LIFTING THE VEIL OF MONA LISA: A MULTIFACETED INVESTIGATION OF THE "BEYOND A REASONABLE DOUBT" STANDARD

Zhuhao Wang and Eric Zhi***

TABLE OF CONTENTS

INTRODUCTION 122

I. THE MYTH OF "PROOF BEYOND A REASONABLE DOUBT" 127
- *A. The Unsettled Origin Of The Reasonable Doubt Standard* 128
- *B. The Unclear Definition Of The Reasonable Doubt Standard* .. 133

II. GLORIES OF THE REASONABLE DOUBT STANDARD 139
- *A. The Reasonable Doubt Standard As Constitutionally Required In Criminal Cases* 140
- *B. The Reasonable Doubt Standard As An Essential Tool To Restrain Wrongful Convictions* 143

III. DARKNESS OF THE BEYOND A REASONABLE DOUBT STANDARD 152
- *A. The Confusion Surrounding The Reasonable Doubt Standard* 152
- *B. Negative Byproducts Of The Reasonable Doubt Standard* 154

---

* Zhuhao Wang (汪诸豪) is an associate professor of law at the Institute of Evidence Law and Forensic Science, China University of Political Science and Law; LL.M., University of Pennsylvania Carey Law School, J.D., Indiana University Bloomington Maurer School of Law, S.J.D. candidate, Northwestern University Pritzker School of Law. Email: wangzhuhao@cupl.edu.cn. Special thanks to Professor Ronald J. Allen, who taught me evidence law and provided helpful comments and suggestions on this article. I am also grateful for inspiring comments from Edward J. Imwinkelried, Edward K. Cheng, Andrew Ligertwood, Julia Simon-Kerr, Barbara Shapiro, Alex Biedermann and Yuguan Yang.

** Eric Zhi (支金成) is an associate attorney at Kirkland & Ellis LLP; J.D., University of Chicago, A.B.D. in Economics, University of Toronto. I would like to thank colleagues and friends at University of Chicago and University of Toronto for their comments and support. All errors are mine.

IV. THE EFFECT OF THE BEYOND A REASONABLE DOUBT STANDARD IN PRACTICE: EMPIRICAL MEASUREMENT AND ITS LIMITS 156
A. *Thinking Empirically: The Problem Of Empirical Study On The Reasonable Doubt Standard* 157
i. *Data on wrongful convictions* 159
ii. *Data on wrongful acquittals* 160
B. *An Alternative Empirical Model For Testing The Effect Of The Reasonable Doubt Standard On Restraining Wrongful Convictions* 161

V. A CASE STUDY ON CHINA 163
A. *The Unique Opportunity Of China As A Test Case* 164
B. *The Chinese Judicial Data Under Study* 167
C. *The Roadmap: Recap Of The Three Indicators In The Context Of China* 169
i. *Indicator 1: Acquittal rate* 170
ii. *Indicator 2.1: Number of criminal cases that were reversed and remanded* 171
iii. *Indicator 2.2: Number of cases that were reversed and acquitted in China's "second instance trial"* 171
iv. *Indicator 3: Non-prosecution rate* 172
D. *Analysis: Application And Implications Of The Chinese Judicial Data* 174
i. *Reflection on Indicator 1: Impact of the reasonable doubt standard on acquittal rate* 174
ii. *Reflection on Indicator 2.1: Impact of the reasonable doubt standard on criminal cases of reversal and remand* 180
iii. *Reflection on Indicator 2.2: Impact of the reasonable doubt standard on cases of reversal and direct acquittal* 182
iv. *Reflection on Indicator 3: Impact of the reasonable doubt standard on non-prosecutions* 185
E. *Limitations Of The Chinese Case Study* 187

VI. REFLECTION AND CONCLUSION 189

ABSTRACT

For a long period of time, the golden standard in judicial fact-finding of criminal cases in the United States and many other countries has been the "Beyond a Reasonable Doubt" (BARD) standard – every person accused of a crime is presumed to be innocent unless, and until, his or her guilt is established beyond a reasonable doubt.[1] *The BARD standard's undergirding principle is one of error distribution, where wrongful conviction of the innocent is a much greater wrong than failed conviction of the guilty. This concept was famously expressed by the English jurist William Blackstone in 1760s: "It is better that ten guilty persons escape than that one innocent suffer."*[2] *This principle is widely regarded as "the Mount Everest of legal mantras;"*[3] *however, a closer look at the BARD standard reveals many potential doubts: What is meant by eliminating reasonable doubt? Where does this standard come from? There is no clear answer so far. Given the passage of time, does the Blackstone Principle still accurately reflect the policy or value choice of today's society? How can empirical studies be done on the effectiveness of the BARD standard in practice?*

This article seeks to give the BARD standard the careful attention it deserves by querying it from historical, epistemological, and empirical perspectives. Even though each of these three areas of inquiry involves its own unique controversy or difficulty, we can come even closer to an understanding of BARD by studying its origins, its ideological underpinnings, and its workings in practice. In particular, this article proposes a novel empirical formula to approximate the effect of BARD in restraining wrongful convictions through the lens of indirect but available statistical data on acquittals, reversals, and non-prosecutions. The best way to study this venerable common-law notion is to examine its recent implementation in a civil law system, namely, China's legal system. We identify operations of the criminal justice system in China between 2007 and 2018 as the most suitable for focusing this empirical study because China began implementing the BARD standard in 2013 as a complement to its existing standard of proof in criminal cases – the "reliable and sufficient evidence" standard. Our study shows that although implementation of the BARD standard in the Chinese

---

[1] *See* Coffin v. United States, 156 U.S. 432, 453 (1895).

[2] 4 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAW OF ENGLAND 352 (1769).

[3] Laura I. Appleman, *A Tragedy of Errors: Blackstone, Procedural Asymmetry, and Criminal Justice*, 128 HARV. L. REV. F. 91 (2015). *See* the National Registry of Exonerations, https://www.law.umich.edu/special/exoneration/Pages/about.aspx (*last visited* September 22, 2021) (Even so, according to U.S. National Registry of Exonerations, a project collecting and publicizing information about all known exonerations of innocent criminal defendants in the United States, there were 2,869 exonerations between 1989 and September 2021, and those wrongfully convicted had collectively lost more than 25,600 years to prison).

criminal justice system has been generally helpful in restraining wrongful convictions, its effect is limited. The two standards are effectively treated the same in practice. China's move to the BARD standard was only a formal change, rather than a substantive modification to its criminal justice system. Our empirical study is a demonstration that the formal standard of proof may not matter all that much. Courts have an intuitive understanding of the high standard of proof required in the criminal context and are going to apply it accordingly, no matter whether we call it "reliable and sufficient" or "beyond a reasonable doubt".

INTRODUCTION

It is difficult, if not impossible, to so define the term "reasonable doubt" as to satisfy a subtle and metaphysical mind, bent on the detection of some point, however attenuated, upon which to hang a criticism. --- Supreme Court of Virginia (1905)[4]

The truth is that no one has yet invented or discovered a mode of measurement for the intensity of human belief. Hence there can be yet no successful method of communicating intelligibly . . . a sound method of self-analysis for one's belief. --- John Henry Wigmore (1940)[5]

The same points are true in criminal cases: The effect of burdens of persuasion cannot be determined analytically and neither can the effect of a change in the burden of persuasion be determined analytically. They are both empirical questions. --- Ronald J. Allen (2014)[6]

The long-standing standard of proof for criminal trials in Anglo-American legal systems—"beyond a reasonable doubt" (BARD)—is a *magic* phrase. It is like a charm that, when spoken, drives the legal system. Take the criminal justice system in the United States: it is *magically important in theory* that as the cornerstone to American criminal jurisprudence, an accused shall be presumed innocent until the government proves guilt beyond a reasonable doubt. A criminal defendant shall be found not guilty and thus acquitted if the BARD standard is not satisfied.[7] The standard is also *magically powerful in*

[4] McCue v. Commonwealth, 49 S.E. 623, 629 (Va. 1905).
[5] 9 JOHN HENRY WIGMORE, EVIDENCE A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW 325 (3d ed. 1940).
[6] Ronald J. Allen, *Burdens of Proof*, 13 LAW, PROBABILITY AND RISK 195, 212 (2014).
[7] *See, e.g.*, *In re* Winship, 397 U.S. 358, 363 (1970) (quoting Coffin v. United States, 156 U.S. 432, 453 (1895)); *see also* Daniel Epps, *The Consequences of Error in Criminal Justice*, 128 HARV. L. REV. 1065, 1073 (2015) (noting that if even one juror insists that the

practice: in the federal court system, as well as in most state courts, the criminal trial judge instructs the jurors that in order to convict the defendant, they need to unanimously find that the prosecutor has proved beyond a reasonable doubt every fact necessary to constitute the crime charged. A judge's failure to so instruct a jury constitutes automatic grounds for reversing a conviction.[8]

While it is magically important and magically powerful, the BARD standard is also *magically mysterious*. The origins and early development of this standard are still the subject of controversy, even after centuries of usage. And, as is more commonly known, not only does no settled definition of BARD exist, but efforts to firmly define it have been met with resistance; as the chief judge of a U.S. Court of Appeals once said, "I find it rather unsettling that we are using a formulation that we believe will become less clear the more we explain it."[9] This article seeks to meet this challenge, not by *explaining* the BARD standard but by querying it from multiple perspectives: historical, epistemological, and, finally, empirical. Even though no one, it seems, can fully explicate this concept—which slips away from us with any attempt to do so—we can come closer to an understanding of it by studying its origins, its ideological underpinnings, and its effect in practice.

Each of these areas of inquiry involve a unique controversy or difficulty. With regard to the origins and definition of BARD, the issue is a lack of agreement; how we define BARD, and how it even came to be, has been the topic of much discussion. BARD is accepted in many different criminal justice systems as the proper formula to set out the burden of persuasion, but no one can pin down its precise birth date[10] or convincingly demarcate its early development.[11] Everyone knows that BARD is a high

---

BARD standard is not satisfied, the result is a hung jury and the defendant may not be punished, though he may be retried).

[8] *See* Fed. R. Crim. P. 31(a); Apodaca v. Oregon, 406 U.S. 404 (1972); Johnson v. Louisiana, 406 U.S. 356 (1972); Michael H. Glasser, Comment, *Letting the Supermajority Rule: Nonunanimous Jury Verdicts in Criminal Trials*, 24 FLA. ST. U. L. REV. 659, 671 (1997) (noting that only two states permit conviction by non-unanimous juries). *See generally In re* Winship, 397 U.S. 358 (1970).

[9] Jon O. Newman, *Beyond "Reasonable Doubt,"* N.Y.U. L. REV. 979, 984 (1993).

[10] Most scholars nowadays agree that the Boston Massacre trial (1770) in which five British soldiers were indicted and tried for murder of five civilians in front of the State House of colonial Massachusetts in Boston was the earliest instance presently known in which the beyond a reasonable doubt was articulated. *See* Anthony A. Morano, *A Reexamination of the Development of the Reasonable Doubt Rule*, 55 B.U. L. REV. 507, 508 (1975). *But see* John Wilder May, *Some Rules of Evidence: Reasonable Doubt in Civil and Criminal Cases*, 10 AM. L. REV. 642 (1876) (claiming that the reasonable doubt standard originated in the Irish Treason in 1798).

[11] *See, e.g.*, Miller W. Shealy, Jr., *A Reasonable Doubt about "Reasonable Doubt,"* 65 OKLA. L. REV. 225, 270 (2013) ("There are currently two lines of thought regarding the development of the 'proof beyond a reasonable doubt' standard. For lack of a better form

standard of proof, higher than the "preponderance of evidence" and the "clear and convincing evidence" standards in civil cases, but no one knows exactly how high in degree it reaches.[12] Nor is there common understanding on what "reasonable doubt" means.[13]

When we turn to the BARD standard's ideological and epistemological underpinnings, we find that the issue shifts from a lack of agreement among scholars to an inherent contradiction or paradox: the BARD standard supports a desired *distribution of errors*—with more wrongful acquittals than wrongful convictions, in keeping with the Blackstone principle—but in doing so it increases the overall error rate.[14] Our discussion about this incongruity between what we term the *first* and *second* goals of the criminal justice system (error reduction and error distribution, respectively) is supported by a series of illustrative diagrams.[15] From an epistemological perspective, there are four (and only four) possibilities with respect to the result of a criminal case adjudicated: correct conviction, correct acquittal, wrongful conviction, and wrongful acquittal.[16] Since adjudicating errors will occur from time to time, the question at hand is which sort of error—wrongful acquittal or wrongful conviction—is more serious and, thus, more earnestly to be avoided.[17] The Blackstone principle holds firmly that convicting an innocent person is a significantly more costly mistake than acquitting a guilty one, with Sir William Blackstone stating that "it is better that ten guilty persons escape than that one innocent suffer."[18] The BARD standard helps

---

of expression, the first view will be termed the 'received view.' The second view is the theological analysis recently offered by Professor James Q. Whitman in his book *The Origins of Reasonable Doubt: Theological Roots of the Criminal Trial.*"). For a more detailed discussion, see Part I-A.

[12] *See, e.g.*, United States v. Fatico, 458 F. Supp. 388, 409–10 (E.D.N.Y. 1978), aff'd, 603 F.2d 1053 (2d Cir. 1970).

[13] *See, e.g.*, Thomas V. Mulrine, *Reasonable Doubt: How in the World Is It Defined?* 12 AM. U. J. INT'L L. & POL'Y 195 (1997); Shealy, *supra* note 11, at 225 ("The [U.S.] Supreme Court has failed to define the concept of 'reasonable doubt' with any precision. The Court tolerates conflicting definitions of 'reasonable doubt.' It permits some jurisdictions to forbid any definition of 'reasonable doubt,' while giving others wide latitude to define the concept in ways that are contradictory."). For a more detailed discussion, see Part I-B.

[14] *See* Epps, *supra* note 7, at 1068.

[15] *See* Figures 1-3 in Part II-B.

[16] For a more detailed discussion, see Part II-B.

[17] *See* LARRY LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW: AN ESSAY IN LEGAL EPISTEMOLOGY 1 (2006) [*Hereinafter* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW].

[18] 4 WILLIAM BLACKSTONE, COMMENTARIES 352 (1769); *See also* Alexander Volokh, *n Guilty Men*, 146 U. PA. L. REV. 173, 174 (1997) ("The ratio 10:1 has become known as the 'Blackstone ratio.'"). *But see* JEREMY BENTHAM, PRINCIPLES OF JUDICIAL PROCEDURE 169 (1829); Daniel Epps, *The Consequences of Error in Criminal Justice*, 128 HARV. L. REV. 1065 (2015); Marvin Zalman, *The Anti-Blackstonians: Thinking About Different Strokes*, 48 SETON HALL L. REV. 1319 (2018).

ensure that this imbalance occurs.[19] And yet, as we show, not only do wrongful acquittals come with problems of their own, but—more concerningly—when the standard of proof becomes higher, the total number of wrongful findings increases.[20] Implementation of the BARD standard risks overshadowing the ultimate goal of any criminal justice system – to find the truth and promote overall accuracy in trial outcomes.[21]

Finally, the question of how BARD works in practice lands us in the realm of empirical inquiry, in which the issue is a lack of measurable evidence. Law is ultimately a profession that shapes the way in which individuals and institutions behave, and any legal rule—but especially one as powerful and controversial as the BARD standard—should ideally be examined empirically for observable outcomes. As Larry Laudan once complained, the legal field is faced with a "paucity of empirical information" and thus must "fall back on armchair hunches about the likely effects of various rules and procedures."[22] Data on the criminal justice system's wrongful convictions and wrongful acquittals could allow us to assess: (a) whether the actual ratio between wrongful convictions and wrongful acquittals accords with the Blackstone principle; and (b) whether the total number of erroneous verdicts overburdens the criminal justice system. However, such data is difficult if not impossible to come by.[23] As Keith A. Findley describes:

---

[19] *See, e.g.*, Rinat Kitai, *Protecting the Guilty*, 6 BUFF. CRIM. L. REV. 1163 (2003); ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 172–213 (2005); Richard L. Lippke*, Punishing the Guilty, not Punishing the Innocent*, 7 J. MORAL PHILOSOPHY 462 (2010); Youngjae Lee, *Deontology, Political Morality, and the State*, 8 OHIO ST. CRIM. L. 385 (2011); Alec Walen, *Proof Beyond a Reasonable Doubt: A Balanced Retributive Account*, LA. L. REV. 355 (2015).

[20] For a more detailed discussion, see Part III-B.

[21] For a more detailed discussion, see Part III-B.

[22] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 8. Note: For his estimates on wrongful acquittals, Laudan relied primarily on Harry Kalven and Hans Zeisel's *1966* study of judge and jury agreement and disagreement. HARRY KALVEN & HANS ZEISEL, THE AMERICAN JURY 8-9 (1971). For example, turning to Kalven and Zeisel's study, Laudan notes the rate at which judges would have convicted when juries acquitted, which Kalven and Zeisel found to be about 17% of the criminal cases. Laudan further notes that, among the cases included in Kalven and Zeisel's study, the judges reported that they believed that only about 15% of the acquittals were clear acquittals, and 85% were close cases. Laudan then takes at face value that this means that the close cases—85% of the acquittals— "are close enough to warrant an assumption that these are probably factually guilty defendants." *See* LARRY LAUDAN, THE LAW'S FLAWS: RETHINKING TRIAL AND ERRORS? 58–59 (2016) [*Hereinafter* LAUDAN, THE LAW'S FLAWS]. Such an analysis is not logically rigorous.

[23] *See* Keith A. Findley, *Reducing Error in the Criminal Justice System*, 48 SETON HALL L. REV. 1265, 1266 (2018).

> Only in the exoneration context do we *sometimes* have the ability to discover error, through the production of new evidence of innocence (especially DNA evidence). This tells us that error does occur, and does so at above a trivial level, which can help us understand the conditions and contributing factors that can lead to false positives. It cannot, however, tell us a rate. For false negatives (false acquittals), . . . we do not even have that.[24]

In the absence of direct data, this article proposes a novel empirical formula, maybe the only one in existence, to approximate the effect of the BARD standard in practice through the lens of indirect, but statistically available, data on acquittals, reversals, and non-prosecutions. Our foundational assumption is that the more effective the BARD standard is in restraining wrongful convictions, the more that the "unobservable" number of wrongful convictions will decline, while the observable number of acquittals, reversals, and non-prosecutions will rise, and vice versa. Unfortunately, even with the turn to indirect data, the possible real-world laboratories for empirical observation are exceedingly limited. Such a study needs a clear benchmark date on which the BARD standard was implemented in (or removed from) a given criminal justice system, as well as some length of time both pre- and post-benchmark, from which to collect data. Moreover, all other (non-BARD) factors would need to remain relatively stable throughout the entire period of observation. Without this combination of a clear benchmark date and a *ceteris paribus* environment, even empirical analysis on the indirect but observable data is largely impossible.[25]

The solution to this dilemma of how to investigate the BARD standard empirically presented itself, surprisingly, not in a common law country—where BARD has been in place for more than two centuries and no clear benchmark date exists—but in the civil law country of China. As with other civil law countries, China developed a criminal standard of proof different from that of BARD; however, in 2013, China officially adopted the BARD standard—complementing its existing standard of proof, "reliable and sufficient evidence"—in a move that otherwise left the legal system unchanged.[26] This development offers the necessary "kicker" for our study, allowing us to observe changes in acquittal, reversal, and non-prosecution

---

[24] *Id.* at 1275 (emphasis added).

[25] Note: *Ceteris paribus* is a Latin phrase that means "all other things being equal" or "all other things being normal." For a detailed discussion on this notion, *see* Johan van Benthem, Patrick Girard, & Oliver Roy, *Everything Else Being Equal: A Modal Logic for Ceteris Paribus Preferences*, 38 J. PHILOS. LOGIC 83, 84 (2009).

[26] *See* Zongzhi Long, *"Beyond Reasonable Doubt" in the Chinese Legal Context*, 1 PEKING U. L.J. 339 (2014).

rates after the benchmark date and thus extrapolate the effect of the BARD standard in restraining wrongful convictions. Using data that has been made available as part of China's move toward judicial transparency and digitalization of court files nationwide, we conducted a series of analyses with simple regression and fixed effect models to draw further insights on the effect of the BARD standard in the Chinese criminal justice system.[27]

Our study shows that, although implementation of the BARD standard in the Chinese criminal justice system has been generally helpful in restraining wrongful convictions, this effect is limited. Instead, the BARD standard and the "reliable and sufficient evidence" standard have been effectively treated as the same in practice. This finding suggests that the BARD standard itself is not all powerful or indispensable (it is not, in the end, magical).

This article proceeds in six parts. Part I examines the myth of the BARD standard, including its disputed origins and definition. We suggest that the BARD standard is like Mona Lisa's smile, deadly attractive but ambiguous. In Parts II and III, we examine the highs and lows of the BARD standard, respectively: on the one hand, it is constitutionally based and is an essential tool to restrain wrongful convictions; on the other hand, it is surrounded by confusion and can lead to both excessive wrongful acquittals and pressure on overall accuracy of trial outcomes. Part IV introduces our empirical approach on studying the effect of the BARD standard in restraining wrongful convictions, after which Part V presents the Chinese case study. Finally, Part VI concludes our findings that how the formal standard of proof is phrased may not matter very much. Words effectively cannot describe the standard adequately.

I. THE MYTH OF "PROOF BEYOND A REASONABLE DOUBT"

Proof beyond a reasonable doubt (BARD) is a fundamental requirement of Anglo-American criminal law and other common law systems.[28] And yet, the standard is plagued by imprecision and a lack of

---

[27] For a more detailed discussion, see Part V-D.

[28] *See* Alec Walen, *Proof Beyond a Reasonable Doubt: A Balanced Retributive Account*, 76 LA. L. REV. 355, 356, 363 (2015) ("The standard of proof for criminal convictions in the United States—and in many other jurisdictions—is proof beyond a reasonable doubt.... The English-speaking world fully embraces the BARD standard, and is joined by a growing swath of Western Europe, including Germany, Sweden, and most recently, Italy."); *see also* Thomas V. Mulrine, *Reasonable Doubt: How in the World is It Defined?*, 12 AM. U. J. INT'L L. & POL'Y 195, 214-18 (1997) (Australia, Canada, and the U.K.); Warren Young, Neil Cameron, & Yvette Tinsley, *Juries in Criminal Trials*, *Part Two*, 54 (N.Z. Law

clarity, with even its origins disputed. We begin our discussion with an examination of the contested origins and definition of BARD.

A. *The unsettled origin of the reasonable doubt standard*

Many researchers would agree on at least two points regarding the birth of the BARD standard: first, it has ancient roots, with classical Roman law and medieval canon law requiring a high standard of proof in criminal cases for the protection of innocent defendants;[29] and second, it crystallized at the end of the eighteenth century,[30] with the earliest use of the precise phrase appearing in the Boston Massacre trials of 1770.[31] However, what happened between these two nodes on the timeline remains unclear; at least three schools of thought exist on this issue.

A leading viewpoint, known as the "Received View"[32] and represented most ably by scholars Barbara Shapiro and Anthony Morano, among others, is that the reasonable doubt standard evolved from the common law jury trial system that spawned in England during the twelfth century.[33] In

---

Comm'n, Preliminary Paper No. 37, 1999) (N. Z.); Criminal Procedure Act 51 of 1977 § 105A(2)(a)(i) (S. Afr.).

[29] *See, e.g.*, Shealy, *supra* note 11, at 269–70 ("Shortly after AD 382, when Christianity had been declared the official faith of the Roman Empire, the emperors ruled that a verdict could only be had if it was based on 'indubitable evidence' (*Indiciis Indubitatis*). . . . By the early ninth century, the canon lawyers also required a high standard of proof. Following the Gospel of Matthew, they demanded evidence that was 'clear as the noon-day sun' (*luce meridiana clarior*)." [Italics added]).

[30] *See* JOHN H. LANGBEIN, RENÉE LETTOW LERNER & BRUCE P. SMITH, HISTORY OF THE COMMON LAW: DEVELOPMENT OF ANGLO-AMERICAN LEGAL INSTITUTIONS 696–97 (2009).

[31] *See id.* at 697; Morano, *supra* note 10, at 508. *But see* May, *supra* note 10, at 656 (noting that Earlier historical literature long focused on a series of treason trials in Ireland in 1798 as the first appearance of the "beyond a reasonable doubt" standard.). It is worth noting that in the Boston Massacre trials of 1770, the reasonable doubt standard was urged by prosecutors who were trying to *lower* their burden of persuasion from an often unattainable task of having to persuade the jury beyond all doubt, an unusual way of using the standard from today's perspective. In contrast, in the Irish Treason Cases in 1798, the reasonable doubt standard was urged by defense lawyers who were endeavoring to *raise* the prosecution's burden of persuasion, which is in accordance with common practices nowadays. *See* Newman, *supra* note 9, at 981-982.

[32] Note: the phrase "Received View" captures the emphasis of this theory on the origin of BARD: toward the end of the Middle Ages, the common law jury underwent its epochal transformation from active neighborhood investigators to passive triers of fact who *received* and evaluated evidence presented at trial. *See* Shealy, *supra* note 11, at 271–77.

[33] *See* BARBARA J. SHAPIRO, BEYOND "REASONABLE DOUBT" AND "PROBABLE CAUSE" 3 (1991) (tracing the development of the reasonable doubt standard); Barbara J. Shapiro, *"To a Moral Certainty": Theories of Knowledge and Anglo-American Juries 1600-1850*, 38 HASTINGS L.J. 153, 155 (1986); Morano, *supra* note 10, at 507; *see also* Theodore

the medieval era of intensely interdependent agricultural communities in England, jurors were drawn from neighborhoods near the contested events.[34] The norm was that a jury of the locality would contain witness-like persons who would know most of the facts by themselves, or if not, that these jurors would at least be well positioned to investigate the facts on their own (a "self-informing jury").[35] "The medieval jury came to court not to listen, but to speak, not to hear evidence but to deliver a verdict formulated in advance."[36] In contrast, medieval judges knew significantly less about the facts of cases adjudicated than did the jurors, and judges accepted this "rough verdict [ ] without caring to investigate the logical processes [of the jury decision-making], if logical they were, of which that verdict was the outcome."[37] Toward the end of the medieval era, due to societal expansion, the volume of both civil and criminal cases surged.[38] Jurors no longer did their own investigations; nor did they always have sufficient personal knowledge of the defendant to make decisions without hearing additional evidence or testimony from witnesses.[39]

As more and more jurors became *passive* viewers of fact and increasingly relied on witnesses and documents presented at trial that had to be evaluated for truthfulness and accuracy (under the "progressive dethronement of the jury"),[40] it became necessary to develop *standards of evaluation* by which jurors could test and weigh evidence. This process of developing "jury control"[41] marks the beginning of the formal "reasonable

---

Waldman, *Origins of the Legal Doctrine of Reasonable Doubt*, 20 J. HIST. IDEAS 299 (1959); Steve Sheppard, *The Metamorphoses of Reasonable Doubt: How Changes in the Burden of Proof Have Weakened the Presumption of Innocence*, 78 NOTRE DAME L. REV. 1165, 1165-169 (2003).

[34] *See* John H. Langbein, *Historical Foundations of the Law of Evidence: A View from the Ryder Sources*, 96 COLUM. L. REV. 1168, 1170 (1996).

[35] *See* James Bradley Thayer, *The Jury and Its Development*, 5 HARV. L. REV. 295, 302–05 (1892) ("From the beginning of our records, we find cases, in a dispute over the genuineness of a deed, where the jury are combined with the witnesses to the deed.").

[36] *See* Langbein, *supra* note 34, at 1170.

[37] *See* FREDERICK POLLOCK & FREDERIC MAITLAND, THE HISTORY OF ENGLISH LAW BEFORE THE TIME OF EDWARD I 660-61 (2d ed. 1898).

[38] *See* Daniel Klerman, *Was the Jury Ever Self-Informing?*, 77 S. CAL. L. REV. 123, 145 (2003) ("Undoubtedly, increased mobility and other social changes played a large role.").

[39] *See* BARBARA J. SHAPIRO, BEYOND "REASONABLE DOUBT" AND "PROBABLE CAUSE" 4 (1991); Langbein, *supra* note 34, at 1170–71; Thayer, *supra* note 35, at 305 ("As late, certainly, as 1489. . . . [] we find witnesses to deeds still summoned with the jury. I know of no later case.").

[40] *See* A. W. B. Simpson, *The Horwitz Thesis and the History of Contracts*, 46 U. CHI. L. REV. 533, 600 (1979).

[41] *See* BARBARA J. SHAPIRO, BEYOND "REASONABLE DOUBT" AND "PROBABLE CAUSE" 6 (1991); *see also* LANGBEIN ET AL, *supra* note 30, at 698.

doubt" rule in England.[42] The articulation of the BARD standard of proof was part of a "broader phenomenon of increasingly detailed judicial instructions to the jury, which had the effect of transforming into 'law' matters that had previously been remitted to jury discretion."[43]

Influenced by the Protestant Reformation in the sixteenth century and the Enlightenment in the seventeenth and eighteenth centuries, "philosophers and jurists [in England] came to realize that in [juridical proof and other] human affairs (as opposed, say, to mathematics or logic), no full certainty was to be had." The next best thing, according to such philosophers as John Locke and John Wilkins, was what they called "moral certainty," a standard of proof that was short of absolute certainty but more precise than mere opinion, and that had evolved from religious and scientific arguments concerning proof.[44] Laudan noted that these philosophers "dubbed this sort of certainty as 'moral,' not because it had anything to do with ethics or morality but because they wanted to contrast it with 'mathematical' certainty of the sort traditionally associated with rigorous demonstration. Morally certain beliefs could not be proven beyond all doubt, but they were nonetheless firm and settled truths, supported by [adequate evidence]."[45] Although in theory they were open to the doubt of skeptics, such beliefs were firm enough that reasonable men, employing their senses and rational faculties, would not doubt them in practice.[46] Laudan further noted that, "[f]rom this [notion of moral certainty] arose the [precise] notion that a guilty verdict required the jury to believe 'beyond a reasonable doubt' [(BARD)] in the guilt of the accused."[47]

A second school of thought on the origin of the BARD standard holds that "the reasonable doubt formula was originally concerned" not with protecting the innocent from conviction, but "with protecting the souls of the *jurors* against damnation."[48] James Whitman, the leading scholar on this theological line of thought, describes it as follows:

> [C]onvicting an innocent defendant was regarded, in the older Christian tradition, as a potential mortal sin. The

[42] BARBARA J. SHAPIRO, BEYOND "REASONABLE DOUBT" AND "PROBABLE CAUSE" 6 (1991).
[43] *See* LANGBEIN ET AL, *supra* note 30, at 698.
[44] LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 32–33; *see* BARBARA J. SHAPIRO, BEYOND "REASONABLE DOUBT" AND "PROBABLE CAUSE" 7 (1991).
[45] LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 33.
[46] *Id.*
[47] *Id*; *see also* Barbara J. Shapiro, *"Beyond Reasonable Doubt": The Neglected Eighteenth-Century Context*, 8 LAW & HUMAN. 19 (2014) (arguing that the concepts "moral certainty" and "beyond a reasonable doubt" mean the same thing and that beyond reasonable doubt language was to be found earlier in fields other than law.).
[48] JAMES Q. WHITMAN, THE ORIGINS OF REASONABLE DOUBT: THEOLOGICAL ROOTS OF THE CRIMINAL TRIAL 2-4 (2008).

> reasonable doubt rule was one of many rules and procedures that developed in response to this disquieting possibility. *It was originally a theological doctrine, intended to reassure jurors that they could convict the defendant without risking their own salvation, so long as their doubts about guilt were not "reasonable."* . . . As we shall see, medieval and early modern Christians experienced great anxiety about the dangers that acts of judgment presented for the soul. . . . The story of the reasonable doubt rule is simply an English chapter in this long history of safer way theology, a history in which Christian theologians worried for centuries over the nature of judging, over the problems of doubt. . . .[49]

According to Whitman, "[j]urors in medieval England were simply terrified to [render a] conviction."[50] "The 'reasonable doubt' standard developed so that English jurors could, in fact, [convict defendants] in appropriate cases."[51] The standard "arose in the face of this religiously motivated *reluctance to convict*, taking its now-familiar form during the 1780s. It is still with us today, *a living fossil from an older moral world*."[52]

Finally, the third view on the origin of the BARD standard has been voiced by legal historian John Langbein, who notes that before the last quarter of the eighteenth century, when the BARD standard crystallized, many Old Bailey[53] cases in England seemed impossible to square with a high standard of proof.[54] The Sessions Papers[55] report on many eighteenth-century and

---

[49] *Id.* at 2–4 (emphasis added).

[50] *See* WHITMAN, *supra* note 48, at 3 ("[M]edieval and early modern Christians experienced great anxiety about the dangers that acts of judgment presented for the soul [of the judge]."); Shealy, *supra* note 11, at 279.

[51] *See* Shealy, *supra* note 11, at 279.

[52] *See* WHITMAN, *supra* note 48, at 4 (emphasis added).

[53] The Central Criminal Court of England and Wales is commonly referred to as the "Old Bailey.".

[54] *See* JOHN H. LANGBEIN, THE ORIGINS OF ADVERSARY CRIMINAL TRIAL 261–62 (2003) [hereinafter LANGBEIN, THE ORIGINS OF ADVERSARY CRIMINAL TRIAL]; *See also* LANGBEIN ET AL, *supra* note 30, at 696–98.

[55] "Sessions Papers are chronologically organised files of documents relating to the work of local magistrates' courts (named as "Justices of the Peace") [in England and Wales]. . . . They include several different types of documents, relating to the prosecution of criminals, the administration of poor relief and the settlement laws, and other aspects of local government. The documents, which can be in rough draft or final form, were generated primarily by the Justices' clerks, but also by Justices of Peace [judges] themselves, as well as by parish officers, litigants, accused criminals, and paupers." *See Sessions Papers: Justices' Working Documents (PS),* LONDON LIVES, https://www.londonlives.org/static/PS.jsp#:~:text=Sessions%20Papers%20are%20chronologically%20organised,the%20clerk%20of%20the%20peace. (last visited Sept. 27, 2020).

earlier criminal cases in which the prosecution was not able to prove a case against the accused, but the accused was still convicted through a partial verdict, mitigation, or recommendation of mercy.[56] English sources of the early eighteenth century, just decades before the precise notion of the BARD standard emerged, reveal that the assumption "was *not* that [the accused] was innocent until the case against him was proved beyond a reasonable doubt, but that if he were innocent he ought to be able to demonstrate it for the jury by the quality and character of his reply to the prosecutor's evidence."[57] In the late eighteenth and early nineteenth century, just as the BARD standard emerged and gained popularity, the adversarial criminal process experienced rapid development in England,[58] suggesting that the emergence of the BARD standard was related in some way to the growing presence and influence of defense lawyers in criminal trials. For example, in the Irish Treason Cases in 1798, defense lawyers were the first to urge the BARD standard on the courts, in an attempt to raise the prosecution's burden of persuasion.[59] Even though he cannot conclude just *how* the emergence of the BARD standard was related to the growing lawyerization of Old Bailey trials in the last quarter of the eighteenth century,[60] Langbein finds that *the rise of the adversary criminal trial* may indeed have affected the development of the BARD standard of proof in different ways; "for example, by disposing the judges to . . . further safeguard against the failings of the adversary procedure."[61] By "setting a high standard of proof and instructing the jury about it, [the judges] directed the attention of the jury toward any weaknesses of the prosecution's case, and away from the focus of the old altercation trial, which was on how well the defendant answered the charges and the accusing evidence."[62]

So, who is right and who is wrong here regarding the origin and early development of the BARD standard of proof? Our view is that the three schools of thought are all legitimate; they just observe the same subject from different perspectives. The origin and early development of the BARD standard were ultimately complicated and multilayered, unlikely to be fully explained by any single theory.

---

[56] *See* LANGBEIN ET AL, *supra* note 30, at 696–98.
[57] *See* J. M. BEATTIE, CRIME AND THE COURTS IN ENGLAND: 1660-1800, 341 (1986) (emphasis added).
[58] *See* LANGBEIN ET AL, *supra* note 30, at 452.
[59] *See* May, *supra* note 10, at 656–57, quoted in CHARLES MCCORMICK, LAW OF EVIDENCE §321, at 682 n.3 (1st ed. 1954); *see also* Newman, *supra* note 9, at 981–82.
[60] For instance, the Sessions Papers, Langbein's major source for research on this matter, do not indicate that the defense counsel played much role in the development of the BARD standard in 1780s England. *See* LANGBEIN, THE ORIGINS OF ADVERSARY CRIMINAL TRIAL, *supra* note 54, at 265.
[61] *See id.* (emphasis added).
[62] *See* LANGBEIN ET AL, *supra* note 30, at 698.

B. The unclear definition of the reasonable doubt standard

The mystery surrounding the BARD standard does not stop with its origins. Perhaps more surprising—if also more well known among the legal community—is that the mystery extends to the BARD standard's very definition, which remains ambiguous despite the standard's central role in the criminal justice system. People have seemingly become accustomed to such unclarity, as "many judges [are] inclined to the view that the less said by way of explanation of this [reasonable doubt] standard, the better."[63]

This unclarity was not always the case. From the late nineteenth century to the second half of the twentieth century, the common law world widely recognized a singular definition of "proof beyond a reasonable doubt."[64] This definition was delivered by Chief Justice Lemuel Shaw of the Supreme Judicial Court of Massachusetts, in *Commonwealth v. Webster*[65] in 1850, and stated:

> Then, what is reasonable doubt? It is a term often used, probably pretty well understood, but not easily defined. It is not merely possible doubt; because every thing relating to human affairs, and depending on moral evidence, is open to some possible or imaginary doubt. It is that state of the case,

---

[63] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 34.

[64] *See* 9 WIGMORE, ON EVIDENCE § 2497 (Chadbourn rev. 3d ed. 1981) ("From time to time, various efforts have been made to define more in detail this elusive and undefinable state of mind. One that has received frequent sanction and has been quoted innumerable times is that of Chief Justice Shaw of Massachusetts, on the trial of Dr. Webster for the murder of Mr. Parkman. . . ."); *see also* JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 76 (5th ed. 2009) ("Chief Justice Shaw of the Massachusetts Supreme Judicial Court crafted the traditional definition of 'beyond a reasonable doubt,' which served for more than a century as the basis for many reasonable doubt jury instructions."). *But see* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 34 ("Although language [of Justice Shaw] to this effect became standard by the late-nineteenth century, judges would often expand on and embellish Shaw's phraseology. Sometimes they would say that a belief was true beyond reasonable doubt when it was 'highly probable' or when the jurors had an 'abiding conviction' about it, or when their 'consciences were satisfied' that conviction was the right thing to do. But even in those early days of the BARD rule, many judges inclined to the view that the less said by way of explanation of this relatively new standard, the better.").

[65] Commonwealth v. Webster, 59 Mass. 295 (Sup. Ct. Mass. 1850). Coincidentally or not, it is interesting to observe that seventy years after the earliest use of the precise phrase "proof beyond a reasonable doubt" appeared in the *Boston Massacre* trials of 1770, the first widely recognized definition of this standard also emerged from Boston, Massachusetts.

> which, after the entire comparison and consideration of all the evidence, leaves the minds of jurors in that condition that they cannot say they feel an abiding conviction, to a *moral certainty*, of the truth of the charge. *The burden of proof is upon the prosecutor*. *All the presumptions of law independent of evidence are in favor of innocence*; and *every person is presumed to be innocent until he is proved guilty*. If upon such proof there is reasonable doubt remaining, the accused is entitled to the *benefit* of it by an acquittal. For it is not sufficient to establish a probability, though a strong one arising from the doctrine of chances, that the fact charged is more likely to be true than the contrary; but the evidence must establish the truth of the fact to a reasonable and *moral certainty*, a certainty that convinces and directs the understanding, and satisfies the reason and judgment, of those who are bound to act conscientiously upon it. This we take to be proof beyond reasonable doubt. . . .[66]

Justice Shaw was in concurrence with the "Received View," which equates proof beyond a reasonable doubt with moral certainty. Notably, in his explanation, Justice Shaw combined this standard of proof with a cluster of other defendant-friendly rules, namely *the prosecutorial burden of proof*, *the presumption of innocence*, and *the benefit of doubt*, all of which are cornerstones of the criminal justice system.[67] What he did here is an application of the doctrine of *noscitur a sociis*: the expression "beyond a reasonable doubt" draws meaning and clarity from the associated rules mentioned in the same instruction. Maybe that helps explain why Shaw's definition of the BARD standard was considered to be thorough and convincing for upward of a century.

This definition only lasted until 1994, when the Supreme Court of the United States negated Shaw's definition in the case of *Victor v. Nebraska*, holding that his terminology is archaic, unhelpful, and misleading.[68] The condemnation of the Court in *Victor* focused on Shaw's reference to "moral certainty," the same phrase that had shepherded the BARD standard's

[66] *Id.* at 320 (emphasis added).

[67] *Id.* at 320. *See e.g.*, Scott E. Sundby, *The Reasonable Doubt Rule and the Meaning of Innocence*, 40 HASTINGS L.J. 457, 457 (1989) ("The presumption has been called the 'golden thread' that runs throughout the criminal law, heralded as the 'cornerstone of Anglo-Saxon justice,' and identified as the 'focal point of any concept of due process.' Indeed, the presumption has become so central to the popular view of the criminal justice system, it has taken on 'some of the characteristics of superstition.'").

[68] Victor v. Nebraska, 511 U.S. 1 (1994).

evolution in the age of Enlightenment.[69] The Court opined: "[t]he risk that jurors would understand 'moral certainty' to authorize convictions based in part on *value judgments regarding the defendant's behavior* is particularly high in cases where the defendant is alleged to have committed a repugnant or brutal crime."[70] Justice Sandra Day O'Connor, writing for the Court, said emphatically that "this Court does not condone the use of the *antiquated* 'moral certainty' phrase. . . ."[71] Justice Ruth Bader Ginsburg ventured that "the phrase 'moral certainty', though not so misleading as to render the instructions [to jurors] unconstitutional, should be avoided as *unhelpful* in defining reasonable doubt. . . ."[72] Even Justice Harry Blackmun, writing in a dissent, stated that there exists "the real possibility that such language would lead jurors reasonably to believe that they could base their decision to convict upon *moral standards or emotion in addition to or instead of evidentiary standards*."[73] In a nutshell, the Court in *Victor* noted that "'moral certainty,' standing alone, might not be recognized by modern jurors as a synonym for 'proof beyond a reasonable doubt.'"[74]

The Supreme Court of the United States in *Victor* noted what is *not* "proof beyond a reasonable doubt," disconnecting BARD from "moral certainty," a phrase that once grounded the concept and linked it to its philosophical origins.[75] However, it did not elaborate on what BARD *is*.[76]

---

[69] Here, it is worth noting that the phrase "moral certainty" means quite different things in Justice Shaw's definition of BARD (the version of the Enlightenment era) from its usage in the definitional rejection made by the Court in *Victor*. *See Victor*, 511 U.S. at 13–15 ("We are somewhat more concerned with Sandoval's argument that the phrase 'moral certainty' has lost its historical meaning, and that a modern jury would understand it to allow conviction on proof that does not meet the beyond a reasonable doubt standard. Words and phrases can change meaning over time: A passage generally understood in 1850 may be incomprehensible or confusing to a modern juror. . . . [M]oral certainty would be understood by modern jurors to mean a standard of proof lower than beyond a reasonable doubt."); *see also* Shealy, *supra* note 11, at 245.

[70] *Victor*, 511 U.S. at 24 (emphasis added).

[71] *Id.* at 4 (emphasis added).

[72] *Id.* (emphasis added).

[73] *Id*. at 38 (emphasis added).

[74] *Id.* at 14.

[75] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 35.

[76] But note: Justice Ginsburg did offer a definition of the BARD standard in her concurrence in *Victor*. She advocated the proposed definition as set forth by the Federal Judicial Center. *See Victor*, 511 U.S. at 26–27. The proposed definition of reasonable doubt by the Federal Judicial Center is as follows:

> [T]he government has the burden of proving the defendant guilty beyond a reasonable doubt. Some of you may have served as jurors in civil cases, where you were told that it is only necessary to prove that a fact is more likely true than not true. In criminal cases, the

Maybe that was because the Supreme Court was not yet ready to launch such a definition, or because it believed that the concept is quite self-explanatory. Or maybe the Court deliberately avoided clarifying the definition of BARD in order to keep the standard flexible for dealing with situations of varying complexity and to defer some discretionary power of interpretation to the jurors in individual cases. No matter what the reason was, the inaction of the Supreme Court of the United States in defining BARD has caused confusion among lower courts and beyond in interpreting the standard of proof in criminal cases, as shown in various, inconsistent explications in jury instructions.[77]

Among the various efforts to interpret BARD, one *unofficial* interpretation has gained ground for being straightforward, easy to understand, heuristically sensible, and popular among the general public. That interpretation involves seeing "proof beyond a reasonable doubt" as a mathematically high probability, a number toward the high end on a scale that ranges from 0 to 1 (i.e., from 0 percent to 100 percent).[78] One obvious inspiration for this approach is found in the commonly used interpretation of the civil law standard of proof: preponderance of evidence. That is normally taken to mean that the trier of fact should favor the plaintiff over the defendant in a civil case if the case of the plaintiff is, more likely than not. That is, if it

---

> government's proof must be more powerful than that. It must be beyond a reasonable doubt. . . . Proof beyond a reasonable doubt is proof that leaves you *firmly convinced* of the defendant's guilt. There are very few things in this world that we know with absolute certainty, and in criminal cases the law does not require proof that overcomes every possible doubt. If, based on your consideration of the evidence, you are firmly convinced that the defendant is guilty of the crime charged, you must find him guilty. If on the other hand, you think there is a real possibility that he is not guilty, you must give him the benefit of the doubt and find him not guilty."

FED. JUDICIAL CTR., PATTERN CRIMINAL JURY INSTRUCTIONS 28 (1987) (emphasis added). Unfortunately, Justice Ginsburg was unable to convince any other justices to join her. The recommendations of the Federal Judicial Center remain just that, recommendations. *See* Shealy, *supra* note 11, at 251–53.

[77] *See, e.g.*, Shealy, *supra* note 11, at 254–68 ("[*Victor*] did not bring closure or clarity to the lower courts; difficulties around. One finds a polyglot of definitions, terms, and rationales. Federal and state courts tolerate a wide variety of instructions on 'reasonable doubt.' Even those courts which prefer not to define 'reasonable doubt' will, on occasion, tolerate substantial differences among trial judges who attempt to do so."). For example, the Fourth and Seventh Circuits upheld their district courts' refusals to define "reasonable doubt" even upon request by a defendant. In contrast, the Sixth and Eighth Circuits have model jury instructions defining reasonable doubt.

[78] For one of the earliest articles discussing quantification of standards of proof in juridical fact-finding, *see* John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065 (1968).

has a probability greater than 0.5 (i.e., 50 percent).[79] The criminal standard, of course, is meant to be much higher than the civil standard, so the relevant probabilities would have to be much higher for conviction (0.9 or 0.95 are commonly cited unofficial estimates).[80]

We agree that this probability-based approach makes intuitive sense. As the BARD standard recognizes, we cannot be completely certain (a probability of 1.0) about anything to do with human affairs or the guilt of an accused party. This impossibility invites the suggestion that the standard of proof for conviction be defined as a threshold probability, a numerical degree of confidence in the party's guilt on the part of triers of fact.[81] However, in practice, most trial judges would hesitate (if not immediately refuse) to adopt this numerical approach in jury instructions due to an inherent problem: any specification of a degree of belief necessary for a finding of guilt involves an *explicit* admission that wrongful convictions will inevitably occur.[82] For instance, a confidence of 95 percent in the guilt of the accused would generally suggest that as many as one in every twenty innocent defendants will be wrongly convicted. As Larry Laudan neatly describes in his book *Truth, Error, and Criminal Law: An Essay in Legal Epistemology*:

> While acknowledging in the abstract that no method of proof is infallible and thus admitting in principle that mistakes will occur from time to time, the judiciary has an entrenched resistance to any explicit admission that the system has a built-in tolerance for wrongful convictions. . . . What should be clear is that identifying BARD with any level of probability less than unity would explicitly acknowledge that the system officially condones a certain fraction of

---

[79] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 44. *But see* Edward K. Cheng & Michael S. Pardo, *Accuracy, Optimality, and the Preponderance Standard*, 14 LAW, PROBABILITY & RISK 193 (2015) (arguing the traditional probability understanding of the civil standard of proof, above 0.5, is a mistake. One hypothesis can be more probable than a competing hypothesis—that is, supported by the preponderance of the available evidence—even if the former hypothesis falls short of the 0.5 threshold.).

[80] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 44.

[81] *Id.*

[82] *See generally* Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530 (1989); Daniel Shaviro, *A Response to Professor Callen*, 65 TUL. L. REV. 499 (1991) (Professor Shaviro made the point that in the past legal professionals had been "squeamish" about frankly acknowledging the fact that they implicitly tolerate some wrongful verdicts.). *But see* Speiser v. Randall, 357 U.S. 513, 525–26 (1958) ("There is always in litigation a margin of error, representing error in factfinding, which both parties must take into account. Where one party has at stake an interest of transcending value – as a criminal defendant his liberty—this margin of error is reduced as to him by the process of placing on the other party the burden . . . of persuading the factfinder at the conclusion of the trial of his guilt beyond a reasonable doubt.").

> wrongful convictions. That, in turn, would supposedly threaten the ordinary person's faith in the justice system. Better, it seems, to avoid any talk of probability as the standard for conviction than to acknowledge publicly that the system expressly permits incorrect judgments of guilt.[83]

In addition, even if the worries about the broader social message being sent by the use of probabilities could be quelled, people disagree about what number on the scale of probabilities should or could be assigned to BARD.[84] Therefore, even if BARD, as a numerical high probability, makes intuitive sense, it is controversial and indeterminable in practice.

Another problem here is that the law applies burdens of persuasion like the BARD standard to individual elements, not to causes of action as a whole.[85] So, even if people agreed to narrow down the BARD standard to a range of high probabilities (e.g., 0.9–0.95), the application of burdens of persuasion would yield the *conjunction paradox of proof*.[86] If, for example, four independent elements are each proved to a 0.9 probability, the probability of them all being true is only about 0.66, which does not sound like proof beyond a reasonable doubt.

The trouble of considering BARD as a mathematically high probability goes even further. As Brian Leiter and Ronald J. Allen stated in their article *Naturalized Epistemology and the Law of Evidence*:

> The expected utility theorist may respond by criticizing the *law* and arguing that it is the conjunction of elements that should be found to a specific level. This, too, yields unacceptable consequences, by making the level of proof of

---

[83] LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 45–46.

[84] *See* United States v. Fatico, 458 F. Supp. 388, 409–10 (E.D.N.Y. 1978), aff'd, 603 F.2d 1053 (2d Cir. 1970) (Judge Jack B. Weinstein described a survey he did on his colleagues' different understandings of the probability equivalents of the various standards of proof.). *Also see* Kaplan, *supra* note 80, at 1073–74 ("Probably the most important reason why we do not attempt to express reasonable doubt in terms of quantitative odds, however, is that in any rational system the utilities (or disutilities) that determine the necessary probability of guilt will vary with the crime for which the defendant is being tried, and indeed with the particular defendant. In a criminal trial, as in any decision process, we must consider the utilities associated with differing decisions of the particular case at issue-not ties over many disparate types of criminal cases.").

[85] *See In re* Winship, 397 U.S. 358, 364 (1970) (". . . the Due Process Clause protects the accused against conviction except upon proof beyond a reasonable doubt of *every fact necessary to constitute the crime* with which he is charged.") (emphasis added).

[86] *See,* Michael S. Pardo, *The Paradoxes of Legal Proof: A Critical Guide*, 99 B.U. L. REV. 233, 266–81 (2019); Ronald J. Allen, *Artificial Intelligence and the Evidentiary Process: The Challenges of Formalism and Computation*, 9 A.I. & L. 99, 109 (2001).

> specific elements turn on the fortuity of the *number* of elements in a cause of action. Take the example of theft and murder. Theft has considerably more elements than murder. To convict for theft requires on average that intent to steal be established to a higher probability than intent to kill for a murder conviction. This strikes all legal observers as both unacceptable and absurd. . . .[87]

In essence, the idea that proof beyond a reasonable doubt might be seen as a mathematically high probability may have great heuristic value, but it likely can never become an official explanation or definition of the BARD standard. At the end of the day, people have mixed feelings about this standard. On the one hand, this standard has been almost uniformly accepted as the standard of proof for criminal cases in numerous legal systems for a long period of time; on the other hand, as Erik Lillquist has said, "what the standard really requires and why we use it at all both remain unclear."[88] Researchers still disagree on its origin, and the precise meaning of BARD remains elusive. Like Leonardo da Vinci's famous portrait *Mona Lisa*, no matter how many times we view this vivid painting, we just cannot get enough of her ambiguous smile and we strive ceaselessly to understand it. Based on this context, to better understand the reasonable doubt standard, we decided to take a closer look at its highs and lows.

II. GLORIES OF THE REASONABLE DOUBT STANDARD

Proof beyond a reasonable doubt has been accepted in many different criminal justice systems as the proper formula to set out the standard of proof and deemed as one of the most fundamental requirements in law. But why is BARD treated as the golden standard? We identified two holy auras surrounding it. First, in the United States, proof of a criminal charge beyond a reasonable doubt is *constitutionally* required, since the accused in a criminal prosecution has at stake interests of immense importance (e.g., his or her good name, freedom, or even life), which triggers protection under the Due Process Clause of the Constitution. And second, the BARD standard is widely considered to reflect the *Blackstone principle* (a foundational maxim in common law countries and beyond), which states that convicting an innocent

[87] Brian Leiter & Ronald J. Allen, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1504–05 (2001) (emphasis added).
[88] *See* Erik Lillquist, *Recasting Reasonable Doubt: Decision Theory and the Virtue of Variability*, 36 UC DAVIS L. REV. 85, 87 (2002).

person is much worse than acquitting a guilty person.[89] We will take a closer look at each of these glories of BARD in turn.

A. *The reasonable doubt standard as constitutionally required in criminal cases*

In the United States, the BARD standard has been widely used in both state and federal courts by custom and common law precedent since the mid-nineteenth century,[90] and the Supreme Court discussed reasonable doubt in a series of cases in the late nineteenth and early twentieth centuries.[91] For example, in *Miles v. United States* (1880), the Court found no error in the trial court's jury instruction of this case: "The prisoner's guilt must be established beyond reasonable doubt."[92] In *Coffin v. United States* (1895), the Court held that:

> [I]t is the duty of the judge, in all jurisdictions, when requested, and in some when not requested, to explain [the presumption of innocence] to the jury in his charge. The usual formula in which this doctrine is expressed is that every man is presumed to be innocent until his guilt is proved beyond a reasonable doubt.[93]

Nonetheless, it was not until *In re Winship* in 1970 that the Supreme Court formally *constitutionalized* the BARD standard.[94]

In *Winship*, for the first time, the Court explicitly held that "the *Due Process Clause* protects the accused against conviction except upon proof beyond a reasonable doubt of every fact necessary to constitute the crime with which he is charged."[95] To support this holding, the majority opinion of the

---

[89] Daniel Epps, *One Last Word on the Blackstone Principle*. 102 VA. L. REV. 34, 35 (2016).

[90] *See,* Commonwealth v. Webster, 59 Mass. 295, 320 (1850); Commonwealth v. Costley, 118 Mass. 1, 16 (1875).

[91] *See,* Miles v. United States, 103 U.S. 304, 312 (1880); Coffin v. United States, 156 U.S. 432, 452–53 (1895); Davis v. United States, 160 U.S. 469, 488 (1895); Holt v. United States, 218 US. 245, 253 (1910); Wilson v. United States, 232 U.S. 563, 569–70 (1914).

[92] Miles v. United States, 103 U.S. 304, 309 (1880).

[93] Coffin v. United States, 156 U.S. 432, 495 (1895).

[94] *In re* Winship, 397 U.S. 358 (1970).

[95] *Id.* at 364 (emphasis added). Note: The Due Process Clause provides that no person shall be "deprived of life, liberty, or property, without due process of law." The Fifth Amendment applies this limitation to the Federal Government, and the Fourteenth Amendment imposes the same restriction on the States. Here in *Winship*, because the issue presented before the Court was whether proof beyond a reasonable doubt is required during the adjudicatory stage in the New York Family Court where a juvenile was charged with

Winship Court, delivered by Justice William Brennan, illustrated two points. First, based on a long history and virtually unanimous adherence to the reasonable-doubt standard in common law jurisdictions, "expressions in many [earlier] opinions of this Court indicated that it has long been *assumed* that proof of a criminal charge beyond a reasonable doubt is *constitutionally* required."[96] Second, the majority opinion in *Winship* stated:

> The accused during a criminal prosecution has at stake interests of immense importance, both because of the possibility that he may lose his liberty upon conviction and because of the certainty that he would be stigmatized by the conviction. Accordingly, a society that values the good name and freedom of every individual should not condemn a man for commission of a crime when there is reasonable doubt about this guilt. . . . "*Due process* commands that no man shall lose his liberty unless the Government has borne the burden of . . . convincing the factfinder of his guilt." To this end, the reasonable-doubt standard is indispensable, for it "impresses on the trier of fact the necessity of reaching a subjective state of certitude of the facts in issue."[97]

---

an act which would constitute a crime if committed by an adult, it was more about the Due Process Clause of the Fourteenth Amendment. Nonetheless, the Court of *Winship*, when holding BARD standard of proof in criminal trials is required by the Due Process Clause ("DP"), did not further discuss about DP of Fifth and Fourteenth Amendments separately.

[96] *Id.* at 361–63 (emphasis added). Here, the majority opinion of the *Winship* Court listed a number of supporting cases previously adjudicated in the Supreme Court of the United States and quoted lines in opinions of three particular cases of the Court to demonstrate that BARD standard in criminal trials has a constitutional root: Leland v. Oregon, 343 U.S. 790, 802–03 (1952) (dissenting opinion) ("It is the duty of the Government to establish . . . guilt beyond a reasonable doubt. This notion – basic in our law and rightly one of the boasts of a free society – is a requirement and a safeguard of *due process* of law in the historic, procedural content of '*due process*'."); Brinegar v. United States, 338 U.S. 160, 174 (1949) ("Guilt in a criminal case must be proved beyond a reasonable doubt and by evidence confined to that which long experience in the common law tradition, to some extent embodied in the *Constitution*, has crystallized into rules of evidence consistent with that standard. These rules are historically grounded rights of our system, developed to safeguard men from dubious and unjust convictions, with resulting forfeitures of life, liberty and property."); and Davis v. United States, 160 U.S. 469 (1895) ("[The requirement is implicit in] *constitutions* . . . [which] recognize the fundamental principles that are deemed essential for the protection of life and liberty.").

[97] *In re* Winship, 397 U.S. at 363-64 (quoting Speiser v. Randall, 357 U.S. 513, 525–26 (1958) (emphasis added)) (quoting Norman Dorsen and Daniel Rezneck, *In Re Gault and the Future of Juvenile Law*, 1 FAM. L. Q. 1, 26 (1967)). *But see* Lawrence Rosenthal, *Second Thoughts on Damages for Wrongful Convictions*, 85 CHI.-KENT L. REV. 127, 160 (2010) ("[D]ue process has always concerned itself with protection of the innocent, *but it accommodates society's interest in maintaining reasonably practicable means for*

This historical interpretation of the BARD standard together with the argument about its link to the Due Process Clause was *creative*: the Constitution of the United States neither explicitly addresses the standard of proof in criminal trials nor mentions reasonable doubt,[98] as Justice Hugo Black pointed out in his dissent in *Winship*.[99] Still, the Court opinion in *Winship* went further in defending BARD, declaring it to be not just constitutionally but also ethically warranted, a prime instrument for reducing adjudicative errors (wrongful convictions). In his concurrence, Justice John Harlan put *Winship* at the very heart of the modern Court's "reasonable doubt" jurisprudence:

> [T]he choice of the standard for a particular variety of adjudication does, I think, reflect a very fundamental assessment of the comparative social costs of erroneous factual determinations.[100] . . . I view the requirement of proof beyond a reasonable doubt in a criminal case as bottomed on a fundamental value determination of our society that *it is far worse to convict an innocent man than to let a guilty man go free*.[101]

Justice Harlan's concurrence in *Winship* embedded a significant amount of information regarding adjudicative errors and relied on a social creed that has ancient roots as well as famous interpretations in history: the Blackstone principle. To understand this principle, we need to first understand the concept of error distribution.

---

punishing the guilty as well. It is a small wonder that the Court has been reluctant to use the Due Process Clause to impose novel procedural reforms even when they have the potential to reduce the risk of erroneous convictions.") (emphasis added).

[98] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 34.

[99] *In re* Winship, 397 U.S. at 358 ("[Contrary to the majority opinion of this case], the Court has never clearly held, however, that proof beyond a reasonable doubt is either expressly or impliedly commanded by any provision of the Constitution. The Bill of Rights, which in my view is made fully applicable to the States by the Fourteenth Amendment . . . does by *express language* provide for, among other things, a right to counsel in criminal trials, a right to indictment, and the right of a defendant to be informed of the nature of the charges against him. And in two places the Constitution provides for trial by jury, but *nowhere in that document is there any statement that conviction of crime requires proof of guilt beyond a reasonable doubt*.") (emphasis added).

[100] *Id.* at 370.

[101] *Id.* at 372 (emphasis added).

B. *The reasonable doubt standard as an essential tool to restrain wrongful convictions*

From an epistemic point of view, there are four, and only four, possibilities with respect to the result of a criminal case adjudicated: correct conviction (a wrongdoer being found guilty), correct acquittal (an innocent person being found not guilty), wrongful conviction (an innocent person being found guilty), and wrongful acquittal (a wrongdoer being found not guilty, or escaping from punishment through ways other than a formal trial).[102] A core aim (and, some scholars[103] argue, the primary purpose) of any criminal justice system is the "[r]ectitude of decision,"[104] meaning the correct application of substantive law to the true facts, thus making truthful verdicts (epistemically correct convictions and correct acquittals) and avoiding false verdicts (wrongful convictions and wrongful acquittals). In this article, we phrase this aim as the "first goal of criminal justice"—*error reduction*.[105] Here, the term *error* is used in a purely epistemic sense, unlike the procedural errors often mentioned by appellate courts.[106]

And yet, triers of fact will sometimes, despite their best efforts, be wrong in their factual conclusions of the case at issue, namely finding an innocent person guilty or a wrongdoer not guilty, thus making wrongful convictions and wrongful acquittals, respectively.[107] In that case, the question

---

[102] Note: here, we use the phrase "wrongful acquittal" to include events that are not an acquittal per se (a formal trial verdict of not guilty) but that represent other kinds of failure by the justice system to impose punishment on a wrongdoer whom authorities initially and correctly identified as guilty, such as a prosecutor's decision not to press charges after an arrest or a judge's order dismissing an indictment.

[103] *See, e.g.*, LAUDAN, THE LAW'S FLAWS, *supra* note 22; Ronald J. Allen & Larry Laudan, *Why Do We Convict As Many Innocent People as We Do?: Deadly Dilemmas*, 41 TEX. TECH L. REV. 65, 65 (2008).

[104] *See* WILLIAM TWINING, THEORIES OF EVIDENCE: BENTHAM AND WIGMORE 14–16 (1985).

[105] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 1.

[106] For a higher or appellate court, an "error" occurs in a trial when some rule of evidence or procedure has been violated, misinterpreted, or misapplied. A trial can avoid such "errors" by scrupulously following the letter of the current rules governing the admissibility of evidence and procedures. This process, however, is no guarantee of an epistemically correct outcome (with the truly guilty and the truly innocent being correctly identified). For the purpose of our discussion, an error occurs when a materially (i.e., truly) innocent person is deemed guilty (named as "false positive") or when a materially guilty person fails to be found guilty (named as "false negative"). By the same token, for the purpose of our discussion, the phrase "correct verdict" means an epistemically or materially correct verdict, not a procedurally correct one, even though these two concepts are overlapping (a procedurally correct verdict may very likely also be an epistemically correct one).

[107] *In re* Winship, 397 U.S. 358, 370 (1970).

arises of which type of error—a wrongful conviction (false positive) or a wrongful acquittal (false negative)—is more serious, and is a consideration that involves an assessment of the comparative social disutility of each outcome and a determination of which is better avoided.[108] In this article, we phrase this question as the "second goal of criminal justice"—*error distribution*.

As previously stated, the social norm on this question is the view that "it is far worse to convict an innocent man than to let a guilty man go free."[109] Or, according the English jurist William Blackstone in his 1760s *Commentaries on the Laws of England*—"[i]t is better that ten guilty persons escape than that one innocent suffer."[110] This adage is perhaps the most revered in the English criminal law, "exalted by [succeeding generations] of judges and scholars alike[111] as 'a cardinal principle of Anglo-American jurisprudence.'"[112] Of course, no one insists that the criminal justice system

---

[108] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 1.

[109] *See In re* Winship, 397 U.S. at 372.

[110] 4 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAW OF ENGLAND 352 (1769).

[111] *See, e.g.*, Laura I Appleman, *A Tragedy of Errors: Blackstone, Procedural Asymmetry, and Criminal Justice*, 128 HARV. L. REV. F. 91, 91 (2015) (describing the Blackstone ratio as the "Mount Everest of legal mantras").

[112] *See*, Epps, *supra* note 7, at 1067–68 (citing United States v. Greer, 538 F.2d 437, 441 (D.C. Cir. 1976). *But see* Lawrence Rosenthal, *Eyewitness Identification and the Problematics of Blackstonian Reform of the Criminal Law*, 110 J. CRIM. L. & CRIMINOLOGY 181, 220 (2020) ("Blackstone's ratio is not uncontroversial; it is far from clear that it accurately captures the costs and benefits of error in the criminal justice system."). Blackstone was not the first to invent a maxim on error distribution. Similar maxims were "in the air" of English legal thought long before Blackstone published his *Commentaries*. Epps, *supra* note 7, at 1080. As Alexander Volokh has amusingly documented in his article *n Guilty Men*, many legal and religious thinkers over the centuries have endorsed similar ratios. *See* Alexander Volokh, *n Guilty Men*, 146 U. PA. L. REV. 173, 175–76 (1997). For example, in 1470, Sir John Fortescue, who served as Chief Justice of the King's Bench during the reign of Henry VI, published *De Laudibus Legum Angliae* (Commendation of the Laws of England), one paragraph of which stated: "Indeed, one would much rather that *twenty guilty persons should escape the punishment of death, than that one innocent person should be condemned*, and suffer capitally." JOHN FORTESCUE, COMMENDATION OF THE LAWS OF ENGLAND 45 (Francis Grigor ed. & trans., Sweet & Maxwell 1917) (c. 1543). As another example, in a work written in the 1670s but not published until 1736, English jurist Sir Matthew Hale provided: "it is better *five guilty persons should escape unpunished, than one innocent person should die*." 2 MATTHEW HALE, THE HISTORY OF THE PLEAS OF THE CROWN 289 (George Wilson ed., 1778) (1736).

The popularity of Blackstone's version of the maxim and the frequency with which it is cited are likely due to Blackstone's enormous influence on the Anglo-American legal profession more generally. According to Wilfrid Prest, demand for reprinted, abridged, and translated versions of Blackstone's written work, the *Commentaries on the Laws of England*, was "almost inexhaustible" in the eighteenth and nineteenth centuries. Some commentator described him as "the core element in the British Enlightenment," comparing him to Montesquieu, Beccaria, and Voltaire. *See* WILFRID PREST, WILLIAM BLACKSTONE:

should produce *exactly* ten false acquittals for every false conviction.[113] But the maxim does convey a *value preference in error distribution*: wrongful conviction of the innocent is a much greater wrong than failed conviction of the guilty. Therefore, "in distributing criminal punishment, we must strongly err in favor of false negatives . . . in order to minimize false positives."[114]

Although Blackstone is cited most often nowadays, the directive to avoid the erroneous punishment of the innocent has much older roots, "at least as far back as the *Old Testament*," and is beyond the contour of the common law system.[115] For example, in Genesis, Abraham pleads with God to spare Sodom in order to avoid "slay[ing] the righteous with the wicked . . . ."[116] God agrees that he will not destroy the city if as few as ten righteous persons there are found.[117] Similarly, in Exodus, God commands, "the innocent and righteous slay thou not. . . ."[118] As another example, medieval Continental judges developed the rule "*in dubio pro reo*" that means, "in doubt you must decide for the defendant,"[119] which was linked to the notion that judges

---

LAW AND LETTERS IN THE EIGHTEENTH CENTURY 307–08 (2008). Across the Atlantic, within the American academia, practice, as well as within the judiciary, the *Commentaries* had a substantial impact as well. According to Albert Miles, with the scarcity of law books on the frontier, they were "both the only law school and the only law library most American lawyers used to practice law in America for nearly a century after they were published [in the 1760s]." *See* Albert S. Miles, *Blackstone and his American Legacy*, 5 AUSTL. & N.Z.J. L. & EDUC. 46, 56 (2000). American subscribers to the first edition of *Commentaries*, and later readers who were profoundly influenced by it, include James Iredell, John Marshall, James Wilson, John Jay, John Adams, James Kent, and Abraham Lincoln, among others. *See generally,* Albert Alschuler, *Sir William Blackstone and the shaping of American Law*, 144 NEW L.J. 896, 897 (1994).

[113] Note:

> If the ratio were treated as the *sole* benchmark for the success of the criminal justice system, it would permit absurd results. For example, a system that, out of 100 trials, convicted 9 innocent people and acquitted 91 guilty people (and thus rendered no accurate adjudications at all) would technically comply with the principle.

Epps, *supra* note 7, at 1068 n.6 (citing Allen & Laudan, *supra* note 103, at 75–77).

[114] *See* Epps, *supra* note 7, at 1067-68 (In his article, Epps refers to the general rule of value preference as the "Blackstone principle," carefully separating this principle from the "Blackstone ratio" of 10:1. Epps argues that it is the former and not the latter that "accords with most people's deeply felt intuitions about criminal justice.").

[115] *See generally*, Alexander Volokh, *n Guilty Men*, 146 U. PA. L. REV. 173, 173–74, 177–78 (1997); Epps, *supra* note 7, at 1077–81. This paragraph owes a significant debt to Volokh and Epps whose articles pointed the way to numerous sources.

[116] *Genesis* 18:25.

[117] *Id.* at 18:32.

[118] *Exodus* 23:7.

[119] *See* WHITMAN, *supra* note 48, at 122–23.

needed "proof 'clearer than the midday sun' before sending a person to blood punishment."[120]

The most common and straightforward argument for the Blackstone principle is that "the disutility of convicting an innocent person far exceeds the disutility of finding a guilty person to be not guilty. . . ."[121] Justice Harlan, concurring in *Winship*, stressed this point:

> Because the standard of proof affects the comparative frequency of [false positives and false negatives], the choice of the standard to be applied in a particular kind of litigation should, in a rational world, reflect an assessment of the comparative social disutility of each. . . . [F]or example, we view it as no more serious in general for there to be an erroneous verdict in the defendant's favor than for there to be an erroneous verdict in the plaintiff's favor [in a civil case]. . . . In a criminal case, on the other hand, we do not view the social disutility of convicting an innocent man as equivalent to the disutility of acquitting someone who is guilty."[122]

Justice Harlan's argument emphasizes the severity of harm from wrongful convictions in individual cases: a wrongfully convicted defendant can lose his liberty or even his life, and also faces the stigma of being officially branded as a wrongdoer.[123] More specifically, Keith Findley has explained why wrongful convictions at all levels, from petty misdemeanors to capital murder, are costly and profoundly damaging:

> [R]egardless of the sentence imposed, *every* prosecution and every conviction is a devastating experience. The stress of accusation alone is overwhelming. The expense of defense can be enormous. The loss of one's good name, of friendships and family relationships, of employment, of savings, of the ability to find future employment and

[120] *Id.* at 123.

[121] *See* Lawrence B. Solum, *Presumptions and Transcendentalism: You Prove It! Why Should I?*, 17 HARV. J. L. & PUB. POL'Y 691, 701 (1994).

[122] *In re* Winship, 397 U.S. 358, 371–72 (1970) (Harlan, J., concurring). It is important to note that typical civil remedies (damages or injunctions) are less severe than criminal sanctions, meaning that false positives in the civil arena are less costly.

[123] *See, e.g.*, Michael D. Pepson & John N. Sharifi, *Lego v. Twomey: The Improbable Relationship Between an Obscure Supreme Court Decision and Wrongful Convictions*, 47 AM. CRIM. L. REV. 1185, 1195 (2010) ("The rationale undergirding the requirement of an enhanced standard of proof in criminal cases is bottomed on the nature of what is at stake (i.e., the accused's liberty or, at times, life) . . . .").

> housing, and the corrosive effects of being marginalized and disbelieved by one's own government—all are regular features of false convictions, *regardless of the sentence imposed*. And then consider the enormous losses occasioned by imprisonment. Beyond those direct punishments, the collateral consequences of convictions tally literally in the thousands. . . . A criminal conviction, especially a felony conviction, marks a person for life, making it enormously difficult to obtain employment, housing, and education.[124]

Other arguments have been made from an ideological perspective that the consequences of convicting an innocent person are more *morally* serious than that of acquitting a guilty person. For example, the nineteenth-century reformist Sir Samuel Romilly observed, "[w]hen, therefore, the guilty escape, the Law has merely failed. . . . But when the innocent become the victims of the Law . . . it creates the very evil it was to cure, and destroys the security it was made to preserve . . . ."[125] More recently, Ronald J. Allen and Larry Laudan suggested, "executing a truly innocent person is . . . a gross violation of the social contract that promises the government will protect the life, liberty, and pursuit of happiness of its citizens."[126] Also, Michael Risinger has pointed out:

> Viewing the *state* as having more responsibility for harm done directly to the immediate subjects of its acts than for the harm done indirectly by its failures to act, or by its choices to act one way rather than another, has a long tradition, especially in situations where the latter harm is done by the subsequent choice of an independent human agent.[127]

Standard of proof has been conceived of as one of the most appropriate tools for distributing errors in criminal cases.[128] As Michael L. DeKay has put it, "[h]igher standards of proof lead to more erroneous

---

[124] Keith A. Findley, *Reducing Error in the Criminal Justice System*, 48 SETON HALL L. REV. 1265, 1293 (2018) (second emphasis added) (internal citations omitted).
[125] *See* SIR SAMUEL ROMILLY, *Observations on the Criminal Law as It Relates to Capital Punishments, and on the Mode in Which It Is Administered*, in 1 THE SPEECHES OF SIR SAMUEL ROMILLY IN THE HOUSE OF COMMONS 106, 165–66 (London, James Ridgway and Sons, 1820).
[126] Allen & Laudan, *supra* note 103, at 66.
[127] D. Michael Risinger, *Tragic Consequences of Deadly Dilemmas: A Response to Allen and Laudan*, 40 SETON HALL L. REV. 991, 1020 (2010) (emphasis added).
[128] *See, e.g.*, LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 66–69.

acquittals and fewer erroneous convictions, all else being equal."[129] At the heuristic level, the diagram in Figure 1 demonstrates this important point.[130] The horizontal axis in Figure 1 is a scale representing the likelihood of guilt that the trier of fact (jurors in a jury trial or a judge in a bench trial) eventually assigns to the case at trial, which ranges from 0 (the leftmost point), meaning no likelihood of guilt at all, to 1 (the rightmost point), meaning certainty of guilt. The vertical axis represents the (total) number of cases that are assigned with a particular likelihood of guilt by the trier of fact, starting from zero. Curve I is a representation of all the cases in which a criminal defendant involved is truly innocent, meaning if the trier of fact knew all the facts to certainty, the defendant would be found "not guilty." Curve II is a representation of all the cases in which a criminal defendant is, conversely, truly guilty.[131]

Now, suppose that we draw a vertical line in this diagram to represent the standard of proof in criminal cases (the "SoP" line). The curves to the left side of the SoP line represent criminal cases in which a defendant is *acquitted*, whether the defendant is truly innocent or truly guilty, since the likelihood of guilt eventually assigned by the trier of fact in these cases fails to reach the SoP line. By the same token, the curves to the right side of the SoP line represent criminal cases in which a defendant is *convicted*, whether the defendant is truly innocent or truly guilty, since now the likelihood of guilt assigned by the trier of fact exceeds the SoP line.

---

[129] Michael L. DeKay, *The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 L. & SOCLAW & SO. INQUIRY 95, 97 (1996).

[130] The following portion of this section, including explanations of Figures 1, 2 and 3, owes a significant debt to Ronald J. Allen and Larry Laudan, whose materials educated both authors of this article on this matter. *See generally* Allen, *supra* note 6; LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17.

[131] The precise, real shape of Curve I and Curve II is not important for our purpose. In reality, with regard to either curve, its peak might be higher, its tails shorter, its dispersion less, or its position along the horizontal axis different from the representation. Figures 1-3 are for demonstrative purposes only.

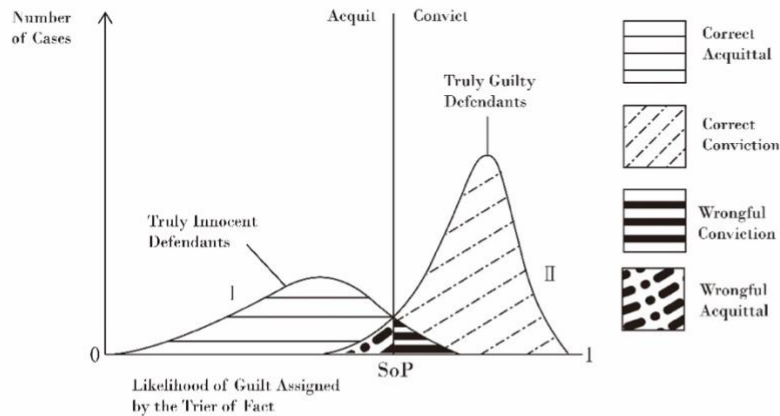

Figure 1

Figure 1 thus captures both correct and erroneous fact-findings. With regard to Curve I, the distribution of cases that criminal defendants deserve to win, the right tail represents cases in which a criminal defendant is truly innocent, but the likelihood of guilt eventually assigned by the trier of fact after hearing all the evidence is unfortunately high. Thus, the fraction of Curve I on the right side of the SoP line (the area of dark solid lines) represents *wrongful convictions*. In contrast, the fraction of Curve I on the left side of the SoP line (the area of light solid lines) represents correct acquittals since the likelihood of guilt assigned to those cases is below the SoP line.

Conversely, with regard to Curve II, the distribution of cases for truly guilty defendants, the left tail represents cases in which the likelihood of guilt eventually assigned by the trier of fact after hearing all evidence is unfortunately low. Thus, the fraction of Curve II on the left side of the SoP line (the area of dark dashed lines) represents *wrongful acquittals*, while the fraction of Curve II on the right side of the SoP line (the area of light dashed lines) is correct convictions.

Overall, the areas hashed with light lines in the diagram represent correct case fact-findings, including both correct acquittals and correct convictions, and areas hashed with dark lines represent wrongful case fact-findings, including both wrongful acquittals and wrongful convictions. According to the first goal of the criminal justice system, *error reduction*, the light-lined areas in the diagram, correct acquittals, should be much larger than the dark-lined areas, wrongful convictions, and the larger the better.[132] With

[132] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 1.

respect to the second goal of the criminal justice system, *error distribution*, in order to follow the Blackstone principle, the area of dark dashed lines, wrongful acquittals, should be much larger than the area of dark solid lines, wrongful convictions.[133]

Now, Figure 1 shows a nearly equal size for the two dark-lined areas, which means an approximately equal number of wrongful convictions and wrongful acquittals, obviously not in accordance with the Blackstone principle.[134] If we want to reduce the frequency of wrongful convictions in the diagram, we can slide the SoP line to the right, that is, toward a higher standard of proof. All other factors being equal, the farther to the right end of the diagram that we move the standard of proof, the less likely we will convict someone who is truly innocent. By doing so, we will also acquit more criminal defendants who are truly guilty.

When we move the SoP line to the right, to the standard of "beyond a reasonable doubt" (see Figure 2), the Blackstone principle of heavily preferring wrongful acquittals to wrongful convictions is satisfied.[135] Conversely, all other factors being equal, if we move the SoP line to the left and a lower standard of proof (see Figure 3), we would convict more guilty persons but also more innocent persons, which is contrary to the Blackstone principle.[136]

---

[133] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 1; BLACKSTONE, *supra* note 2, at 352.
[134] *See* BLACKSTONE, *supra* note 2, at 352.
[135] *See id.*; Coffin, 156 U.S. at 453 (explaining the beyond a reasonable doubt standard).
[136] *See* BLACKSTONE, *supra* note 2, at 352.

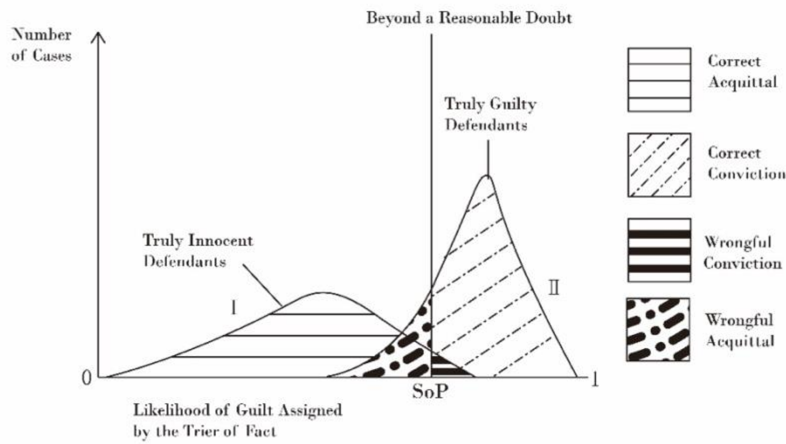

Figure 2

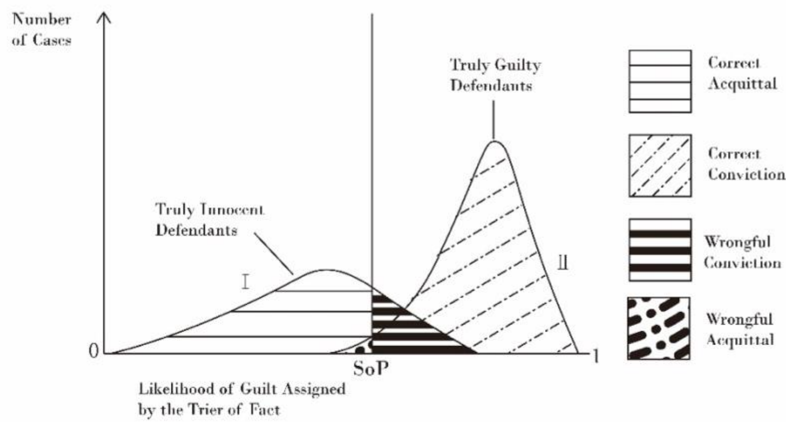

Figure 3

Therefore, as the three diagrams show, theoretically, a standard of proof could be conceived of as a mechanism for distributing errors. And, in theory, the BARD standard of proof is an essential tool in restraining wrongful

convictions in criminal cases and achieving the aspiration of the Blackstone principle.[137]

III. DARKNESS OF THE BEYOND A REASONABLE DOUBT STANDARD

Even though the BARD standard has holy auras surrounding it, that does not mean it has no weaknesses. Indeed, some of its problems are obvious, overshadowing its glories. Based on current research,[138] we will discuss two negative aspects of BARD: the confusion surrounding the standard and the negative byproducts of its application.

A. *The confusion surrounding the reasonable doubt standard*

As discussed in Section I, the definition of the BARD standard of proof is still not settled. After the Supreme Court of the United States in *Victor*[139] rejected the use of the traditional definition of BARD as "moral certainty" but failed to define what it is, this defect has been more exposed than ever. Various *unofficial* definitions of BARD have been developed.[140] While most complex notions can be defined in a variety of ways, this multivalence of BARD presents problems, as Laudan has pointed out:

> What we face here are not different glosses on the same notion but fundamentally different conceptions of the kind and level of proof necessary to convict someone. To make matters worse, courts have faulted all these definitions as either wrong or misleading or unintelligible. Versions that some courts have found acceptable, even exemplary, have been dismissed by other courts as violating the constitutional rights of the accused.[141]

This situation has prompted many federal and state appellate courts in the United States "to insist that trial judges should not define BARD for jurors in

---

[137] *Id.*
[138] *See, e.g.*, LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17; Allen & Laudan, *supra* note 103; Larry Laudan, *Is Reasonable Doubt Reasonable?*, 9 LEGAL THEORY 295 (2003); Shealy, *supra* note 11; Federico Picinali, *Can the Reasonable Doubt Standard be Justified? A Reconstructed Dialogue*, 31 CAN. J. L. JUR. 365 (2018).
[139] *See* Victor, 511 U.S. at 13–15.
[140] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 36–47. This section owes a significant debt to Larry Laudan, whose book argues that the BARD standard of proof is indescribable, vague, and lacking in empirical support.
[141] *Id.* at 47.

their instructions."[142] As the Seventh Circuit Court of Appeals put it, "[w]e have repeatedly admonished district courts not to define 'reasonable doubt.'"[143] A few years later, the same court reiterated that "'[r]easonable doubt' must speak for itself."[144] BARD is seen as "self-defining" and thus not in need of further commentary.[145] Thus, the Seventh Circuit has gone on record that the idea of reasonable doubt is so transparent that definitions of BARD—attempts to make "the clear more clear"—can only confuse.[146] This viewpoint is highly representative among courts.[147] However, in the eyes of Laudan and other critics, this argument for avoiding definition is "[t]he [u]ltimate [a]ct of [d]esperation."[148] As Laudan stated:

> [T]his is false on its face. Juries frequently *request* that judges explain to them what reasonable doubt is. That would not occur were it clear and self-evident. More to the point, we have already seen that different judges and different legal jurisdictions have profoundly different understandings of BARD. If judges cannot agree among themselves about this crucial notion, and it is clear that they cannot, it is a dangerous act of self-deception (or worse) to suggest that lay jurors, completely unschooled in the law, will have some common, shared understanding of this doctrine.[149]

---

[142] *Id.*
[143] U.S. v. Martin-Tregora*,* 684 F.2d 485, 493 (7th Cir. 1982).
[144] U.S. v. Glass, 846 F.2d 386, 387 (7th Cir. 1988).
[145] U.S. v. Lawson*,* 507 F.2d 433, 443 (7th Cir. 1974), *overruled on other grounds by* U.S. v. Hollinger 553 F.2d 535, 541 (7th Cir. 1997).
[146] Lawson*,* 507 F.2d 433 at 442.
[147] *See, e.g.*, Henry A. Diamond, *Reasonable Doubt: To Define, or Not to Define*, 90 COLUM. L. REV. 1716, 1719–20 (1990) ("The Courts of Appeals for the Fourth Circuit, Fifth Circuit, and Seventh Circuit have held that jury instructions defining reasonable doubt are not required and should be avoided. . . . At the state level, the highest courts of Illinois, Mississippi, Texas, and Wyoming have held that jury instructions defining reasonable doubt are not required. The Oklahoma Court of Criminal Appeals has even held that any attempt by a trial judge to define reasonable doubt automatically constitutes reversible error, and the Kentucky Rules of Criminal Procedure prohibit jury instructions that attempt to define reasonable doubt.") (internal citations omitted).
[148] *See* LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 47.
[149] *Id.* at 49 (emphasis in original). Laudan even has direct evidence that jurors are deeply confused about BARD. He mentions in his book that "in a study of some six hundred Michigan jurors, Geoffrey Kramer and Doorean Koenig discovered that *a quarter* believed that 'you have a reasonable doubt if you can see *any* possibility, no matter how slight, that the defendant is innocent.' Not surprisingly, roughly the same proportion of jurors (a quarter) agreed that 'to find the defendant guilty, beyond a reasonable doubt, you must be *100 percent* certain of the defendant's guilt.'" *Id.* at 49–50 (emphasis in original).

Laudan called this situation "a fundamental confusion at the core of criminal jurisprudence" and lamented that "the concept of proving guilt beyond a reasonable doubt—the only accepted, explicit yardstick for reaching a just verdict in a criminal trial—is *obscure, incoherent, and muddled*."[150] We are impressed by Laudan's courage here. Nonetheless, even though the confusion surrounding the BARD standard of proof is high, no alternatives have yet arisen to explain the standard of proof in criminal cases to jurors and judges—at least none that we have so far found.[151] To us, the lack of clarity in the BARD standard is more of a hard cost than a fixable problem: it serves to remind us that we live in an imperfect world. As Dean John Henry Wigmore has observed, "[t]he truth is that no one has yet invented or discovered a mode of measurement for the intensity of human belief. Hence there can be yet no successful method of communicating intelligibly . . . a sound method of self-analysis for one's belief."[152]

B. Negative byproducts of the reasonable doubt standard

As a comparison of Figures 1 and 2 indicates, all other factors being equal, a higher standard of proof in criminal trials would lead to fewer wrongful convictions but more wrongful acquittals.[153] While this distribution satisfies the Blackstone principle,[154] wrongful acquittals have their own serious social costs. One of the earliest critics to raise this concern was Jeremy Bentham; although he agreed that judges should err in favor of wrongful acquittals, he nonetheless warned "against those sentimental exaggerations which tend to give crime impunity, under the pretext of [e]nsuring the safety of innocence."[155] More recently, Laudan reminded us:

---

[150] *Id.* at 30 (emphasis added).

[151] Here, in particular, we considered a solution provided by Larry Laudan, using proof by the "clear and convincing evidence" (CACE) standard in criminal cases, but we had to reject his suggestion, as most other commentators have, finding that CACE is a much less attractive option than BARD as the standard of proof in criminal cases. *Id.* at 64–65. For reasons to reject the use of CACE as the standard of proof in criminal cases, *see e.g.,* Keith A. Findley, *Reducing Error in the Criminal Justice System*, 48 SETON HALL L. REV. 1265, 1270 (2018) ("There is indeed nothing intrinsically clearer about 'clear and convincing' than there is about 'reasonable doubt,' and Professor Laudan points to nothing providing the clarity he asserts we need.").

[152] WIGMORE, *supra* note 5 at 325.

[153] Figure 2 indicates, in comparison to Figures 1 and 3, when the SoP line slides toward to the right end of the horizontal axis, the area of dashed lines (acquitted cases),) inevitably becomes larger.

[154] *See* BLACKSTONE, *supra* note 2, at 352.

[155] JEREMY BENTHAM, A TREATISE ON JUDICIAL EVIDENCE 198 (1825). As Bentham saw it: "Public applause has been, so to speak, set up to auction. At first it was said to be better to save several guilty men, than to condemn a single innocent man; others, to make the maxim

> [The] very occurrence of [wrongful acquittals] shows that *the justice system is failing* as a device for finding out the truth and for ensuring that those who commit crimes pay for their misdeeds. . . . "We must never forget that . . . the acquittal of ten guilty persons is exactly ten times as great a *failure of justice* as the conviction of one innocent person."[156]

Even those who support the Blackstone principle seem to recognize that it likely imposes significant costs in terms of a lack of deterrence[157] and possibly increased recidivism.[158] Furthermore, except for those so-called victimless crimes,[159] wrongful acquittals mean that victims who have faced harm, or even death at the hands of wrongdoers will never receive justice. The physical harm of the crime is thus followed by psychological (and in some cases also financial) harm as a result of the criminal justice system in failing to do its job.[160] Repercussions can also ripple out from there, affecting the victim's

---

more striking, fixed on the number *ten*; a third made this ten a hundred, and a fourth made it a thousand. All these candidates for the prize of humanity have been outstripped by I know not how many writers, who hold, that, in no case, ought an accused person to be condemned, unless the evidence amount to mathematical or absolute certainty. According to this maxim, nobody ought to be punished, lest an innocent man be punished." *Id.*

[156] LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 70. (quoting CARLETON KEMP ALLEN, LEGAL DUTIES AND OTHER ESSAYS IN JURISPRUDENCE 286–87 (1931)) (emphasis added).

[157] *See* Epps, *supra* note 7, at 1092; *see also* Rosenthal, *supra* note 112, at 225 ("Indeed, the Blackstone ratio itself implies an awareness of a tradeoff – it may not be possible to minimize the rate of false convictions without unacceptably increasing false acquittals. Thus, even for Blackstonians, simply minimizing the rate of false convictions will not do.").

[158] *See* Charles H. Rose III, *Should the Tail Wag the Dog? The Potential Effects of Recidivism Data on Character Evidence Rules*, 36 N.M. L. REV. 341, 344–45 (2006) (summarizing recidivism rates among various categories of crimes and specific crimes such as rates as high as 50 percent for property crimes and 7 percent for homicide).

[159] For example, drug possession might be a victimless crime. *See* e.g., Epps, *supra* note 7, at 1091 ("The problem is that for many such crimes, the social benefit of individual convictions is more attenuated and complex than the benefit corresponding to more paradigmatic crimes involving obvious victims.").

[160] *See* Eidell Wasserman & Carroll Ann Ellis, *Impact of Crime on Victims*, in NATIONAL VICTIM ASSISTANCE ACADEMY TRACK 1: FOUNDATION-LEVEL TRAINING 6–1 (2010), https://ce4less.com/Tests/Materials/E075Materials.pdf ("The trauma of victimization can have a profound and devastating impact on crime victims and their loved ones. It can alter the victim's view of the world as a just place and leave victims with new and difficult feelings and reactions that they may not understand.").

family members, loved ones, and friends, reducing their faith and confidence in the criminal justice system.[161]

Another potential, though less obvious, negative byproduct of the BARD standard is its likely burden on the goal of *error reduction*. Because the BARD standard is designed to achieve a weighted *error distribution* in accordance with the Blackstone principle,[162] it may actually have a negative effect on the goal of *error reduction*. This problem reflects a strict interpretation of the Blackstone principle, which tells us that "in distributing criminal punishment, we must strongly err in favor of false negatives (failures to convict the guilty) in order to minimize false positives (convictions of the innocent), *even if doing so significantly decreases overall accuracy*."[163] Even in our demonstrative diagrams, we can see that when SoP is the reasonable doubt standard (Figure 2), the total number of wrongful cases of all types (represented by the dark areas of the diagram) is greater than when SoP is somewhere lower than the BARD standard (Figure 1). Similarly, the total number of correct cases (represented by the light areas) is smaller in Figure 2 than in Figure 1. Thus, a quick comparison between Figures 1 and 2 tells us that when the BARD standard of proof is applied in criminal cases, the overall accuracy of judicial fact-findings may decline, which contradicts the goal of *error reduction*. And yet, theoretically speaking, *error reduction* is even more fundamental in jurisprudence than is *error distribution*, which functions primarily as a systematic remedy to address the cases that are left out when we strive for *error reduction*.

IV. THE EFFECT OF THE BEYOND A REASONABLE DOUBT STANDARD IN PRACTICE: EMPIRICAL MEASUREMENT AND ITS LIMITS

Both those for and against applying the reasonable doubt standard in criminal cases have sound arguments on their side. More seriously, as multiple researchers have pointed out, wrongful convictions and wrongful acquittals are two very different types of harm, and proponents and critics of the Blackstone principle may simply be "talking past each other."[164] When theoretical understandings are at an impasse, it makes sense to seek out real-life applications and empirical evidence. The legal field is no exception, since "law is ultimately a profession that shapes the way in which individuals and

[161] *Id.* ("Crime has significant, yet varying consequences, on individual crime victims, their families and friends, and communities. The impact of crime on victims results in emotional and psychological, physical, financial, social and spiritual consequences.").
[162] *See* BLACKSTONE, *supra* note 2, at 352.
[163] *See* Epps, *supra* note 7, at 1068 (emphasis added).
[164] *See id.* at 1092; Zalman, *supra* note 18, at 1325–26.

institutions behave. . . ."[165] If possible, *any* legal rule or theory should be examined empirically for observable outcomes.[166] Such empirical data can help expose the distance between the law-in-book and the law-in-action,[167] facilitating further improvement of the rule. For a concept like the BARD standard—whose very existence and definition are both mysterious and controversial—empirical data is even more valuable. The empirical question we would like to ask is what is the effect of the BARD standard in judicial practice? Unfortunately, very few extant empirical studies would allow us to analyze how well the BARD standard functions in real cases. The rest of this section attempts to develop such a study to fill the void.

A. *Thinking Empirically: The Problem of Empirical Study on the Reasonable Doubt Standard*

In an ideal criminal justice system, no error of any kind would happen: no criminal could escape the punishment of law, no innocent person would ever be convicted, and the total number of convictions would equal the total number of crimes. However, such a perfect state is impossible in practice. Litigation involves human decision-making, in which errors (both false positives and false negatives) are inevitable.[168] In the pre-trial context, law enforcement and the legal system both face budgetary restraints, which lead to certain crimes and charges being prioritized while others are deemed "de minimis" and not charged.[169] Errors may also happen at trial; these generally

---

[165] *See* Zalman, *supra* note 18, at 1326.

[166] *See, e.g.*, George J. Stigler, *The Development of Utility Theory. II.*, 58 J. POL. ECON. 373, 395 (1950) ("This third criterion of congruence with reality should have been sharpened, sharpened into the insistence that theories be examined for their implications for observable behavior, and these specific implications compared with observable behavior. The implication of the diminishing marginal utility of money, that people will not gamble, should have been used to test this assumption, not to reproach the individuals whose behavior the theory sought to describe.").

[167] *See generally* Roscoe Pound, *Law in Books and Law in Action*, 44 AM. L. REV. 12 (1910).

[168] Karl Mason, *Learning from Error in the Criminal Justice System*, U.S. DEPARTMENT OF JUSTICE OFFICE OF JUSTICE PROGRAMS (Sep. 16, 2014), https://www.ojp.gov/sites/g/files/xyckuh241/files/archives/blogs-2014/2014-blog-learningfromerror.htm.

[169] *See* Ioannis Lianos, et al., *An Optimal and Just Financial Penalties System for Infringements of Competition Law:* A *Comparative Analysis*, CTR. FOR L., ECON. & SOC'Y UCL FAC. OF LAWS 149 (May 2014), https://www.fne.gob.cl/wp-content/uploads/2014/11/Estudio.pdf.

involve wrongful convictions, wrongful acquittals, procedural errors, or evidentiary errors.[170] Finally, most post-trial errors occur in sentencing.[171]

For our purposes, we focus on epistemological errors in the verdict of a criminal trial. If the trial verdict is in accordance with the truth, then it is epistemologically correct (either a correct conviction or a correct acquittal); if the trial verdict is contradictory with the truth, then it is epistemologically erroneous (either a wrongful conviction or a wrongful acquittal). Ideally, an empirical study on the impact of the BARD standard on a criminal justice system—in terms of its effectiveness on the goal of error distribution and its burden on the goal of error reduction—would make use of a complete set of empirical data on wrongful convictions and wrongful acquittals in a designated criminal justice system where the BARD standard is implemented. This data would allow us to check: (a) whether the actual ratio of wrongful acquittals and wrongful convictions is in accordance with the Blackstone principle;[172] and (b) whether the total number of wrongful acquittals and wrongful convictions overburdens the criminal justice system.

However, in reality, no one has such comprehensive data. Both wrongful convictions and wrongful acquittals are unknown data for the most part; occasionally, wrongful convictions have been exposed through various channels, like through the advancement of forensic technologies such as DNA typing, reminding us that wrongful convictions indeed happen in criminal cases.[173] Wrongful acquittals are even less visible, even though we sense that their frequency of occurrence (more likely than not) would be even higher than that of wrongful convictions. This deadly dilemma of the lack of data on

---

[170] *See* Fritz Allhoff, *Wrongful Convictions, Wrongful Acquittals, and Blackstone's Ratio*, 43 Australasian J. of Legal Phil. 39, 39 (stating that "criminal law can err in either of two ways: it can wrongly convict the innocent, or it can wrongly acquit the guilty."); *Wrongful Convictions*, NATIONAL INSTITUTE OF JUSTICE, https://nij.ojp.gov/topics/justice-system-reform/wrongful-convictions (last visited Sep. 30, 2021) (explaining that " [a] conviction may be classified as wrongful [if] . . . [t]here were procedural errors that violated the convicted person's rights."); *Evidentiary Errors*, NORTH CAROLINA CENTER ON ACTUAL INNOCENCE, https://www.nccai.org/causes/evidence-error (last visited Sep. 30, 2021) (arguing that "[i]mproper collection, labeling, or preservation of evidence, or the less than complete processing of a crime scene, can be the first steps in the road to conviction for an innocent person. Once the evidence reaches the lab, errors can multiply.").

[171] *See* Brandon Hasbrouck, *Saving Justice: Why Sentencing Errors Fall Within the Savings Clause, 28 U.S.C. § 2255(e)*, 108 GEO. L.J. 287, 289 (2019) (stating that "[i]nmates are sitting in federal prisons serving unlawful sentences. Many will die in those prisons serving 'unjust' sentences.").

[172] *See* BLACKSTONE, *supra* note 2, at 352.

[173] *See* EDWARD CONNORS ET AL., CONVICTED BY JURIES, EXONERATED BY SCIENCE: CASE STUDIES IN THE USE OF DNA EVIDENCE TO ESTABLISH INNOCENCE AFTER TRIAL iii (1996).

both wrongful convictions and wrongful acquittals indeed deserves further discussion.

i. Data on wrongful convictions

Wrongful convictions are at best partially observable. Rarely and randomly, some wrongful convictions have been belatedly exposed by DNA evidence and/or confessions from the real perpetrators,[174] but the data on such cases are scarce, or in Chinese "可遇而不可求" (something that we cannot ask for). Data collectors may thus have to rely on societal efforts to expose wrongful convictions. In recent years, an increasing number of privately funded projects and pro bono clinics have aimed to exonerate the wrongfully convicted and prevent future injustice.[175] Such projects often attract media attention (every story of redemption is encouraging) in addition to providing belated justice to the wrongfully convicted, their families, and to society as a whole. The National Registry of Exonerations—a project collectively led by the University of California Irvine Newkirk Center for Science and Society, University of Michigan Law School, and Michigan State University College of Law—has collected and publicized information about all known exonerations of innocent criminal defendants in the United States since 1989.[176] As of October 2021, according to the U.S. National Registry of Exonerations, there have been 2,872 exonerations since the project's inception, and those wrongfully convicted have collectively lost more than 25,600 years to prison.[177]

While those numbers are significant, they hardly represent all wrongful convictions. Exoneration projects are restrained by budgetary

[174] *See*, *e.g.*, Thomas McGowan, *Freeing the innocent and identifying true perpetrators*, INNOCENT PROJECT BLOG (Aug. 1, 2008), https://innocenceproject.org/freeing-the-innocent-and-identifying-true-perpetrators (quoting a former exoneree after he served 23 years for a rape he did not commit as saying, "I never saw Kenneth Wayne Woodson [the actual perpetrator]; I don't know if he ever saw me. He went to prison a year later than I did. I'm glad he *confessed*, but I think the only reason he did is because of the DNA hit.").
[175] Many exoneration projects have been developed through the growth in forensic technology. For example, DNA technology inspired the development of the Innocence Project, which has helped 375 DNA exonerees to date, of which twenty-one had been serving time on death row, and forty-four had pled guilty to crimes they did not commit. Furthermore, the exonerated had served an astonishing fourteen years on average at the time of their exoneration. *See DNA Exonerations in the United States*, INNOCENCE PROJECT, https://www.innocenceproject.org/dna-exonerations-in-the-united-states (last visited Sep. 30, 2021).
[176] *About The Registry*, The National Registry of Exonerations, https://www.law.umich.edu/special/exoneration/Pages/about.aspx (last visited Sept. 30, 2021).
[177] *Id.*

concerns (of both financial and human capital) and thus may have to focus primarily on high-profile cases, such as murder or rape, not low-profile cases, whether felony or misdemeanor, that are mainly handled by local attorneys and are "passively or actively concealed from public attention."[178] In addition, while groundbreaking forensic technology has also contributed to the fight against wrongful convictions, it is also not a panacea in this situation; DNA typing, for example, is still circumstantial evidence, which cannot not directly tell us who did what wrong things.[179] As such, from an epistemological point of view, wrongful conviction data are at best partially observable.

ii. Data on wrongful acquittals

Just as with wrongful convictions, wrongful acquittals can be revealed through ex-post self-reporting by the wrongdoers, individuals who have "luckily" avoided punishment in the criminal justice system.[180] However, not surprisingly, there exists little incentive for such confessions; likewise, minimal incentive exists for social efforts to uncover wrongful acquittals because such individuals are constitutionally protected against double jeopardy and res judicata.[181] Therefore, reliable data sources of wrongful acquittals are almost nonexistent.

Perhaps largely due to such severe shortage of data on wrongful convictions and acquittals, little meaningful work has been done in terms of empirical studies in this field.[182] How, then, do we get around the problem of insufficient sample data?

---

[178] *Exonerations in the United States, 1989–2012*, National Registry of Exonerations (May 20, 2012), https://www.law.umich.edu/special/exoneration/Documents/exonerations_us_1989_2012_summary.pdf.

[179] Kevin Keller, *The Cognitive Psychology of Circumstantial Evidence*, 105 MICH. L. REV. 241, 241 (2006) ("Empirical research indicates that jurors routinely undervalue *circumstantial evidence* (*DNA*, fingerprints, and the like) and overvalue direct evidence (eyewitness identifications and confessions) when making verdict choices, even though false-conviction statistics indicate that the former is normally more probative and more reliable than the latter.") (emphasis added).

[180] *See e.g.*, O.J. SIMPSON ET AL., IF I DID IT: CONFESSIONS OF THE KILLER XLV (2007).

[181] *See* U.S. CONST. AMEND. V ("No person shall . . . be subject for the same offence to be twice put in jeopardy of life or limb . . . ."); *see also The Double Jeopardy Clause as a Bar to Reintroducing Evidence*, 89 Yale L.J. 962, 962 (1980).

[182] *See, e.g.*, LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW, *supra* note 17, at 8 ("In writing this book, I have been constantly frustrated by the paucity of empirical information that would allow us to reach clear conclusions about how well or badly our legal methods are working. Where there are reliable empirical studies with a bearing on the issues addressed here, I will make use of them. Unfortunately, given the dearth of hard evidence, the analysis

B. An alternative empirical model for testing the effect of the reasonable doubt standard on restraining wrongful convictions

Because the ideal model of empirical study on BARD is not possible due to the insufficiency of the sample data (particularly on wrongful acquittals), we developed an alternative approach to *approximate* the effect of the reasonable doubt standard in restraining wrongful convictions in a criminal justice system.

The harsh reality is that an empirical study on the impact of the BARD standard on total *error reduction* (the first goal of criminal justice) and *error distribution* (the second goal of criminal justice) cannot be meaningfully done since it would need data on wrongful acquittals. To account for this, we narrowed our scope of analysis to the effect of the BARD standard on restraining wrongful convictions, the standard's most important function as designed. Even then, we had to continue adjusting from the ideal empirical model because of the limited availability of direct data (confessions and exonerations) on wrongful convictions. Thus, we decided to focus on statistically available but indirect data, choosing three indicators that could approximate changes in wrongful convictions in a criminal justice system (collectively, "Indicators"):

i. Indicator 1: The percentage of criminal cases that are awarded with the verdict of not guilty ("acquittal rate")

ii. Indicator 2: Appellate courts' reversals on criminal trial judgments based on a failure to satisfy the BARD standard ("reversals")

iii. Indicator 3: Non-prosecution rate

We hypothesize that: (a) more acquittals hint at less wrongful convictions, and vice versa; (b) more reversals hint at less wrongful convictions, and vice versa; and (c) a higher rate of non-prosecution hints at less wrongful convictions, and vice versa. These suppositions are based in common sense: when the total number of pending criminal cases is fixed, the presence of more acquittals, reversals, and non-prosecutions would mean that fewer cases will receive a verdict of guilty, thus leaving less chance for wrongful convictions to occur.[183]

---

in this book will fall back on armchair hunches about the likely effects of various rules and procedures far more often than I would have liked.").

[183] This adjusted empirical study is aimed to show whether changing the formal standard of proof changes outcomes. What it does not tell us is whether those outcomes are correct

While collective data of these three *Indicators* should not be difficult to gather in a criminal justice system with open data access for the public, it is still difficult, if not impossible, to isolate the impact of the BARD standard from that of other factors including the presumption of innocence, burden of proof, and other defendant friendly rules. All these factors are embedded in a criminal proceeding, protecting the constitutional rights of the accused and preventing wrongful convictions.[184] We need more information than just the three *Indicators* to ascertain the impact of the BARD standard on restraining wrongful convictions.

To implement our study, we needed to satisfy two more conditions. First and foremost, we needed a clear benchmark date on which the BARD standard of proof was added to (or excluded from) a criminal justice system. This date would allow us to collect data for two comparable periods of time; the empirically observable periods of time that immediately preceded and followed the benchmark. By comparing collected data of the three *Indicators* during these periods, we can approximate the impact of the BARD standard on the designated criminal justice system. The second necessary condition is that all other factors in the designated criminal justice system should remain in the same or at least similar condition throughout the entire period of observation. Only if all other variables stayed relatively stable could we attribute the observed variations in the *Indicators* to the impact of the BARD standard.

Still, even if these two conditions are met, correlation does not by itself imply causation;[185] and the collected data of the three *Indicators* are not direct evidence of wrongful convictions. That is why we emphasize here that this empirical study can only approximate the effectiveness of BARD in restraining wrongful convictions.

These necessary conditions for such a study severely limit its likelihood of being implemented. Because no country would allow such an experiment with its criminal justice system, we must turn to countries with pertinent, extant data. And yet, as discussed in the beginning of this article, the BARD standard originated in common law countries centuries ago and

---

or incorrect. For example, more acquittals could also possibly mean more wrongful acquittals. However, that is not the subject of our study. Here, we focus on more acquittals meaning less convictions (supposing that the total number of criminals trials in the scope of study is fixed), thus less *opportunities* for having wrongful convictions.

[184] *See e.g.*, *Model Charge: Burden of Proof Presumption of Innocence, Reasonable Doubt*, NEW HAMPSHIRE JUDICIAL BRANCH, https://www.courts.nh.gov/model-charge-burden-proof-presumption-innocence-reasonable-doubt (last visited Sep. 30, 2021) (stating that "[u]nder our constitutions, all defendants in criminal cases are presumed to be innocent until proven guilty beyond a reasonable doubt. The burden of proving guilt is entirely on the State. The defendant does not have to prove his innocence.").

[185] *See, e.g.*, Randall Morck & Bernard Yeung, *Economics, History, and Causation*, 85 BUS. HIST. REV. 39, 61 (2011).

lacks a clear benchmark date.[186] Thus, such a study has to be done in a country in which criminal proceedings have recently been changed with regard to the BARD standard, in which sufficient and available data on the three *Indicators* exist for the periods of time surrounding the benchmark event, in which the legal system has otherwise been largely unchanged throughout the period of observation. Such a list of countries is short, indeed, but such a case does exist.

V. A CASE STUDY ON CHINA

We turn to China, whose criminal procedural law underwent a significant change in 2013, making possible a rare research opportunity for an empirical study on the BARD standard. In that year, China, for the first time, introduced and implemented the reasonable doubt standard in its criminal proceedings through an amendment to its *Criminal Procedure Law* that took effect as of January 1, 2013.[187] Together with China's free access to nationwide judicial data and the otherwise relatively unchanged status of the Chinese criminal justice system at this time, this development presents an unique opportunity to observe and evaluate the effect of the BARD standard on restraining wrongful convictions in the nation.[188]

We begin this section by introducing the unique judicial data set available in China and explaining why it enables us to run an empirical study as laid out in Part IV.[189] We also delve into definitions of a host of quantitative measures that will recur throughout the remainder of this article. We then apply and analyze the collected Chinese judicial data through the three

[186] *See infra* Introduction and Part I-A.

[187] Zhonghua Renmin Gonghe Guo Xingshi Susong Fa (中华人民共和国刑事诉讼法) [Criminal Procedure Law of the People's Republic of China] (promulgated by the Standing Comm. Nat'l People's Cong., July 1, 1970, rev'd Mar. 17, 2012, effective Jan. 1, 2013), art. 53 (China), http://www.gov.cn/flfg/2012-03/17/content_2094354.htm.

[188]*See* Zuigao Renmin Fayuan Guanyu Shiyong "Zhonghua Renmin Gonghe Guo Xingshi Susong Fa" De Jieshi (最高人民法院关于适用《中华人民共和国刑事诉讼法》的解释) [The Interpretation of the Supreme People's Court concerning the Implementation of the Criminal Procedure Law of the People's Republic of China] (promulgated by Jud. Comm. Sup. People's Ct., Nov. 5, 2012, effective Jan. 1, 2013), art. 105 (China), https://www.spp.gov.cn/sscx/201502/t20150217_91462.shtml; Zhonghua Renmin Gonghe Guo Zhuxi Ling: Quanguo Renmin Daibiao Dahui Guanyu Xiugai "Zhonghua Renmin Gonghe Guo Xingshi Susong Fa" De Jueding" (中华人民共和国主席令：全国人民代表大会关于修改《中华人民共和国刑事诉讼法》的决定) [The order of the President of the Central People's Government of the People's Republic of China, announcement regarding the amendment to the Criminal Procedure Law of the People's Republic of China], (March 14, 2012), http://www.gov.cn/flfg/2012-03/15/content_2092191.htm.

[189] *See supra* Part IV.

Indicators defined in Part IV,[190] and we conclude this section by presenting our findings and the limitations of the study. Ultimately, in this Chinese case, the BARD standard appears to be more of a dogmatic tool than a pragmatic method to be used in the prevention of wrongful convictions.

A. *The unique opportunity of China as a test case*

China presents a unique opportunity to study the effectiveness of the BARD standard on restraining wrongful convictions for a number of reasons. First and foremost, to the best of our knowledge, China is the only country to have launched the BARD standard into its criminal justice system in the twenty-first century.[191] The reasonable doubt standard was first introduced in the *Criminal Procedure Law of the People's Republic of China* in 2012, made effective as of January 1, 2013, as an answer to the political and public outcry that followed the exposure of several wrongful convictions in the first decade of the century.[192] The BARD standard was added to Article 53 as a new subclause to supplement the original standard of proof, which had required evidence to be "reliable and sufficient" (证据确实充分标准) for fact-finders to reach a guilty verdict of any criminal case.[193] This particular legislative moment offers the benchmark (January 1, 2013) that we needed to activate an empirical study on the effect of BARD.[194] Besides offering the necessary

---

[190] *Id.*

[191] Japan introduced BARD into its criminal justice system with a traceable record, doing so in 1945. *See Reforms in Japanese Criminal Procedure under Allied Occupation,* 24 WASH. L. REV. 401, 412 (1949).

[192] *See* Wang Minyuan (王敏远) et al., Bitan: Cuoan, Sixing Yu Fazhi (笔谈:错案、死刑与法治) [Discussions: Erroneous Cases, Death Penalty and Rule of Law], 3 (中外法学) [PEKING UNIV. L.J.] 565, 582 (2015); *see generally* Chen Yongsheng, *A Perspective of China's Criminal Misjudgment*: *An* A*nalysis of* Twenty Wrongful Convictions *That Shocked The Whole Country As A Sample*, 3 CHINA LEGAL SCIENCE 45 (2007); Chen Yongsheng (陈永生), Woguo Xingshi Wupan Wenti Toushi: Yi 20 Qi Zhenjing Quanguo De Xingshi Yuanan Wei Yangben De Fenxi (我国刑事误判问题透视——以20起震惊全国的刑事冤案为样本的分析) [A Perspective of China's Criminal Misjudgment: An Analysis of Twenty Wrongful Convictions That Shocked The Whole Country As A Sample], 6 Zhengfa Luntan (政法论坛) [TRIBUNE OF POLITICAL SCIENCE AND LAW] 3 (2011).

[193] *See* Wei Hanqing, *The Integration and Practice of "Beyond Reasonable Doubt" in an Inquisitorial Legal System—Taking China as an Example*, 68 STOCKHOLM UNIV. RESEARCH PAPER 1, 2 (2019).

[194] Here, we chose January 1, 2013, the effective date of the BARD standard in China, as the benchmark for the empirical study, including for Indicator 2 (appellate reversal rates). We understand, however, that the time delay between conviction and appeal means that cases that resulted in conviction under the old SoP would still wind their way through the

benchmark date, this legislative moment allows a clear comparison of data on the three *Indicators* (annual acquittal rate, number of criminal cases reversed, and non-prosecution rate) in a defined period of time before and after the benchmark, which in turn allows us to extrapolate the impact of BARD. Collectively speaking, such a scenario offers unparalleled empirical insight into the effectiveness of BARD on restraining wrongful convictions.

Furthermore, this case study fulfills the necessary condition of constancy (*ceteris paribus*) in that throughout the designated period of observation, other than the introduction of the BARD standard, China's criminal justice system was largely unchanged. An indicator of such systematic stability is that before the *2012 Amendments* were promulgated, the *Criminal Procedure Law of China* was most recently amended in the year 1996.[195] The major change to the Chinese criminal justice system subsequent to the *2012 Amendments* emerged in November 2018, when the *Criminal Procedure Law of China* was amended to add a host of new concepts into the Chinese criminal justice system, including a whole new mechanism of plea bargaining (with Chinese characteristics) that would supposedly directly affect the outcome of many criminal cases in China.[196] Because such a major change to Chinese criminal procedure would commingle with our study target and dilute analytic results,[197] we chose to use this date as the cutoff for our observation period. Our study covers a range of twelve years centered around the benchmark date of January 1, 2013, when the BARD standard was implemented: six years after this date (2013–18), to the time when the *Criminal Procedure Law of China* was amended; and six years before this date (2007–12). The six preceding years may be considered as an immediately relevant period that led into the promulgation of BARD, and the six following

appellate courts long after the formal change to the BARD standard. Nonetheless, we cannot determine what other specific timing would be a better changepoint for this study. Thus, using January 1, 2013, as the benchmark is a compromised choice made for this study.

[195] Before the *1996 Amendments* were promulgated, the Criminal Procedure Law of China, enacted in 1979, was never amended. *See* Zhonghua Renmin Gonghe Guo Xingshi Susong Fa (中华人民共和国刑事诉讼法) [Criminal Procedure Law of the People's Republic of China] (promulgated by the Standing Comm. Nat'l People's Cong., July 1, 1979, rev'd Mar. 17, 1996) (China).

[196] *See* Zhonghua Renmin Gonghe Guo Xingshi Susong Fa (中华人民共和国刑事诉讼法) [Criminal Procedure Law of the People's Republic of China] (promulgated by the Standing Comm. Nat'l People's Cong., July 1, 1979, rev'd Oct. 26, 2018) (China); Han Xu (韩旭), 2018 Nian Xingsu Fa Zhong Renzui Renfa Congkuan Zhidu (2018年刑诉法中认罪认罚从宽制度) [The Leniency System of Pleading Guilty and Accepting Penalty in the 2018 Amendment to the Criminal Procedure Law of China], 1 Fazhi Yanjiu (法治研究) [Research on Rule of Law] 35–45 (2019).

[197] *See* Criminal Procedure Law of the People's Republic of China, *supra* note 196.

years may be considered as the period that would reflect the greatest impact of the BARD standard on the Chinese criminal justice system.[198]

In addition, this study has been made possible by China's movement to judicial transparency and digitalization in recent years. Before 2012, most Chinese court files, especially important ones like judgments, were only available in hard copy and were largely inaccessible to the public.[199] In July 2013, the Supreme People's Court of China launched *China Judgments Online* (中国裁判文书网) (CJO), a database and search engine for court files nationwide, with open and free access to the public.[200] According to the *Provisions of the Supreme People's Court on the Issuance of Judgments on the Internet by the People's Courts*, promulgated in November 2013, all levels of courts in China are required to upload and publish judgments on CJO within seven days after the date on which they come into effect and become binding.[201] Today, CJO holds more than 100 million case files, including more than nine million criminal judgments, and has collected judicial data from as far back as 1996.[202]

The clear introduction of the reasonable doubt standard, a determinable study period during which the nation's criminal justice system was stable, and the open access to judicial data in China enables a unique case

---

[198] Admittedly, it is too optimistic to hope that the twelve-year period of the study would offer a truly *ceteris paribus* analytic environment, especially given how fast China has been developing since the launch of the "Reform and Opening Up" (改革开放) national policy in 1978. To take into account the most statistically observable socioeconomic measures (the gross domestic product and changes in the overall crime rate), we further conducted a regression analysis (an analytic method of econometrics) on the data of Indicator 1 (i.e., acquittal rate) to filter out the impact of such variables on this Indicator. We also did a simple regression test on the data of Indicator 3. In both cases, these tests allowed us to extrapolate a less biased estimate of the effectiveness of the BARD standard on the data. In contrast, our study for Indicator 2 (i.e., reversals) did not warrant a regression analysis to consider other social economic factors, partially due to data limitation. *See infra* Part V-D.

[199] *See* QI QI (齐奇), RULE OF LAW IN CHINA AND JUDICIAL TRANSPARENCY (法治中国与司法公开) (2014).

[200] *See* CHINA JUDGMENTS ONLINE (中国裁判文书网), http://wenshu.court.gov.cn (last visited Oct. 4, 2021); *Judicial Transparency of Chinese Courts*, art. 3., THE SUPREME PEOPLE'S COURT OF THE PEOPLE'S REPUBLIC OF CHINA (July 20, 2015), http://english.court.gov.cn/2015-07/20/content_21332354.htm.

[201] *See* Zuigao Renmin Fayuan Guanyu Renmin Fayuan Zai Hulian Wang Gongbu Caipan Wenshu De Guiding (最高人民法院关于人民法院在互联网公布裁判文书的规定) [The Supreme People's Court's Rules on Publishing Judicial Opinions on the Internet], SUPREME PEOPLE'S COURT OF THE PEOPLE'S REPUBLIC OF CHINA (Nov. 29, 2013), http://www.court.gov.cn/zixun-xiangqing-5867.html.

[202] *See* CHINA JUDGMENTS ONLINE, *supra* note 200 (showing real-time updated data shown on the front page) (last visited Oct. 4, 2021).

study to evaluate the effectiveness of the BARD standard in restraining wrongful convictions.

B. *The Chinese judicial data under study*

For nationwide data on acquittals (*Indicator 1*) and non-prosecutions (*Indicator 3*) in China between 2007 and 2018, we relied on statistics disclosed in the *Annual Working Reports* of the Supreme People's Court and the Supreme People's Procuratorate, respectively.[203] These official reports from the two highest judicial authorities in China[204] provide a good overview of the operations of the Chinese criminal justice system, with national data on various aspects of judicial practice from the local to the central level, including national data on the annual acquittal rate and the non-prosecution rate.[205] The *Annual Working Reports* are presented by the chief justice and the chief procurator each year at the annual meeting of the P.R.C. National People's Congress, which then has to review and approve the reports before they become accessible to the general public.[206]

---

[203] *See* The Annual Working Report of the Supreme People's Procuratorate (最高人民检察院工作报告) *See* SUPREME PEOPLE'S PROCURATORATE OF THE PEOPLE'S REPUBLIC OF CHINA, https://www.spp.gov.cn/spp/gzbg/index.shtml (last visited Oct. 4, 2021); The Annual Working Report of the Supreme People's Court (最高人民法院工作报告), SUPREME PEOPLE'S COURT OF THE PEOPLE'S REPUBLIC OF CHINA, http://www.court.gov.cn/zixun.html (last visited Oct. 4, 2021).

[204] *See* Zhuhao Wang & David Caruso, *Is an Oral-Evidence Based Criminal Trial Possible in China?*, 21 INT'L J. EVIDENCE & PROOF 52, 57–58 (2018) (explaining that in China, the concept of the judicial branch of government is more expansive than its counterpart in the United States and many other countries. At the highest level, it includes the Ministry of Public Security, the Supreme People's Court, the Supreme People's Procuratorate, and the Ministry of Justice. To the contrary, in the United States, for example, the judicial branch of government is made up of the court system only).

[205] As of the time of this study, we only have access to aggregate macro-level data. As case specific data points become more available in the future, we are hopeful that our analysis remains relevant.

[206] *See, e.g.*, Lianggao Baogao Zancheng Lv Chuang Shinian Xingao, Ciqian Jiunian Congwei Chaoguo 85% (两高报告赞成率创十年新高，此前九年从未超过85%) [The Approval Rate of Both the Supreme People's Court and the Supreme People's Procuratorate Hits a Ten-Year High, which Has Never Exceeded 85% in the Previous Nine Years], Zhongguo Qingnian Bao (中国青年报) [CHINA YOUTH DAILY] (March 16, 2015), http://www.chinanews.com/gn/2015/03-16/7130454.shtml; Lianggao Baogao Tongguo, Zuigao Fa Fandui Piao Zengduo Zuigao Jian Depiao Lv Fanchao ("两高"报告通过，最高法反对票增多最高检得票率反超) [The Annual Working Report of Both the Supreme People's Court and the Supreme People's Procuratorate Passed, with "No" Votes Against the SPC Increased While "Yes" Votes for the SPP Exceeded the Other], Caixin Wang (财新网) [CAIXIN] (March 15, 2019), http://topics.caixin.com/2019-03-15/101392678.html.

Nationwide data on reversals of criminal cases (*Indicator 2*) in China between 2007 and 2018 are not included in the *Annual Working Reports* of the Supreme People's Court and the Supreme People's Procuratorate, and so we turned to other data sources. The next best available alternative, to our knowledge, is CJO, which collects court files of reversed criminal cases. In addition to CJO, several other popular Chinese online judicial databases exist, including *Gather Law Cases Online* (聚法),[207] *OpenLaw Judgments Search,*[208] *Legal Information Center of Peking University* (北大法宝),[209] and *WuSong Tech* (无讼)[210]. While each database is different, the judicial documents and other materials in these databases largely overlap with the ones in CJO.

Despite the additional features offered by other databases, we chose to use CJO to research and collect data for *Indicator 2* for three primary reasons. First, as mentioned above, CJO is the only judicial database launched and maintained by the Supreme People's Court,[211] thus it is the most authoritative one. Second, while some other databases do occasionally report marginally more cases than CJO from recent years, they all lack substantial amounts of data from the era before 2013.[212] In contrast, CJO has collected case materials as early as 1996 and holds the largest number of criminal judgments and related searchable materials from the era before 2013.[213] Thus,

---

The two judicial authorities even have an internal "competition" for the quality and rate of satisfaction of their Annual Working Reports. All representatives of the National People's Congress of China vote Yes, No, or Abstention at the annual meeting on the question of approving the two reports. Each year, the Chinese media are keen to report on which of the two Annual Working Reports received a higher number of Yes votes from the National People's Congress. For example, in 2015, the Supreme People's Court won with 2,619 Yes votes, beating the Supreme People's Procuratorate, which had 2,529 Yes votes. In 2019, the Supreme People's Procuratorate won with 2,843 Yes votes versus the other's 2,725 Yes votes.

[207] GATHER LAW CASES ONLINE (聚法), https://www.jufaanli.com (last visited Oct. 4, 2021).

[208] OPENLAW JUDGMENTS SEARCH, https://openlaw.cn (last visited Oct. 4, 2021).

[209] LEGAL INFORMATION CENTER OF PEKING UNIVERSITY (北大法宝), http://www.pkulaw.cn (last visited Oct. 4, 2021).

[210] WUSONG TECH (无讼), https://www.itslaw.com (last visited Oct. 4, 2021).

[211] *See* CHINA JUDGMENTS ONLINE, *supra* note 200.

[212] *See generally* Yingmao Tang (唐应茂), *Judicial Disclosure and Its Determinants: Data Analysis Based on China Judgment Online* (司法公开及其决定因素:基于中国裁判文书网的数据分析), CHINESE INSTITUTE FOR SOCIO-LEGAL STUDIES, SHANGHAI JIAO TONG UNIVERSITY (Mar. 26, 2021).

[213] *See* CHINA JUDGMENTS ONLINE, *supra* note 200. As we will address later, CJO still lacks much data for the period between 2007 and 2013, although its records are more complete than any other online judicial database of China.

CJO has the most complete data for criminal cases between 2007 and 2018, the observation period of our study. In addition, CJO allows users to search any arrangement of keywords in any specific portion(s) of a case judgment, an advanced function not provided by competing databases, rather than simply running a full-text search for keyword (e.g., searching "BARD and/or standard of proof and/or reliable and sufficient evidence" in the portion of a judge's reasoning, and searching "reverse and remand" in the judgment portion).[214] Such strong search functions allowed us to track targeted usage of BARD-related keywords including "standard of proof" (证明标准), "beyond a reasonable doubt" (排除合理怀疑), "reasonable doubt" (合理怀疑), "no punishment in doubtful cases" (疑罪从无) (a traditional Chinese phrase with the same meaning as BARD), and "reliable and sufficient evidence" (证据确实充分) (the traditional Chinese standard of proof in criminal cases). We targeted these searches to the relevant portion(s) of appellate court decisions for reversal (e.g., the appellate judges' reasoning), while ignoring irrelevant portion(s) of the same documents (e.g., the summary of both parties' arguments). This allowed us to affirm whether such keywords played a substantial role in a given appellate court's reversal decision.[215]

C. The roadmap: Recap of the three Indicators in the context of China

Theoretically speaking, as applied anywhere in the world, once "beyond a reasonable doubt" is made an official standard of proof in criminal trials, it would be expected to affect every stage of a criminal proceeding—ranging from the pre-trial evidentiary hearing all the way through the appellate courts, where a case may be overturned due to failure of the BARD standard. Since direct data on wrongful acquittals is unobservable and direct data on wrongful convictions is, at best, partially observable, we attempted here to use the three *Indicators* as proxies, to approximate the effectiveness of the BARD standard in the reduction of wrongful convictions in Chinese criminal proceedings based on three hypotheses.

---

[214] *Id.*

[215] An appellate court judge may reverse a criminal trial judgment on various grounds. For our purposes, we focused the data collection for *Indicator 2* on those reversals that were based on the standard of proof in criminal cases not being satisfied.

i. Indicator 1: Acquittal rate

The first hypothesis is *if the reasonable doubt standard is effective, it will safeguard the accused in those cases in which evidence is weak and would not support surviving the BARD standard.* Such cases would be expected to end up with acquittals. If this hypothesis is valid, we would expect the national acquittal rate in Chinese criminal cases to increase in the years after January 1, 2013, as compared to the immediately preceding period. To test *Indicator 1*, we first observed changes in the national annual acquittal rate as published in the *Annual Working Reports of the People's Supreme Court* for the period between 2007 and 2018.[216] However, correlation may not suggest causation; to further distill the effectiveness of BARD, we then performed a multifactor linear regression analysis[217] to test the effect of BARD on changes in the annual acquittal rate by taking into account two major and quantifiable socioeconomic factors, namely the gross domestic product and the overall crime rate in China. If the introduction of these two variables caused the coefficient on the fixed effect of BARD to be statistically insignificant, then the result would suggest that other social economic factors play a bigger role than the BARD standard in changes in the annual acquittal rate in China.[218]

---

[216] The Annual Working Report of the Supreme People's Court, *supra* note 203.

[217] A multifactor regression is a statistical model that describes the relationship between one dependent variable and multiple independent variables, whereas a dependent variable can be viewed as the "outcome" variable and the independent variable can be viewed as the "control" that can be manipulated or changed to have a direct effect on the dependent variable. Regression analysis can be used to analyze the causal relation for the impact of the independent variable(s) on the dependent variable(s). In other words, a regression model may be used to predict the change in the dependent variable, given the changes in the independent variable. As used in this article, the dependent variables will be the Indicators (the proxies for changes in wrongful convictions), and the independent variables will be the introduction of BARD standard, and other changes in socioeconomic conditions that we will discuss below. In short, for simplicity and due to data constraints, we will rely on multiple linear regression models in this article to explain and model the linear relationship between the introduction of BARD and the Indicators, and assess whether the observed relations are statistically significant. *See* JEFFREY WOOLDRIDGE, INTRODUCTORY ECONOMETRICS: A MODERN APPROACH 68–77 (2006) (providing a detailed discussion on regression analysis).

[218] As an integral concept in hypothesis testing, statistical significance is a mathematical way to prove that a certain statistic is reliable. In this study, such statistics would be the relationship between the introduction of BARD and the various Indicators. Statistical significance builds on the presumption of normal distribution and reflects risk tolerance and confidence level. *See id.* at 121–35.

ii. Indicator 2.1: Number of criminal cases that were reversed and remanded

A wrongfully convicted case may be rectified in subsequent appellate proceedings, and judgments of criminal cases issued after 2013 in China may be appealed on the grounds of failing the reasonable doubt standard of proof.[219] If the BARD standard is working effectively as the legislature intended, our second hypothesis is that *we would expect the number and rate of criminal cases that are reversed and remanded (RR) for the reason of failing the standard of proof (SoP) to have increased after 2013, as compared to the immediately preceding period*. To evaluate this hypothesis, we first examined all criminal cases that were reversed and remanded with reasons attributable to a failed standard of proof between 2007 and 2018.[220] Then, among all such RR cases, we examined the number of cases that failed the traditional Chinese standard of proof ("reliable and sufficient evidence" (RS)) and, for cases decided after 2013, the number of cases that failed the BARD standard. If in the post-2013 period, the frequency of BARD being mentioned in RR criminal cases in China is higher than the frequency of RS being mentioned, and if there is an upward trend for a failed SoP as the general reason for RR criminal cases in China in the post-2013 period, then that would suggest that the BARD standard is an effective tool utilized by the Chinese appellate courts to scrutinize criminal trial judgments and grant reversal and remand. Otherwise, depending on further analysis, we could infer that the BARD standard did not add much value to the original criminal standard of proof in China.

iii. Indicator 2.2: Number of cases that were reversed and acquitted in China's "second instance trial"

Similar to *Indicator 2.1* and the hypothesis thereof, we examined how the BARD standard helps reduce wrongful convictions at the appellate level through appellate decisions of "reversal and acquittal," a unique feature in Chinese court settings.[221] While a civil or criminal appellate court in the

[219] *See* Criminal Procedure Law of the People's Republic of China, *supra* note 196, art. 53.
[220] Trial court judgment of a given case (civil or criminal) may be reversed and remanded for reasons other than a failure of the standard of proof. Such RR cases are irrelevant to our study and thus not in the scope of our analysis.
[221] Unlike any common law system, in which trials are particularly reliant on the testimony and cross-examination of witnesses to furnish to the judge and/or jury the relevant facts of the case, "the appearance rate of witnesses to orally testify at criminal trials in China is and has been extremely low. In keeping with common and civil law pre-trial preparation, it is

United States and other common law countries fully delegates the job of fact-finding to the trial court—and thus only grants a reversal and remand if it wants to reverse a trial court judgment—a Chinese civil or criminal appellate court (二审法院, the second instance trial court) has the authority to review both fact and law, and can reverse and directly acquit (RA) the accused if so justified,[222] like in a case with a failed SoP. Thus, while an appellate court in China may reverse and remand a criminal conviction to the lower court on the grounds that the evidence collectively provided by the procurator fails the BARD standard, the appellate court can also reverse and directly acquit the convicted for the same reason.[223]

If the BARD standard is working effectively, our third hypothesis is that *we would expect that the number of criminal cases that are reversed and acquitted at the appellate level for the reason of failing the SoP to increase after 2013, as compared to the immediately preceding period*. To test this hypothesis, we ran an examination similar to what we did for *Indicator 2.1*,[224] but for RA cases rather than RR cases. If in the post-2013 period, the frequency of BARD being mentioned in RA cases in China is higher than the frequency of RS being mentioned, and if there is an increase in the number of RA cases in China in this period as compared to the preceding one, then that again would strongly suggest that the BARD standard is an effective tool of the Chinese appellate courts in scrutinizing trial judgments and granting RA. Otherwise, depending on further analysis, we could infer that the BARD standard did not add much value to the original criminal standard of proof in China (RS).

iv. Indicator 3: Non-prosecution rate

At the beginning of a criminal proceeding, a prosecutor must evaluate the merits of the case and determine whether the evidence against the accused

---

common in China for witnesses to provide written statements at police stations or to procurators prior to trial. . . . Chinese judges decide guilt [or not] on these written witness statements . . . made pre-trial" as well as other documentary evidence. This tradition of written-evidence based criminal trials makes it possible for the Chinese appellate court to review the facts. *See* Wang & Caruso, *supra* note 204, at 52 (describing the role of witness testimony in Chinese criminal trials); *see also* An Ping, (安平), Nie Shubin Zaishen Gaipan Wuzui An (聂树斌再审改判无罪案) [Nie Shubin found not guilty at the last appellate proceeding], Renmin Fayuan Bao (人民法院报), [PEOPLE'S COURT DAILY], (Sept. 12, 2019), https://www.chinacourt.org/article/detail/2019/09/id/4473539.shtml.

[222] *See, e.g.*, *An Ping*, *supra* note 221 (stating that after a retrial "[t]he verdict of the original trial was revoked and [the defendant] was acquitted.").

[223] *Id.*

[224] *See supra* Part V(C)(ii).

is sufficient to warrant a meaningful trial.[225] The mere fact that the judge will explicitly impose the BARD standard presumably has a shadow-effect in deterring prosecution of any case that the prosecutor thinks cannot meet the standard of proof. Admittedly, many other factors (e.g., political considerations)[226] contribute to the non-prosecution rate, and no data can show the reasons behind each such decision. However, *ceteris paribus*, our fourth hypothesis is that *an effective BARD standard will contribute to an increase of the non-prosecution rate post-2013 in China*. To test this hypothesis, we observed changes in the nationwide non-prosecution rate as published by the *Annual Working Reports of the Supreme People's Procuratorat*e between 2007 and 2018,[227] checking whether a statistically significant change marks the number of non-prosecution cases after 2013. As with *Indicator 1*,[228] we also performed a regression analysis to observe whether the BARD standard contributes to these changes in a statistically significant way.

[225] Hannah Meinke, *Prosecution vs. Defense: A Discovery of the Differences*, RASMUSSEN UNIV. (Nov. 11, 2019), https://www.rasmussen.edu/degrees/justice-studies/blog/prosecution-vs-defense (stating that after someone is arrested, a prosecutor must "consider the burden of proof and their ability to successfully prosecute a case. . . . A prosecutor must have sufficient evidence to meet the standard of proof beyond a reasonable doubt used in criminal cases.").
[226] *See e.g.*, Matthew Boyle, *Grassley: Politics in US attorney's decision not to prosecute Holder?*, YAHOO NEWS (June 29, 2012), https://news.yahoo.com/news/grassley-politics-us-attorney-decision-not-prosecute-holder-210409449.html; Josh Gerstein, *Judge orders release of DOJ memo justifying not prosecuting Trump*, POLITICO (May 4, 2021), https://www.politico.com/news/2021/05/04/trump-obstruction-justice-doj-485360.
[227] The Annual Working Report of the Supreme People's Procuratorate, *supra* note 203.
[228] *See supra* Part V(C)(i).

D. Analysis: Application and implications of the Chinese judicial data

i. Reflection on Indicator 1: Impact of the reasonable doubt standard on acquittal rate

In Figure 4, we can observe the total number of acquittals as a percentage of all criminal cases terminated in China from 2007 through 2018. The annual acquittal rate diminished drastically over the period from 2007 through 2012, rebounded at a less drastic rate during the period from 2013 through 2017, and decreased again in 2018.

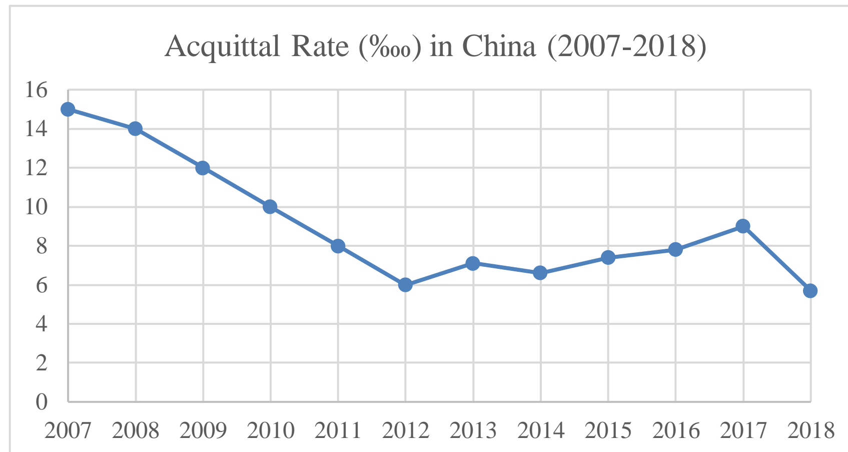

Figure 4

The trends in the acquittal rate in China in the post-2013 period send mixed signals. On the one hand, the upward trend we observed between 2013 and 2017 affirms our hypothesis that the BARD standard, if effective, will safeguard those cases where evidence is too weak to survive the BARD standard. The overall acquittal rate in Chinese criminal cases should increase after 2013, as compared to the immediately preceding period, which is exactly what happened. On the other hand, the acquittal rate suffered a sharp drop in 2018, not only neutralizing the cumulative gains from 2013–17 but also reaching a new low for the acquittal rate across the entire period of observation. Such a sudden reversal indicates either that the implementation of the BARD standard is ineffective or that it alone cannot explain changes in China's annual acquittal rate after 2013.

In order to test whether the BARD standard affected the observed trend in the acquittal rate in the post-2013 period, we designed a treatment

effect regression model, an analytic tool used in the field of econometrics.[229] Specifically, such a model allows us to examine whether the acquittal rate post-2013 is statistically different from that of the pre-2013 period.[230] If the answer is yes, we may draw a preliminary conclusion that BARD may have had some effect in stabilizing a downward trend in the acquittal rate in China since the early 2000s (i.e., the introduction of BARD helped rebound the acquittal rate). A simple treatment effect model is shown in Equation (i):

$$Acquittal\ Rate = \beta_0 + \alpha_{2013} * D + \epsilon;$$

In Equation (i), β_0 represents the average acquittal rate, and D is the dummy indicator[231] for the treatment study, i.e., D is a binary variable, with its value equal to 1 for the period after the promulgation of the BARD standard (2013–18), and its value equal to 0 for the period prior to the promulgation of the BARD standard (2007–12).[232] Furthermore, α_{2013} represents the treatment effect brought by the BARD standard. As an illustration, if α_{2013} equals 0, then the model would suggest that the acquittal rate for 2007–12 would be similar to that of 2013–18. On the other hand, if α_{2013} is negative, then for the years after 2013, we would expect an acquittal rate lower than it would otherwise be for years prior to 2013. Conversely, a positive α_{2013} would suggest that the average post-2013 acquittal rate should be higher than it

---

[229] As used in econometrics, the term "treatment effect" refers to the causal effect of a binary variable on an outcome variable of interest, in this case, the Indicators. Treatment effects are often estimated using social experiments, regression models, matching estimators, and instrumental variables. *See* Joshua D. Angrist, *Treatment Effects*, MIT ECONOMICS, http://economics.mit.edu/files/32 (last visited Oct. 4, 2021) (explaining the core principles of treatment effects); FUMIO HAYASHI, ECONOMETRICS 327–45 (2000) (showing a detailed discussion of fixed effect treatment models).

[230] Following standard practice in hypothesis testing, we will conclude that the two numbers are statistically different if we find the difference between the group means is statistically different from zero (with a pre-specified confidence interval) and thus reject the null hypothesis (H0) (the two groups are the same) and support the alternative hypothesis (HA) (the two groups are different). Also, note that if two numbers are not statistically different, it would not necessarily mean that they are the same.

[231] The "Dummy" indicator denotes the treatment effect we assigned to BARD. Intuitively speaking, it notes the period where China's criminal justice system has adopted the BARD standard. Thus, the value of the dummy indicator would be 0 for years prior to 2013 because there had been no BARD.

[232] If we have access to case specific data points, we may utilize alternative generalized linear regression models that may shed light as to how the BARD standard would affect the propensity of an outcome variable, such as logistic and probit regressions. However, for simplicity and because the outcome variables we study in this article are continuous (either as a percentage based on linear transformations of an aggregate number or as a non-binary percentage change), we rely on the multiple linear regression models assuming that all assumptions of ordinary least squares have been satisfied.

would otherwise be for the period prior to 2013. Given that the average acquittal rate for the first six years is close to 11‱, and the average for the second six years is only 7.5‱, it would be no surprise for the data to confirm the seemingly obvious speculation that the treatment effect of BARD, α_{2013}, may be negative.

However, the more important question here is how the BARD standard has disturbed the overall trend in the acquittal rate. A variant of the simple treatment effect model can be seen in Equation (ii):

$$\Delta Acquittal\ Rate = \beta_0 + \alpha_{2013} * D + \epsilon;$$

In Equation (ii), the dependent variable (our target variable) is now the change in the annual acquittal rate. The equation allows us to see whether the BARD standard affects the rate of change of the acquittal rate. As an illustration, if α_{2013} equals 0, then the model would suggest that the overall trend of change for 2007–12 would be similar to that of 2013–18. On the other hand, if α_{2013} is negative, then for years after 2013, we would expect the acquittal rate to decrease even more than it did in the years prior to 2013. On the other hand, a positive α_{2013} would suggest that changes in the acquittal rate after 2013 are less volatile (i.e., more stable) than those prior to 2013.

Table 1. Regression Results

Regression Statistics	Equation (i)	Equation (ii)
Multiple R	0.557	0.417
R Square	0.310	0.174
Adjusted R Square	0.241	0.091
Standard Error	2.641	0.211
Observations	12	12
Degree of Freedom		
Regression	1	1
Residual	10	10
Total	11	11
Coefficients		
Intercept	10.833*	-0.215*
P-Value	0.000	0.032
α_{2013}	-3.233	0.177
P-Value	0.060	0.177

* Denotes a result that is statistically significant on a 5% significance level.[233]

Table 1 exhibits the regression results for Equations (i) and (ii). Consistent with our observations of the overall data, we find a negative, statistically significant treatment effect for Equation (i), which suggests that the acquittal rates are statistically different across the treatment period (2013–18) and the control period (2007–12). However, as we progress to Equation (ii), we do not observe any statistically significant treatment effect brought by the BARD standard. In other words, we found no strong evidence through the above treatment effect regression analysis that the introduction of the BARD standard affected the overall trend, or rate of change, of the acquittal rate in China.

To further scrutinize these findings, we examined how the change in acquittal rate evolved across the two periods by considering more variables that may contribute to the overall changes in China's judicial system. As shown in Table 2, we introduced two quantifiable variables to our regression analysis: the percentage change in the Chinese GDP growth rate, and the percentage change in the overall crime rate in China. The Chinese GDP was introduced because almost all of the development we have witnessed in China has been more or less driven by its GDP growth.[234] For the latter, we sense that the overall crime rate has a complicated relationship with the acquittal rate. In essence, we ran three sets of "difference-in-difference" regressions, each as shown in Equation (iii):

$$\Delta\, Acquittal\ Rate = \beta_0 + \beta_1 X + \alpha_{2013} * D + \epsilon$$

In Equation (iii), X represents the variable that we used in the regression analysis (change in GDP, change in crime rate, and a combination of the two), and α_{2013} represents the treatment effect variable that may affect the overall independent variable, $\Delta\, Acquittal\ Rate$, after 2013, possibly due to the promulgation of the BARD standard. D is a binary variable that equals to 1 for 2013 and subsequent years but equals to 0 for years earlier than 2013. β_0 represents the average, and ϵ represents the universe of unobservable.

---

[233] As an illustration of hypothesis testing, the null hypothesis (H0) here involves a variable of zero, and the alternative hypothesis (HA) involves a variable that is not zero. On a significance level of 5%, the most popular choice in the field of econometrics, we are willing to mistakenly reject a true null hypothesis (H0) 5% of the time. *See* JEFFREY WOOLDRIDGE, INTRODUCTORY ECONOMETRICS: A MODERN APPROACH 124 (2013).

[234] *See generally* Xiaodong Zhu, *Understanding China's Growth: Past, Present, and Future*, 26 J. ECON. PERSP. 103 (2012) (providing a comprehensive review of China's unprecedented economic transformation since 1978).

Table 2. Percentage Change in Acquittal Rates in China

Year	Acquittal Rate (‱)	Δ Acquittal Rate (%)	Δ in GDP (%)[235]	Δ Overall Crime Rate (%)[236]
2007	15	-27%	11%	11%
2008	14	-7%	-47%	10%
2009	12	-17%	-3%	9%
2010	10	-20%	12%	0%
2011	8	-25%	-11%	8%
2012	6	-33%	-22%	14%
2013	7.1	15%	-1%	13%
2014	6.6	-8%	-6%	16%
2015	8.4	21%	-6%	0%
2016	8.8	5%	-3%	5%
2017	9	2%	0%	0%
2018	5.7	-58%	-3%	0%

Three sets of regression tests were undertaken: (1) the change in acquittal rate on the BARD treatment effect variable and change in crime rate; (2) the change in acquittal rate on the BARD treatment effect variable and the change in annual GPD growth rate; and (3) the change in acquittal rate on the treatment effect variable, the overall change in crime rate, and the change in annual GDP growth rate. Regression results, as summarized in Table 3, reveal that the treatment effect of the BARD standard is not statistically significant

[235] *See The World Bank National Accounts Data: GDP Growth (Annual %)—China (2000-2018)*, THE WORLD BANK, https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?end=2018&locations=CN&start=2000 (last visited Oct. 5, 2021).

[236] *See China Crime Rate & Statistics 1995-2020*, MACROTRENDS, https://www.macrotrends.net/countries/CHN/china/crime-rate-statistics (last visited October 4, 2021).

to the change in the overall acquittal rate. Thus, we cannot negate the null hypothesis (H0) that implementation of the BARD standard is not statistically significant to changes in the annual acquittal rate in China.

Table 3. Regression Analysis

Regression Statistics	***Test 1***	***Test 2***	***Test 3***
Multiple R	0.495	0.496	0.509
R Square	0.245	0.246	0.259
Adjusted R Square	0.094	0.095	0.012
Standard Error	0.212	0.211	0.221
Observations	13	13	13
Degree of Freedom			
Regression	2	2	3
Residual	10	10	9
Total	12	12	12
Coefficients			
Intercept	-0.284	-0.255	-0.284
P-value	*0.027*	*0.013*	*0.035*
□2013	0.210	0.209	0.214
P-value	*0.108*	*0.109*	*0.120*
□ *Crime Rate*	0.632		0.453
P-value	*0.548*		*0.701*
□ *GPD*		-0.251	-0.185
P-value		*0.539*	*0.687*

Thus, to conclude this part of analysis, we did not observe any statistically significant impact of the BARD standard's implementation on the annual acquittal rate in China.

ii. Reflection on Indicator 2.1: Impact of the reasonable doubt standard on criminal cases of reversal and remand

As discussed above, for the purposes of this study, we used CJO to research and collect data on Chinese appellate court decisions to reverse and remand trial judgments of criminal cases between 2007 and 2018.[237] We focused on the reversal and remand (RR) cases attributed to a failure of satisfying standard of proof (SoP) in criminal proceedings generally, and a failure of the BARD standard more specifically, as shown in Table 4.

CJO maintains a consistent case reporting of more than 50 percent of all criminal cases that have been terminated in China since 2014. However, CJO unfortunately contains only 5 percent of all criminal cases that terminated in China between 2007 and 2013, and less than 17 percent of criminal cases reported for the year 2013.[238] Due to such a significant absence of data reporting before 2013, we cannot compare an absolute number of items in Table 4 in the pre-2013 period with corresponding numbers in the post-2013 period. Nonetheless, we can compare and analyze the percentage changes in rates of the items before and after 2013, assuming cases of any given year reported by CJO were randomly selected. Also, because CJO's data reporting since 2014 is more comprehensive, a comparison of both the absolute numbers and their rates of change for any items between 2014 and 2018 (and analysis of data reported in the post-2013 period generally) makes empirical sense. With the admission that the collected data for the year of 2013 and before are incomplete,[239] our analysis may nevertheless provide a baseline estimation of how the standard of proof, in general, and the reasonable doubt standard, in particular, have been used as a pragmatic tool by Chinese criminal appellate courts in granting reversals and remand.

[237] *See* CHINA JUDGMENTS ONLINE, *supra* note 200.
[238] The CJO website was launched by the Supreme People's Court of China on July 1, 2013 and has been accessible to the public since then. *See Judicial Transparency of Chinese Courts*, *supra* note 200, art. 3. As of November 18, 2020, our data-collection cutoff date, CJO reported judgments of 180,128 Chinese criminal cases terminated in 2013, out of a total number of 1,059,752 criminal cases terminated that same year in China. Nonetheless, from the perspective of macro-statistics, even a randomly selected 1 percent of specimens out of a pool with seven-figure numbers would be considered significant in analyzing features of the whole pool.
[239] We also noticed that CJO is continuously improving its data reporting of any given year since 1996. The cutoff date of data collection in CJO for our study was November 18, 2020. Therefore, our article does not reflect data updates that followed this date.

Table 4. Reversal and Remand (Criminal Cases) in China
(based on data reported through CJO)[240]

Year	All Criminal Judgments reported	Reversal and Remand	SoP Related (generally)	SoP Related (generally) as Percentage	RS Related	RS Related as Percentage	BARD Related	BARD Related as Percentage
2007	2628	1	0	0	0	0	N/A	N/A
2008	5023	3	0	0	0	0	N/A	N/A
2009	4999	57	4	7.02%	4	7.02%	N/A	N/A
2010	14880	54	4	7.41%	4	7.41%	N/A	N/A
2011	30677	82	15	18.29%	15	18.29%	N/A	N/A
2012	80201	317	110	34.70%	110	34.70%	N/A	N/A
2013	239718	573	140	24.43%	136	23.73%	4	0.70%
2014	1251477	3889	769	19.77%	721	18.54%	38	0.98%
2015	1406818	5446	1078	19.79%	1001	18.38%	71	1.30%
2016	1624396	9405	1896	20.16%	1785	18.98%	98	1.04%
2017	1504758	10752	2335	21.72%	2138	19.88%	161	1.50%
2018	1547268	11422	2829	24.77%	2586	22.64%	200	1.75%

In Table 4, we can see that the percentage of RR cases attributed to a failed SoP has trended upward since 2013, as compared to the pre-2013 period, a sign that the BARD standard has been an effective tool of Chinese appellate courts to scrutinize criminal trial judgments and grant reversal and

[240] CJO is updating its database constantly. The cutoff time of our data collection is November 18, 2020. All "Percentage" rates are calculated by designated data divided by the total number of RR cases of the same year.

remand.[241] Additional evidence of an effective BARD is that the annual number of RR cases attributed to the BARD standard increased significantly after 2013.

However, our optimism on BARD takes a hit when our attention shifts to the number and percentage of RR cases attributed to a failure of the traditional Chinese standard of proof in criminal cases, "reliable and sufficient evidence" (证据确实充分标准) (RS). To our surprise, the statistics clearly show that RS had a dominant role in RR cases attributed to a failure of the SoP throughout the period of observation in China, regardless of whether the BARD standard had been implemented at the time. Also, to our surprise, a closer look at the post-2013 period shows that the frequency of RS being used by the Chinese appellate courts in granting reversal and remand in that period increased drastically, with a margin much larger than the increase in usage of BARD during the same period. Even though both the number and rate of BARD-related RR cases steadily increased after 2013, both its numbers and its percentages over the years are consistently smaller than those for RS-related cases. All these facts in turn indicate that the role of BARD in criminal RR cases in China is at best minimal.

Therefore, even though the BARD standard had a positive effect in Chinese appellate courts on the granting of reversals and remand, its role is minimal; even though the usage of BARD in RR cases increased steadily after its 2013 implementation, it was overshadowed by the magnitude of growth in usage of the traditional Chinese criminal standard of proof (RS) during this period. This situation suggests that the BARD standard not only is far from ready to replace the traditional Chinese criminal standard of proof but also only plays a marginal, supporting role in judicial fact-finding of criminal cases in China.

iii. Reflection on Indicator 2.2: Impact of the reasonable doubt standard on cases of reversal and direct acquittal

As discussed above, a unique practice in Chinese criminal proceedings allows the Chinese criminal appellate court to grant reversal and acquittal (RA) directly.[242] CJO currently provides the best available data on

[241] Noticeably, percentage changes between 2012 and 2013 in SoP-related RR cases are contrary to the general trend we have described. Nonetheless, we believe that this "noise" is probably caused by insufficient data reporting of CJO in these two years, which should not affect our overall judgment here on the trends in the post-2013 period, as compared with the pre-2013 period.

[242] *See* Wang & Caruso, *supra* note 204 at 57–58. For both civil and criminal cases, the Chinese appellate court can review both questions of law and questions of fact. Thus, in practice, the parties can experience a "trial" all over again in the appellate court in China. Thus, the appeals proceeding is called "the second instance trial" in China.

Chinese criminal appellate court RA decisions, both those attributed to a failure to satisfy the standard of proof in criminal proceedings generally, and those attributed to the failure of the BARD standard more specifically (see Table 5).

Table 5. Reversal and Acquittal (directly) in China
(based on data reported in CJO)[243]

Year	All Criminal Judgments Reported	Reversal and Acquittal (directly)	SoP Related (generally)	SoP Related (generally) as Percentage	RS Related	RS Related as Percentage	BARD Related	BARD Related as Percentage
2007	2628	0	0	N/A	0	N/A	N/A	N/A
2008	5023	0	0	N/A	0	N/A	N/A	N/A
2009	4999	0	0	N/A	0	N/A	N/A	N/A
2010	14880	0	0	N/A	0	N/A	N/A	N/A
2011	30677	0	0	N/A	0	N/A	N/A	N/A
2012	80201	5	0	0%	0	0%	N/A	N/A
2013	239718	6	2	33.33%	1	16.67%	1	16.67%

[243] CJO updates its database constantly. The cutoff time of our data collection was November 18, 2020. All percentage rates are calculated by designated data divided by the total number of RA cases of the same year.

2014	1251477	68	52	76.47%	37	54.41%	12	17.65%
2015	1406818	84	78	92.86%	60	71.43%	10	11.90%
2016	1624396	132	118	89.39%	99	75.00%	11	8.33%
2017	1504758	181	180	99.45%	147	81.12%	26	14.36%
2018	1547268	183	181	98.91%	153	83.61%	24	13.11%

Throughout the period of observation of our study, RA cases accounted for 0.01 percent or less of all criminal cases in China, a much smaller magnitude than RR cases. As discussed in our analysis of *Indicator 2.1*, significant absence of data reporting of CJO before 2013 makes it meaningless to compare absolute numbers of RA cases attributable to SoP, RS, and BARD in the pre-2013 period with corresponding numbers in the post-2013 period. Moreover—and unlike with *Indicator 2.1*—due to the fact that the reported data of RA cases in Table 5 during the entire pre-2013 period is proximate to zero, we also cannot compare and analyze changes in rate (percentage) of numbers before and after 2013. However, the data reporting after 2013 does allow us to compare and analyze the change in absolute numbers and rate of any items in Table 5 between 2013 and 2018.

From Table 5, we can see that after the BARD standard was implemented in Chinese criminal proceedings in 2013, both the total number of RA cases and the number of such cases attributed to a failure of the BARD standard steadily increased, a positive sign that BARD is an effective tool utilized by the Chinese appellate courts to scrutinize criminal trial judgments and grant reversal and acquittal directly. However, when our attention shifts to the number and percentage of RA cases attributed to the failure of the traditional Chinese standard of proof in criminal cases (RS), as compared to RA cases attributed to SoP in general and those attributed to BARD, we have to reevaluate our impression of the impact BARD made. As with *Indicator 2.1*, the RS standard appears to be more common in Chinese appellate court decisions to grant reversal and acquittal in the post-2013 period, relegating the BARD standard to a supporting role.

iv. Reflection on Indicator 3: Impact of the reasonable doubt standard on non-prosecutions

Unfortunately, no available data can shed light on the reasons behind individual decisions to not prosecute cases in China. We nevertheless would expect that the implementation of the BARD standard in 2013 would contribute to an increase in the non-prosecution rate in China due to the assumption that procurators would likely be deterred, at least to some extent, from pressing charges with weak evidence given that an assumed higher standard of proof now puts a heavier burden of proof on the shoulder of the procurator and thus a heightened scrutiny of evidence against the accused.

Table 6 reproduces data published by the Supreme People's Procuratorate of China in its *Annual Working Report* between 2007 and 2018 regarding non-prosecutions collected from all levels of the People's procuratorates in the country.[244] At the first glance, the total number of non-prosecuted cases and the percentage of cases that the procuratorates decided to not prosecute among all criminal cases appear to be on an upward trend in the post-2013 period, a positive sign that the reasonable doubt standard is effective in China.

Table 6. Non-Prosecution Rate in China

Years	Cases of Non-Prosecution	Non-Prosecution Rate (%)	Change in Non-Prosecution (%)
2007	27995	2.52	N/A
2008	29871	2.54	0.79%
2009	25576	2.20	-13.39%
2010	37468	3.16	43.64%
2011	39754	3.20	1.27%
2012	N/A	N/A	N/A
2013	67820	4.87	N/A
2014	75487	5.15	5.75%
2015	76565	5.22	1.36%
2016	N/A	N/A	N/A
2017	111878	6.15	N/A
2018	136970	7.49	21.79%

However, two concerns from Table 6 must be addressed. First, the upward trend of both number and rate of the non-prosecuted cases begins at

[244] *See The Annual Working Report of the Supreme People's Procuratorate*, *supra* note 203.

least as far back as 2007, the earliest year of our observation period, rather than appearing as a new occurrence after the BARD standard was implemented. This fact may suggest that the observed trend relates not to the implementation of BARD but to some other factor(s) in China that we have not yet identified. Second, if we look at percentage change in the non-prosecution cases over the years, we find that the year-to-year changes both before and after 2013 are rather random and volatile, which again may suggest that the BARD standard is irrelevant to such changes. In fact, we performed another treatment effect regression to see whether the year-to-year changes in non-prosecution rates were statistically different in the post-2013 period, and we failed to observe statistically significant evidence that the BARD standard contributes to the changes in non-prosecutions in China. The regression statistics are shown in Table 7.

Table 7. Regression Results

Regression Statistics	
Multiple R	0.045
R Square	0.002
Adjusted R Square	-0.198
Standard Error	0.203
Observations	7
Degree of Freedom	*df*
Regression	1
Residual	5
Total	6
Coefficients	
Intercept	0.081
	0.462
Dummy	0.016
	0.924

Since a People's procuratorate may decide against prosecution for a number of inherently unobservable reasons—including political considerations, policies, lack of evidence, and more[245]—we must conclude

[245] *See, e.g.*, Renmin Jiancha Yuan Dui Naxie Anjian Keyi Jueding Bu Qisu? (人民检察院对哪些案件可以决定不起诉？) [Which cases can the People's Procuratorate decide not to prosecute?], Zhongguo Renda Wang (中国人大网) [NATIONAL PEOPLE'S CONGRESS NET] (Dec. 17, 2000), http://www.npc.gov.cn/npc/c2279/200012/dbb697b0920344709101fd6db667e03f.shtml.

that even though the non-prosecution rate in China increased significantly after 2013, we do not yet have strong evidence that this change had anything to do with the implementation of the BARD standard. These findings affirm our earlier findings for *Indicators 2.1* and *2.2*, that the BARD standard is not frequently used by the Chinese courts, as well as our overall impressions from *Indicators 1*, *2.1*, and *2.2* that, even though the BARD standard has proven occasionally helpful for the Chinese criminal justice system in restraining wrongful convictions, its impact has been very limited so far and not very effective.

E. Limitations of the Chinese case study

By observing and analyzing three *Indicators* (acquittal rate, reversal rate, and non-prosecution rate) across criminal cases in China for the two six-year periods that immediately precede and follow January 1, 2013, the benchmark date of the BARD standard's implementation in China,[246] we were able to conclude that the BARD standard appears to be supportive in averting wrongful convictions in China, but that it does not have a convincing or immediate effect on minimizing wrongful convictions therein.

Our work may be the only empirical study of this issue in existence. That said, this case study of the BARD standard in China has certain limitations. There are four major concerns that we would have liked to better address. First, the empirical data we collected is not flawless. Even though we had a great amount of data available, the problem of lack of data that has troubled many previous researchers in conducting empirical studies on the BARD standard also inevitably exists in our research. The period of time from which we chose to collect data regarding the three *Indicators* from the Chinese criminal justice system is between 2007 and 2018, or the six years before and the six years after the *2012 Amendments to the Criminal Procedure Law* took into effect on January 1, 2013.[247] One may argue that our selection of the twelve-year study period may be inadequate to make any genuine judgment on the subject at issue. We agree that this is a legitimate concern, since a newly adopted rule may take longer than six years to make an observable impact in a nation's criminal justice system. However, with the numerous constraints above-mentioned, there appears to exist no other alternative period of study in Chinese history that would be more conducive to such observation.[248]

---

[246] Criminal Procedure Law of the People's Republic of China, *supra* note 196, art. 53.
[247] *Id.*
[248] As we discussed above, the twelve-year period (2007–18) is the most (if not the only) suitable window to observe the effectiveness of BARD in Chinese criminal proceedings, since the *Criminal Procedure Law of China* was amended again at the end of 2018 with substantial changes (new variables, e.g., a whole new mechanism of plea bargaining with

Second, even though we focused on the period from 2007 to 2018, we only had comprehensive data for the years of 2014 to 2018. The collected data of 2013 barely qualified for our analysis of *Indicators 2.1* and *2.2*, and data from earlier years was even sparser. CJO, the database we used for this research, is the only online, freely accessible judicial database to be authorized and operated by the Supreme People's Court of China, as well as the most comprehensive database available to the public. Unfortunately, CJO was only launched on July 1, 2013,[249] and it has so far been slow to upload case materials that predate its founding. Another potential issue with CJO is that even though it is the best available judicial database in China, and has a fairly large collection of judicial documents and other related materials, its collection of case judgments and related materials (even after 2013) is still incomplete, representing only a portion of all cases terminated each year in China, no matter civil or criminal.[250] One indication of this issue is the way that the CJO database is updated constantly, with new case materials for multiple years being uploaded into the system and made available to the public on a daily basis.[251] The problem is that we do not know what is still missing for the case data of each year, since the CJO website includes no explanation about the missing data or its algorithm in updating the database.[252]

Moreover, a prerequisite for this study is that the study's environment should remain *ceteris paribus*—other variables in the targeted criminal justice system before and after the benchmark should remain more or less in the same condition throughout the period of observation. As discussed above, we assumed that our case study on China satisfies such a requirement because, across the twelve-year period, the *Criminal Procedure Law of China* was only amended once, in 2012, when the reasonable doubt standard was added into the law.[253] However, in fact, all variables of any criminal justice system, including the one in China, are constantly changing. That is just the nature of our world. Also, factors outside of the criminal justice system may also affect the results of our study, making it even more complex. In this study, we conducted regression analyses on *Indicator 1* via the fixed effect treatment model in order to take into account other major variables such as China's GDP and the crime rate. We also performed a reduced form regression analysis on *Indicator 3*. However, in reality, our regression analyses may be too narrow

---

Chinese characteristics) being made to the Chinese criminal justice system. These changes remove the necessary condition of constancy, impairing our ability to continue a study into BARD's impact. *See* Criminal Procedure Law of the People's Republic of China, *supra* note 196.

[249] *See Judicial Transparency of Chinese* Courts, *supra* note 200, art. 3.

[250] *See* CHINA JUDGMENTS ONLINE, *supra* note 200.

[251] *Id.*

[252] *Id.*

[253] Criminal Procedure Law of the People's Republic of China, *supra* note 196, art. 53.

in scope to distill the impact of BARD. Relatedly, even if we were able to identify other important variables that may have affected our study, it would likely be difficult to quantify them in such a way that would allow for a follow-up treatment effect regression analysis.

Finally, we must reiterate that the three *Indicators* we designed are proxies for changes in the number of wrongful convictions. We do not have direct data on wrongful convictions, and the *Indicators* and methodology we used in this study may be improved as data becomes more available in the future.

VI. REFLECTION AND CONCLUSION

If our earlier historical and epistemological analyses of the BARD standard highlight as many unresolved questions as they do answers—pointing to the slipperiness of the BARD standard when trying to pin it down through definition or quantification—our latter empirical analysis is no less ambivalent. Our empirical study into the Chinese case shows that although the 2013 implementation of the BARD standard was helpful in restraining wrongful convictions, such an effect has been very limited. Even that limited finding helps shed new light on the inner workings of the BARD standard, as we will explain.

First, theoretically speaking, a comparison between Figures 1 and 2 in Part II[254] shows that when the vertical SoP line moves toward the right (i.e., when the standard of proof in criminal cases gets higher), its effect on suppressing wrongful convictions is stronger. Since BARD has been broadly recognized as a high standard of proof, it is expected to function as an effective tool in eliminating wrongful convictions. Nonetheless, when people say that BARD is a high standard of proof, they are speaking relatively: BARD is a higher standard of proof than, for example, the "preponderance of evidence" and "clear and convincing standard" of civil cases. Conversely, when it is compared with the degree of absolute certainty, BARD is clearly a lower standard. What lurks here is the first point we would like to illustrate with regard to the Chinese case: before the BARD standard was implemented in China in 2013, the standard of proof in criminal cases had been "reliable and sufficient evidence" (RS), which many Chinese judges, practitioners, and legal scholars consider to be a very high standard as well.[255] When we compare the BARD standard to the RS standard, we find that it is hard to tell whether one is clearly higher than the other. Furthermore, taking into account a local context in which Chinese legal culture has previously emphasized that

---

[254] *See supra* Part II(B).

[255] *See* Chen Ruihua (陈瑞华), Xingshi Susong Zhong De Zhengming Biaozhun (刑事诉讼中的证明标准) [Standard of Proof in Criminal Proceedings], 3 Suzhou Daxue Xuebao (Zhexue Shehui Kexue Ban) (苏州大学学报 (哲学社会科学版)) [JOURNAL OF SOOCHOW UNIVERSITY (PHILOSOPHY & SOCIAL SCIENCE EDITION)] 77–78, 191 (2013).

evidence presented at trial should reflect the objective truth of past events at issue,[256] we find good arguments that the RS standard in Chinese criminal cases "feels" even higher than the BARD standard.[257] If our understanding is correct here, then the implementation of the BARD standard in China in 2013 would not have an effect of shifting the SoP line toward the right (in Figures 1 and 2)[258] and thus would not have decreased wrongful convictions in China. This leaves the question of, if the RS standard is so high, then why did Chinese legislators adopt the BARD standard in the *2012 Amendments to the Criminal Procedure Law*?[259] We believe that they did so because "reliable and sufficient evidence" is too vague, even more so than "beyond a reasonable doubt." For many years, Chinese judges, lawyers, and scholars have complained that the old standard of proof in criminal cases is more of an aspirational motto than a workable solution.[260] Some Chinese legal experts even declared this ambiguous standard of proof to be a direct cause of the wrongful convictions exposed in China.[261] In the *2012 Amendments to the Criminal Procedure Law*, Chinese legislators were able to add the BARD standard into the standard of proof provision, Article 53.[262] The original standard remains in the law,[263] with both standards now co-existing in the

---

[256] *See generally* Fan Chongyi (樊崇义), Keguan Zhenshi Guanjian—Jianlun Xingshi Susong Zhengming Biaozhun (客观真实管见—兼论刑事诉讼证明标准) [Opinions on Objective Truth—Also on the Standard of Proof in Criminal Proceedings], 1 Zhongguo Faxue (中国法学) [CHINA LEGAL SCI.] 114 (2000).

[257] *See, e.g.*, Pan Zhiyong (潘志勇), Paichu Heli Huaiyi Bu Dengyu Zhengju Queshi Chongfen (排除合理怀疑不等于证据确实充分) [Proof Beyond a Reasonable Doubt Is Not Equal to Reliable and Sufficient Evidence], 8 Zhongguo Jiancha Guan (中国检察官) [CHINESE PROCURATORS] 56–57 (2015).

[258] *See supra* Part II(B).

[259] Criminal Procedure Law of the People's Republic of China, *supra* note 196, art. 53.

[260] *See generally* Xu Yang (徐阳), Woguo Xingshi Susong Zhengming Biaozhun Shiyong Guannian Zhi Sikao—Cong Zengqiang Ke Caozuo Xing Dao Zengqiang Caozuo Guocheng De Guifan Xing (我国刑事诉讼证明标准适用观念之思考——从增强可操作性到增强操作过程的规范性) [Thoughts on the Applicability of China's Criminal Standard of Proof – From Enhancing Operability to Enhancing the Standardization of Operating Process], 2 Fashang Yanjiu (法商研究) [STUD. L. & BUS.] 64 (2017).

[261] *See generally* Zhang Baosheng (张保生), Xingshi Cuoan Jiqi Jiucuo Zhidu De Zhengju Fenxi(刑事错案及其纠错制度的证据分析) [Evidentiary Analysis on Wrongful Convictions and the Correction Mechanism], 1 Zhongguo Faxue (中国法学) [CHINA LEGAL SCI.] 90 (2013); Zhang Zongliang (张宗亮), Woguo xingshi Cuoan Yanjiu Zongshu(我国刑事错案研究综述) [A Research Overview of Misjudged Criminal Cases in China], 5 Shandong Jiancha Xueyuan Xuebao (山东检察学院学报) [Journal of Shandong Police College] 20 (2013).

[262] Criminal Procedure Law of the People's Republic of China, *supra* note 196, art. 53.

[263] *Id*.

same provision.[264] We interpret this particular legislative arrangement as a reflection of the intent of the Chinese legislators: they wanted to use "beyond a reasonable doubt" to explain what "reliable and sufficient evidence" actually means. Unfortunately, as discussed in Part I of this article[265] (and perhaps unbeknown to those Chinese legislators), the reasonable doubt standard itself does not have a clear and settled explanation at all. Thus, because the BARD standard was simply added into the criminal procedure law without any further explanation or guidance, fact-finders in Chinese criminal cases still very likely get confused on what the standard of proof means and thus apply it inconsistently.

At the end of the day, we reflect that the RS standard and the BARD standard are effectively treated the same in practice. China's move to the BARD standard was a formal change only and did not alter the substance of its criminal justice practice. The formal SoP may not matter all that much. Whether we call it "beyond a reasonable doubt" or not, courts have some intuitive understanding of the high standard of proof required in the criminal context and are going to apply it accordingly. Words effectively cannot describe the standard adequately.

Based on our multifaceted analysis of the BARD standard, we surmise that, in the end, the reasonable doubt standard has no magic after all. The standard only works well in preventing innocent defendants from being wrongfully convicted when it is applied together with other "defendant-friendly" mechanisms. Over the years, criminal justice systems all over the world have developed numerous "defendant-friendly" principles and rules to protect innocent persons from being wrongfully convicted, including but not limited to the presumption of innocence,[266] the prosecutorial burden of proof,[267] robust evidence rules,[268] the Miranda warning,[269] the privilege against self-incrimination,[270] the Confrontation Clause,[271] the right to

---

[264] *Id.*
[265] *See supra* Part I(B).
[266] *See Model Charge: Burden of Proof Presumption of Innocence, Reasonable Doubt*, *supra* note 184.
[267] *Id.*
[268] *See e.g.*, *The Exclusionary Rule as a Symbol of the Rule of Law*, 67 SMU L. REV. 821, 821 (2014).
[269] *See Amdt. 5.3.2.2.3.2.2 Requirements of Miranda*, CONSTITUTION ANNOTATED LIBRARY OF CONGRESS, https://constitution.congress.gov/browse/essay/amdt5_3_2_2_3_2_2 (last visited Oct. 6, 2021).
[270] U.S. CONST. amend. V.
[271] U.S. CONST. amend. VI.

counsel,[272] and the "beyond a reasonable doubt" standard of proof.[273] All these mechanisms interrelate and, by working together, grow into an enormous safety network in common law countries for protecting innocent persons from wrongful convictions.

Conversely, the criminal justice system of China had to rebuild itself beginning in 1979 with the "Reform and Opening Up" policy, which came after the ten years of the ruinous Cultural Revolution.[274] Although judicial reforms in China have made significant progress over the years, most "defendant-friendly" mechanisms are still either missing or severely underdeveloped in China. For example, for many years the single most important goal of the criminal justice system in China was to combat and crackdown on crime.[275] Even nowadays, it is still debatable whether defendants are presumed innocent or presumed guilty in criminal procedures.[276] Evidence law is still a new subject, and only a few, fragmented evidentiary rules have been established in China.[277] Neither the Miranda warning, nor privilege against self-incrimination, nor anything like the Confrontation Clause exists in China. The adversary system and advocacy are still weak in Chinese criminal trials.[278] And the list goes on. The adoption of the BARD standard is definitely a positive sign, showing that China has started to turn toward establishing a "defendant-friendly" criminal justice system and protecting the fundamental human rights of criminal defendants. Yet, such a singular solution in the absence of an entire supporting mechanism seems too little to be useful. The criminal justice system nowadays in China

---

[272] *Id.*

[273] *See Model Charge: Burden of Proof Presumption of Innocence, Reasonable Doubt*, *supra* note 184.

[274] *See* RANDALL PEERENBOOM, CHINA'S LONG MARCH TOWARD RULE OF LAW 6–7 (2002); *see generally* Pitman B. Potter, *Review: Legal Reform in China: Institutions, Culture, and Selective Adaptation*, 29 LAW & SOCIAL INQUIRY 465 (2004).

[275] *See, e.g.*, Stanley Lubman, *China's Criminal Procedure Law: Good, Bad and Ugly*, THE WALL STREET JOURNAL (March 21, 2012), https://www.wsj.com/articles/BL-CJB-15476.

[276] *See, e.g.*, Huang Meiyan (黄梅艳), Youzui Tuiding De Sibian Zhimei (有罪推定的思辨之美) [The Critical Beauty of Presumption of Guilty], 12 (才智) [ABILITY & WISDOM] 5–6 (2011); Li Youzhong (李友忠), Lun Wuzui Tuiding Yu Youzui Tuiding (论无罪推定与有罪推定) [On Presumption of Innocence and Presumption of Guilty], 3 Yunnan Faxue (云南法学) [YUNNAN LEGAL SCI.] 37–41 (1995); Zhou Zunyou, *Presumed Guilty in China*, SOUTH CHINA MORNING POST (June 3, 2013), https://www.scmp.com/comment/insight-opinion/article/1252146/presumed-guilty-china.

[277] *See* Jia Li & Zhuhao Wang, *A Trail to Modernity: Observations on the New Developments of China's Evidence Legislation Movement in a Global Context*, 21 IND. J. OF GLOB. LEGAL STUD. 683, 683 (2014).

[278] *See, e.g.*, Robert Lancaster & Xiangshun Ding, *Addressing the Emergence of Advocacy in the Chinese Criminal Justice System: A Collaboration between a U.S. and a Chinese Law School*, 30 FORDHAM IN'L L.J. 356, 356–357 (2006).

is still largely hostile to the accused.[279] The BARD standard can only do so much in such an environment, without other "defendant-friendly" mechanisms one would find in a common law BARD ecosystem.

Finally, although we are hesitant to make this observation, there seems to be a cultural, or local, aspect to the standard of proof in criminal cases. What works well (and has such long precedent) in the common law system cannot necessarily be successfully transplanted to China: elements are lost in translation, lost to a lack of history and understanding, and lost to cultural differences. We see this pattern at work in our empirical section in the discussion of *Indicators 2.1* and *2.2*: the RS standard experienced a sudden surge of usage in the post-2013 period after the BARD standard was implemented, which cannot be explained except for cultural reasons.[280] Taking a step back, then, we find it understandable that people of each country are proud of inventions made by their own countrymen, including legal terms, and might not so easily adopt conceptualizations from other countries.

In summary, and to return to the enigmatic smile of Leonardo's *Mona Lisa*, this multifaceted engagement with the BARD standard reveals that the more we peel away the layers of mystery that surround BARD, the more questions and difficulties arise. Even so, we believe that this investigation—historical, epistemological, and especially empirical—offers an important step toward a more contextualized and rigorous understanding of the effect of BARD on restraining wrongful convictions.

---

[279] *See* Wang & Caruso, *supra* note 204, at 56, 62.
[280] *See supra* Part V(C)(ii)–(iii).