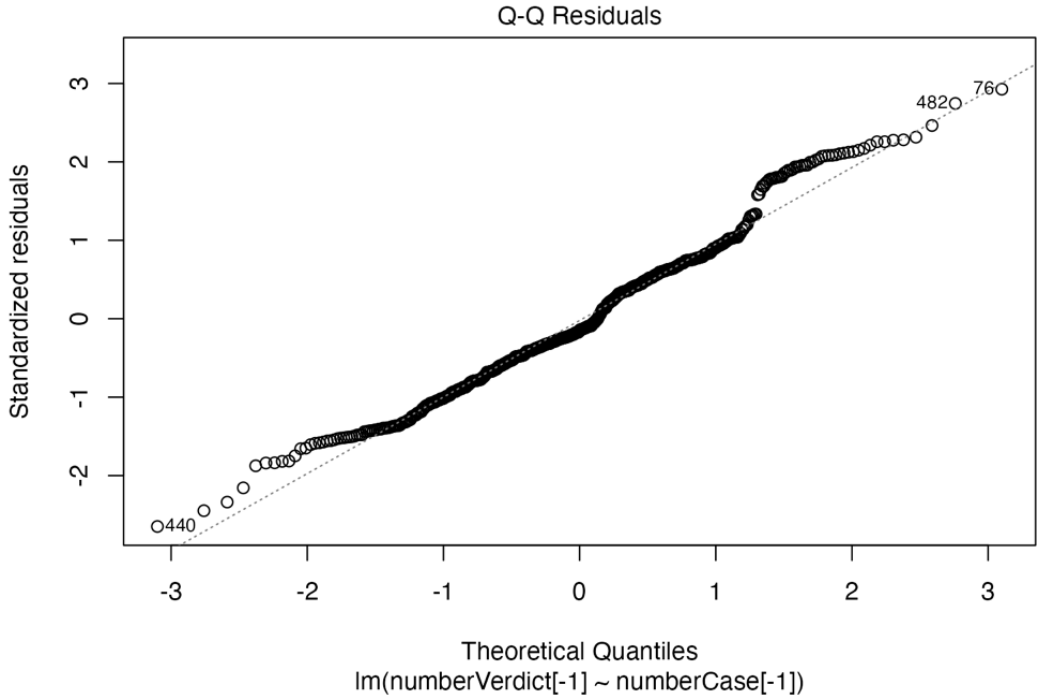


Figure A5.2: model 3 regression diagnostics 2

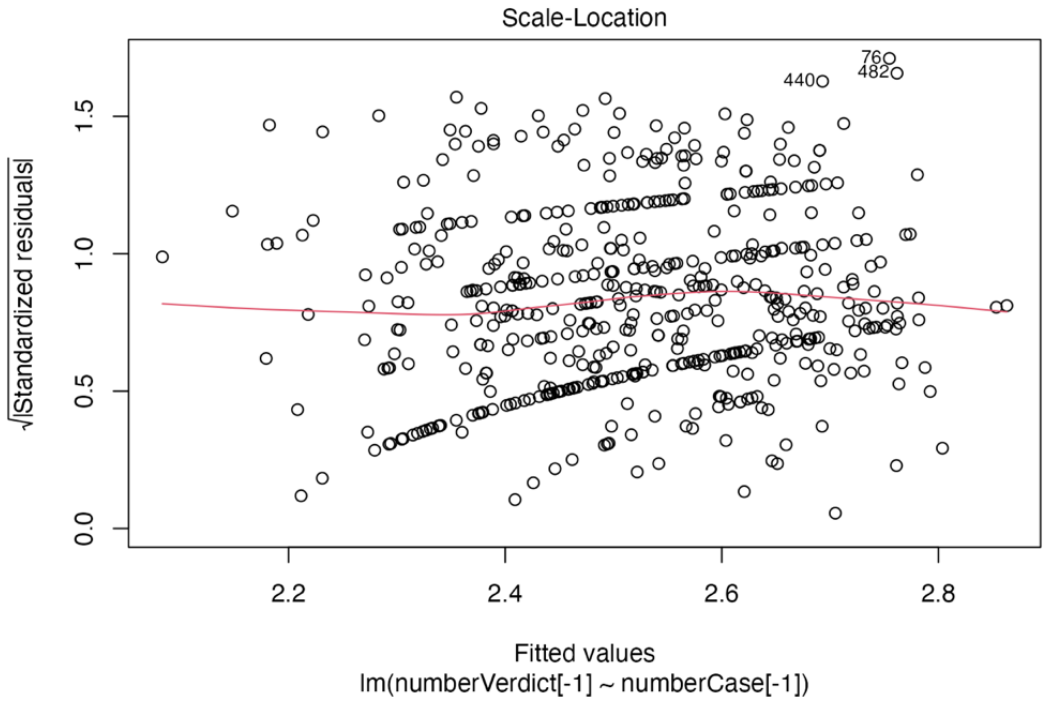


Figure A5.3: model 3 regression diagnostics 3

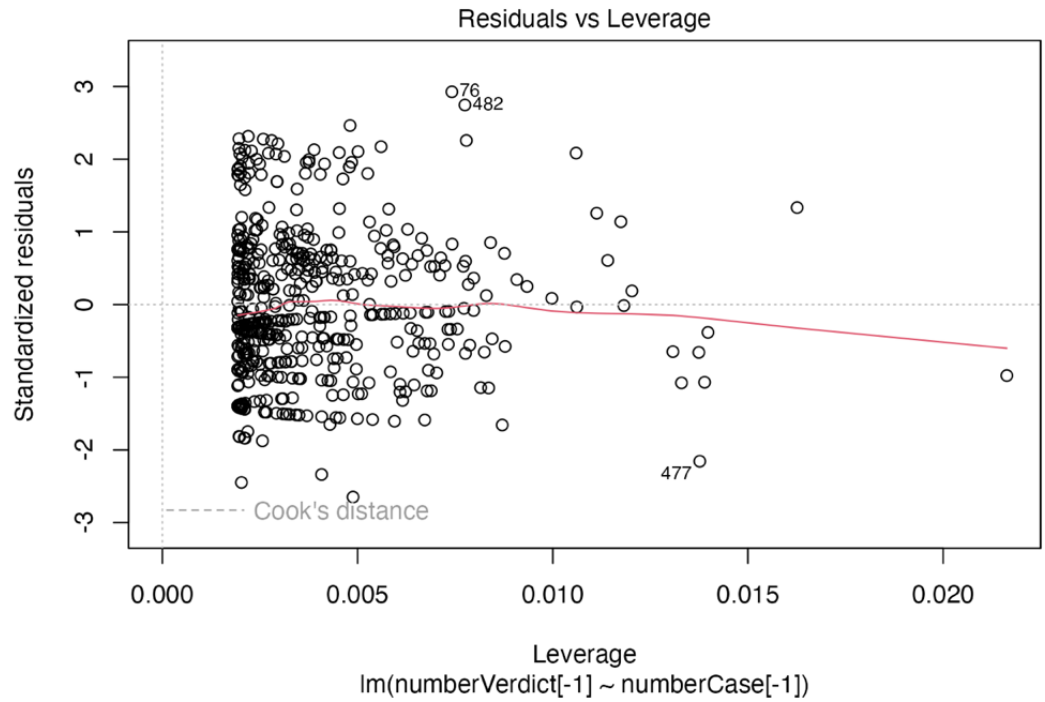


Figure A5.4: model 3 regression diagnostics 4

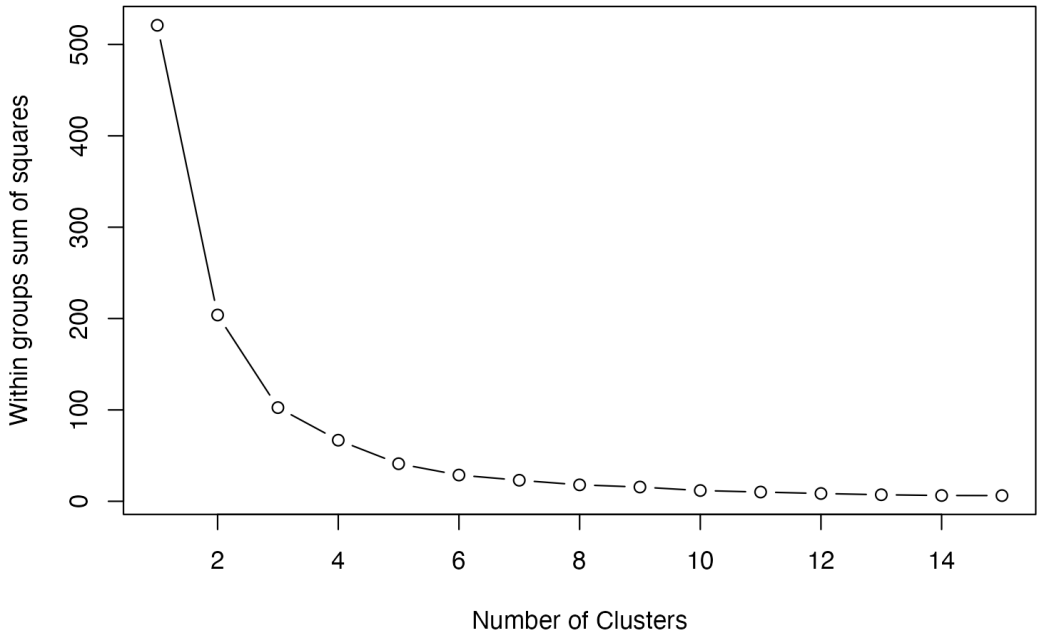


Figure A6.1: elbow method

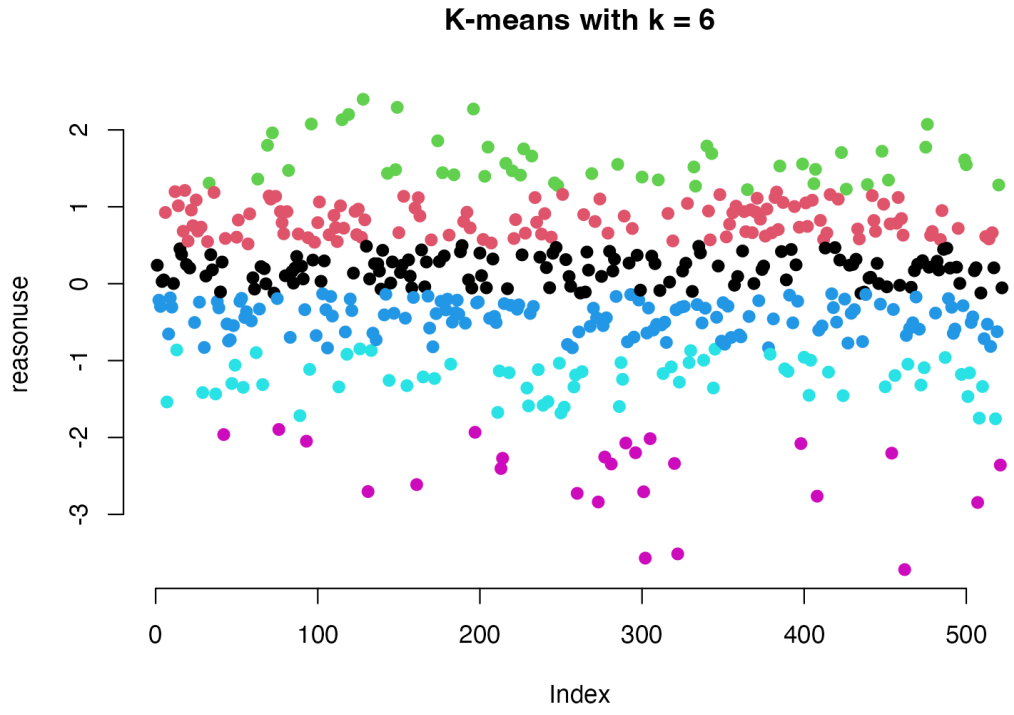


Figure A6.2: clusters

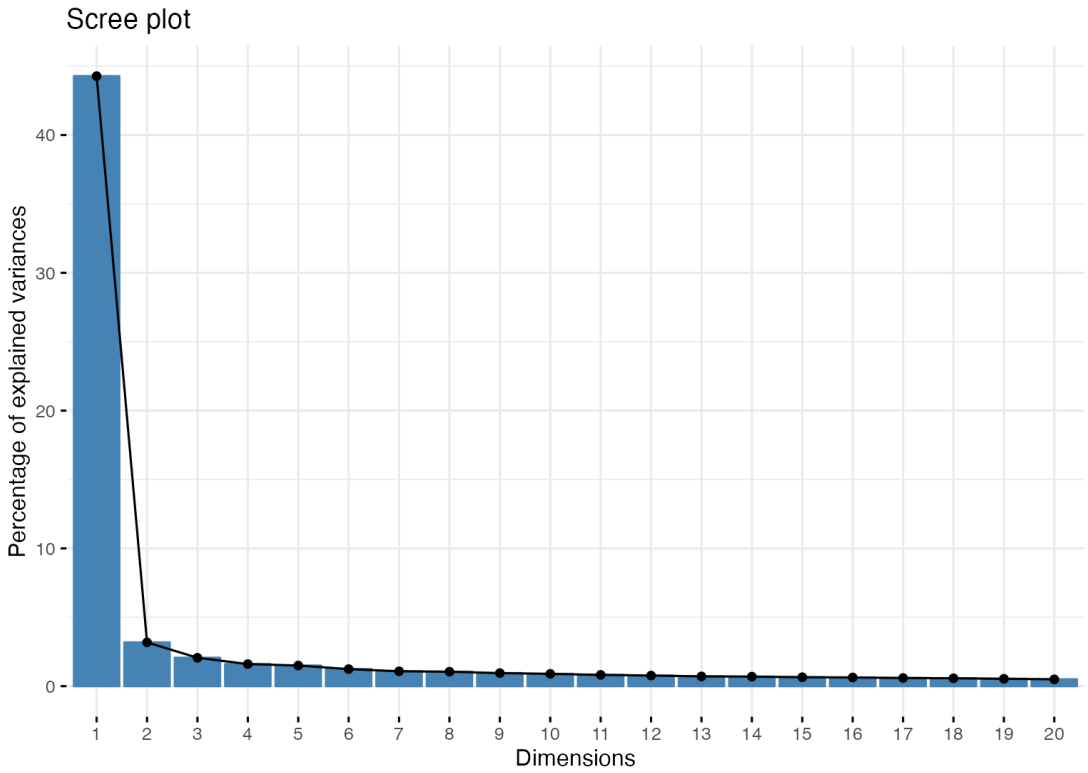


Figure A7.1: reason PCA

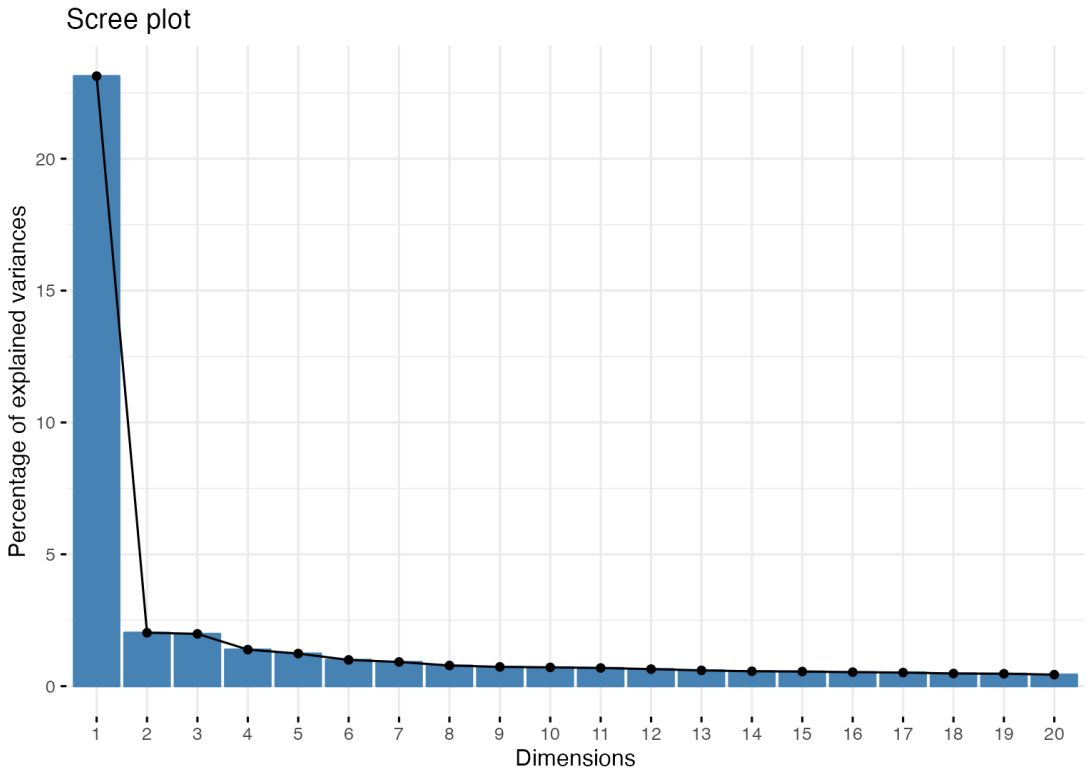


Figure A7.2: case PCA

Tables

Tokenization

Table A1: tokenization

Variance	Percent
刑满释放 （ex-convict）	3.8%
处罚金 （penalty fine）	3.8%
如实 （provide a truthful fact）	3.8%
性质 （severity）	3.8%
持械 （carrying a weapon）	3.8%
滨**区 （this is an area in China）	0.0%
陈*洪 （this is a Chinese name）	0.0%
陈* （this is a Chinese name）	0.0%

Descriptive result

Table A2: descriptive result

	Case	Reason	Sentences	Normalized Sentences
Min	0.03343	0.1805	0.00	0.000
Median	0.22239	0.4707	8.00	2.1792
Mean	0.22262	0.4670	21.02	2.518
Max	1.00000	1.00000	306.00	5.727
Num.Obs.	522	522	518	518

Linear models

Table A3: linear models

	Model 1	Model 2	Model 3
Coefficient	0.279884***	-0.7904	-1.8103**
Num.Obs	522	518	518

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Change the benchmark

Table A4.1: robustness—change benchmark

	Model 1	Model 2	Model 3
No. 1	0.279884***	-0.7904	-1.8103**
No. 55	0.11426*	-1.1626*	-0.7508
No. 166	0.20945**	-0.02763	-0.2158
No. 75	0.43775***	-1.6294**	-1.1743+
No. 434	0.30414***	-2.1481***	-0.8624+
No. 515	0.403635***	-0.5345	0.7304
Num.Obs.	522	518	518

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Text similarity

Table A4.2: robustness

	Model 1	Model 2	Model 3
Jaccard	0.59***	-0.617	0.132
Correlation	0.281703***	-0.7994	-1.8110**
Num.Obs.	522	518	518

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Traditional method

Table A4.3: results of traditional method

	Model
No. 1	-0.66645***
No. 129	-0.78089***
No. 509	-1.1482***
No. 58	0.22397
Num.Obs	518

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

K-means and Logistic regression

Table A4.4.1: k-means clusters

Cluster No.	Value (standardized)
1	1.6081276
2	0.8280640
3	0.1999997
4	-0.4044666
5	-1.2030767
6	-2.4947546

Table A4.4.2: logistic regression

	Model 1	Model 2
Cluster 5	7.208*	-0.161
Cluster 4	11.063**	-0.175
Cluster 3	12.027***	-0.241
Cluster 2	14.157***	-0.284
Cluster 1	20.424***	-0.296
Pr (>Chi-sq)	0.0000***	0.7196
Num.Obs.	522	518

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

PCA

Table A4.5: linear models

	Model 1	Model 2	Model 3
Coefficient	0.2875***	0.3707	0.3541
Num.Obs	522	518	518

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

APPENDIX B

Below is the variables traditional method uses. I divide them into 8 types to illustrate and translate each type into English.

List: variables traditional method uses

Types	Variables
伤情 （injury）	轻微伤、轻伤、重伤、死亡
犯罪数额 （value of property）	较大、巨大、特别巨大
犯罪方法 （crime method）	入户、冒充军警、持枪、持械、转化
犯罪地点 （crime place）	交通工具、银行
犯罪次数和记录 （times and prior crime）	累犯、前科、多次
被告人因素 （defendants' factors）	认罪、自首、坦白、立功
被害人因素 （victim's factors）	谅解、赔偿、和解
其他 （other factors）	黑恶势力、未遂、军用物资、救灾物资、手段特别残忍

The exact meaning of each variable could be found in China's criminal law.

APPENDIX C

This is the policy website page of “Guiding Opinions on Strengthening Case Retrieval to Unify the Application of Law”.

The link is https://law.wkinfo.com.cn/legislation/detail/MTAxMDAxMzgyODI

Figure C1

This is the policy website page of “Guiding Case System”.

The link is https://law.wkinfo.com.cn/legislation/detail/MTAxMDAwMDY0NjM

Figure C2

This is the policy website page of “The Supreme People's Court's Opinions on Accelerating the Construction of Smart Courts”.

The link is https://law.wkinfo.com.cn/legislation/detail/MTAxMDAxMjU5OTk

Figure C3

APPENDIX D

Here is an illustration of cosine similarity for Chinese legal text. I provide a translation for each sentence. However, please remember that translations cannot fit the Chinese texts well.

Table D1: illustration for cosine similarity

	Text 1	Text 2	Text 3
Original Chinese text	我是一个法律人	我是一个刑法学人	我不是一个刑法学人
Translation for text	I am a lawyer	I am a criminal law lawyer	I am not a criminal law lawyer
Tokenization	我/是/一个/法律/人	我/是/一个/刑法学/人	我/不是/一个/刑法学/人
Translation for tokenization	I am a lawyer	I am a criminal law lawyer	I am not a criminal law lawyer
Vector of text	(1, 1, 1, 1, 1)	(1, 1, 1, 0, 1)	(1, 0, 1, 0, 1)
Similarity (Text 1 as benchmark)	1	0.8944272	0.7745967

Notice: the actual factor is decided by the weight of every word. In order to illustrate, I use dummy values to calculate. That will not affect the result of this example.

APPENDIX E

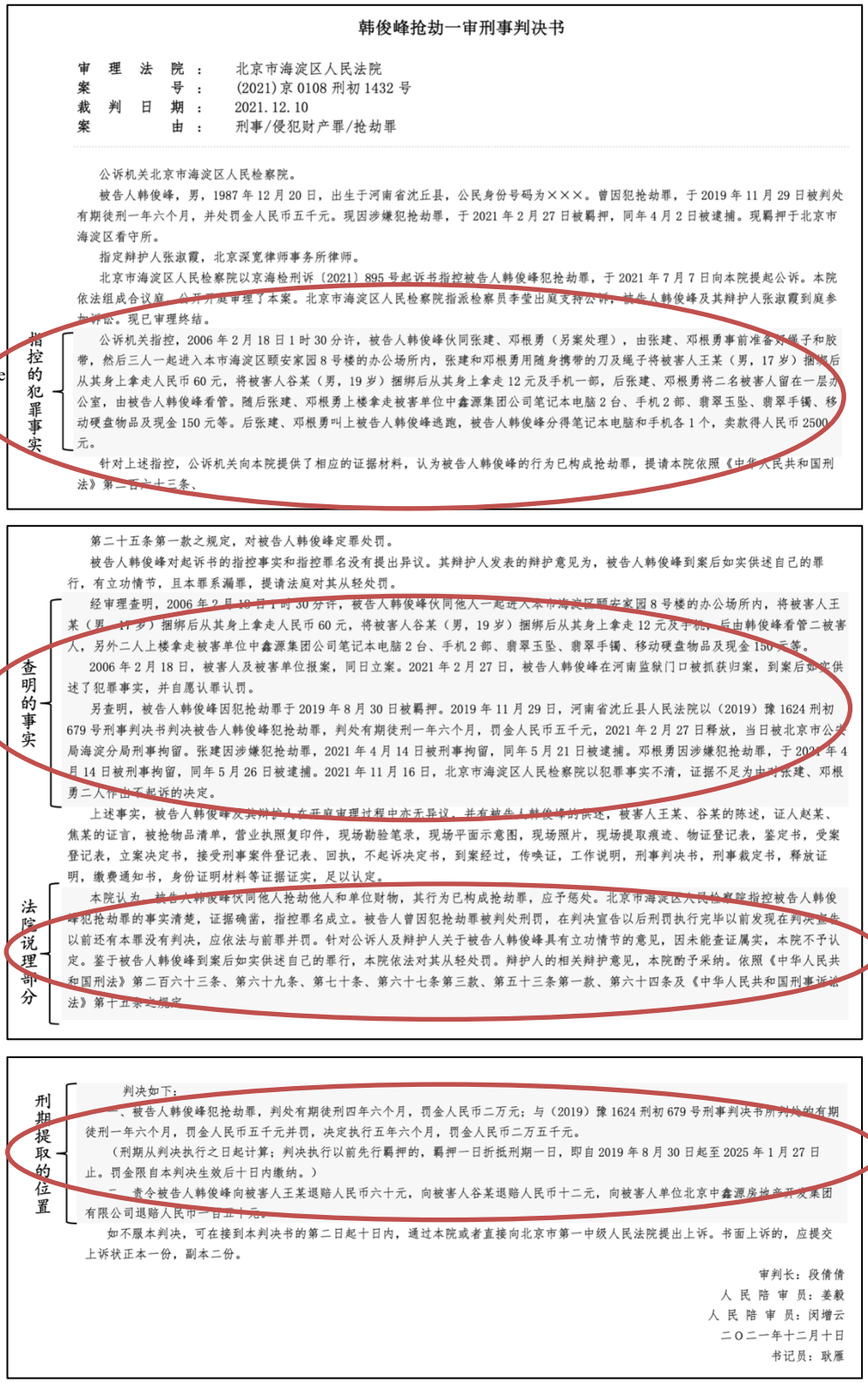
How we divide judicial documents:

Facts Accused by the Authority

Facts Found by the Court

The Court Believes

The Judgment is as Follow



韩俊峰抢劫一审刑事判决书

审 理 法 院 ： 北京市海淀区人民法院
案 号 ： (2021)京 0108 刑初 1432 号
裁 判 日 期 ： 2021. 12. 10
案 由 ： 刑事/侵犯财产罪/抢劫罪

公诉机关北京市海淀区人民检察院。

被告人韩俊峰，男，1987 年 12 月 20 日，出生于河南省沈丘县，公民身份号码为×××。曾因犯抢劫罪，于 2019 年 11 月 29 日被判处有期徒刑一年六个月，并处罚金人民币五千元。现因涉嫌犯抢劫罪，于 2021 年 2 月 27 日被羁押，同年 4 月 2 日被逮捕。现羁押于北京市海淀区看守所。

指定辩护人张淑霞，北京深宽律师事务所律师。

北京市海淀区人民检察院以京海检刑诉〔2021〕895 号起诉书指控被告人韩俊峰犯抢劫罪，于 2021 年 7 月 7 日向本院提起公诉。本院依法组成合议庭，公开开庭审理了本案。北京市海淀区人民检察院指派检察员李莹出庭支持公诉，被告人韩俊峰及其辩护人张淑霞到庭参加诉讼。现已审理终结。

指控的犯罪事实

公诉机关指控，2006 年 2 月 18 日 1 时 30 分许，被告人韩俊峰伙同张建、邓根勇（另案处理），由张建、邓根勇事前准备好绳子和胶带，然后三人一起进入本市海淀区颐安家园 8 号楼的办公场所内，张建和邓根勇用随身携带的刀及绳子将被害人王某（男，17 岁）捆绑后从其身上拿走人民币 60 元，将被害人谷某（男，19 岁）捆绑后从其身上拿走 12 元及手机一部，后张建、邓根勇将二名被害人留在一层办公室，由被告人韩俊峰看管。随后张建、邓根勇上楼拿走被害单位中鑫源集团公司笔记本电脑 2 台、手机 2 部、翡翠玉坠、翡翠手镯、移动硬盘物品及现金 150 元等。后张建、邓根勇叫上被告人韩俊峰逃跑，被告人韩俊峰分得笔记本电脑和手机各 1 个，卖款得人民币 2500 元。

针对上述指控，公诉机关向本院提供了相应的证据材料，认为被告人韩俊峰的行为已构成抢劫罪，提请本院依照《中华人民共和国刑法》第二百六十三条、

第二十五条第一款之规定，对被告人韩俊峰定罪处罚。

被告人韩俊峰对起诉书的指控事实和指控罪名没有提出异议。其辩护人发表的辩护意见为，被告人韩俊峰到案后如实供述自己的罪行，有立功情节，且本罪系漏罪，提请法庭对其从轻处罚。

查明的事实

经审理查明，2006 年 2 月 18 日 1 时 30 分许，被告人韩俊峰伙同他人一起进入本市海淀区颐安家园 8 号楼的办公场所内，将被害人王某（男，17 岁）捆绑后从其身上拿走人民币 60 元，将被害人谷某（男，19 岁）捆绑后从其身上拿走 12 元及手机，后由韩俊峰看管二被害人，另外二人上楼拿走被害单位中鑫源集团公司笔记本电脑 2 台、手机 2 部、翡翠玉坠、翡翠手镯、移动硬盘物品及现金 150 元等。

2006 年 2 月 18 日，被害人及被害单位报案，同日立案。2021 年 2 月 27 日，被告人韩俊峰在河南监狱门口被抓获归案，到案后如实供述了犯罪事实，并自愿认罪认罚。

另查明，被告人韩俊峰因犯抢劫罪于 2019 年 8 月 30 日被羁押。2019 年 11 月 29 日，河南省沈丘县人民法院以（2019）豫 1624 刑初 679 号刑事判决书判决被告人韩俊峰犯抢劫罪，判处有期徒刑一年六个月，罚金人民币五千元，2021 年 2 月 27 日释放，当日被北京市公安局海淀分局刑事拘留。张建因涉嫌犯抢劫罪，2021 年 4 月 14 日被刑事拘留，同年 5 月 21 日被逮捕。邓根勇因涉嫌犯抢劫罪，于 2021 年 4 月 14 日被刑事拘留，同年 5 月 26 日被逮捕。2021 年 11 月 16 日，北京市海淀区人民检察院以犯罪事实不清，证据不足为由对张建、邓根勇二人作出不起诉的决定。

上述事实，被告人韩俊峰及其辩护人在开庭审理过程中亦无异议，并有被告人韩俊峰的供述，被害人王某、谷某的陈述，证人赵某、焦某的证言，被抢物品清单，营业执照复印件，现场勘验笔录，现场平面示意图，现场照片，现场提取痕迹、物证登记表，鉴定书，受案登记表，立案决定书，接受刑事案件登记表、回执，不起诉决定书，到案经过，传唤证，工作说明，刑事判决书，刑事裁定书，释放证明，缴费通知书，身份证明材料等证据证实，足以认定。

法院说理部分

本院认为，被告人韩俊峰伙同他人抢劫他人和单位财物，其行为已构成抢劫罪，应予惩处。北京市海淀区人民检察院指控被告人韩俊峰犯抢劫罪的事实清楚，证据确凿，指控罪名成立。被告人曾因犯抢劫罪被判处刑罚，在判决宣告以后刑罚执行完毕以前发现在判决宣告以前还有本罪没有判决，应依法与前罪并罚。针对公诉人及辩护人关于被告人韩俊峰具有立功情节的意见，因未能查证属实，本院不予认定。鉴于被告人韩俊峰到案后如实供述自己的罪行，本院依法对其从轻处罚。辩护人的相关辩护意见，本院酌予采纳。依照《中华人民共和国刑法》第二百六十三条、第六十九条、第七十条、第六十七条第三款、第五十三条第一款、第六十四条及《中华人民共和国刑事诉讼法》第十五条之规定

刑期提取的位置

判决如下：

一、被告人韩俊峰犯抢劫罪，判处有期徒刑四年六个月，罚金人民币二万元；与（2019）豫 1624 刑初 679 号刑事判决书所判处的有期徒刑一年六个月，罚金人民币五千元并罚，决定执行五年六个月，罚金人民币二万五千元。

（刑期从判决执行之日起计算；判决执行以前先行羁押的，羁押一日折抵刑期一日，即自 2019 年 8 月 30 日起至 2025 年 1 月 27 日止。罚金限自本判决生效后十日内缴纳。）

二、责令被告人韩俊峰向被害人王某退赔人民币六十元，向被害人谷某退赔人民币十二元，向被害人单位北京中鑫源房地产开发集团有限公司退赔人民币一百五十元。

如不服本判决，可在接到本判决书的第二日起十日内，通过本院或者直接向北京市第一中级人民法院提出上诉。书面上诉的，应提交上诉状正本一份，副本二份。

审判长：段倩倩
人 民 陪 审 员：姜毅
人 民 陪 审 员：闵增云
二〇二一年十二月十日
书记员：耿雁