



Alabama Law Scholarly Commons

Working Papers

Faculty Scholarship

5-10-2010

Toward a General Theory of Standards of Proof

Fredrick E. Vars

University of Alabama - School of Law, fvars@law.ua.edu

Follow this and additional works at: https://scholarship.law.ua.edu/fac_working_papers

Recommended Citation

Fredrick E. Vars, *Toward a General Theory of Standards of Proof*, (2010).

Available at: https://scholarship.law.ua.edu/fac_working_papers/716

This Working Paper is brought to you for free and open access by the Faculty Scholarship at Alabama Law Scholarly Commons. It has been accepted for inclusion in Working Papers by an authorized administrator of Alabama Law Scholarly Commons.

THE UNIVERSITY OF
ALABAMA

SCHOOL OF LAW

**Toward a General Theory of
Standards of Proof**

Fredrick E. Vars

Forthcoming, Catholic University Law Review

This paper can be downloaded without charge from the Social
Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=1604065>

The draft of this article is currently undergoing cite checking and revision by the Catholic University Law Review and will be published in final format in Volume 60 Issue 1 of the Catholic University Law Review

Word Count: 18,285

Toward a General Theory of Standards of Proof

Fredrick E. Vars[†]

Abstract

Which standard of proof is best for a particular type of case? This deceptively simple question has been much discussed, but the current state of understanding is unsatisfactory. Statisticians posed a general answer; philosophers and others launched an assault on that answer; practically oriented scholars draw on both strains unsystematically; and courts generally offer little or no reasoning for their decisions. The goal of this article is to outline a systematic and complete justification for selecting one probabilistic standard of proof over another. By training a microscope on one small corner of the law---incapacity will contests---this article demonstrates the relevance of old factors, identifies several new factors, and integrates the factors into an approach that will hopefully guide future inquiry. One important implication is that the choice of proof standard will almost necessarily be tentative: too much is unknown or unknowable.

[†] Assistant Professor, University of Alabama School of Law. Thanks to Ian Ayres, Bill Brewbaker, Adam Cox, Shahar Dillbary, Heather Elliot, Caroline Harada, Paul Horwitz, Dan Joyner, Ron Krotoszynski, Grace Lee, Beth Milnikel, Mike Pardo, David Patton, Roger Perlstadt, Meredith Render, and Norman Stein, each of whom contributed at various points in this article's development. Thanks also to Dean Ken Randall and the University of Alabama Law School Foundation for generous research support.

Table of Contents

Introduction..... 3

I. QUESTION..... 6

II. ANSWER OUTLINE..... 8

III. ASSUMPTIONS..... 14

A. Uniform Distribution..... 14

B. Distribution Is Independent of Proof Standard..... 18

C. Preponderance = 0.5; Clear and Convincing Evidence = 0.75..... 20

D. Courts Correctly Gauge Probability of Incapacity..... 26

E. Wills Executed by Testators with Capacity Are Accurate ($G = 1$)..... 27

F. Will Studies Show How Well Intestacy Matches Testators' Intent ($I \leq 0.31$).... 29

G. Fallback Is Intestacy..... 31

H. Promoting Testator Intent Is Dichotomous..... 32

I. Promoting Testator Intent Is the Only Objective..... 34

J. Every Case Is Litigated to Judgment..... 36

IV. A BETTER ANSWER?..... 42

V. EXTENSIONS..... 45

Conclusion..... 49

Introduction

Did O.J. Simpson kill his ex-wife Nicole Brown Simpson and her friend Ronald Goldman? A criminal jury said no; a civil jury said yes.¹ These apparently inconsistent verdicts can be reconciled because the juries actually answered different questions. The issue before the criminal jury was whether O.J. was proved guilty beyond a reasonable doubt.² The civil jury in turn was asked merely whether it was more likely than not that O.J. killed Nicole and Ronald.³ The two juries together indicated that they believed O.J. probably did it, but there was room for reasonable doubt.⁴

Absolute certainty is generally unattainable in legal proceedings. As a result, triers of fact, like the O.J. juries, are given guidance about how to resolve uncertainty. A primary tool is the standard of proof: the level of confidence or type of evidence required to decide a case one way or another. One proposed method to calculate the optimal standard of proof is based solely on the utilities of the different possible trial outcomes. This method has been formalized in an equation that will be set forth in the body of this article.⁵ However, two points about this method should be emphasized at the outset: (1) the formula generates a probabilistic standard---for example, convict only if you're at least 90% sure the defendant is guilty; and (2) the formula is incomplete.

¹ B. Drummond Ayres Jr., *The Simpson Verdict: The Overview; Civil Jury Finds Simpson Liable In Pair Of Killings*, N.Y. TIMES A1 (Feb. 5, 1997).

² *Id.*

³ *Id.*

⁴ Whether such doubts are still reasonable is questionable after the publisher labeled O.J.'s aborted account of the murders, "If I Did It," a "confession." Julie Bosman, *Simpson Publisher Explains*, N.Y. TIMES A11 (Nov. 18, 2006).

⁵ Equation 4 (accompanying *infra* note 47).

There is a substantial body of literature criticizing the formula on both grounds.⁶ This article builds upon and adds to the literature arguing that the utility of trial outcomes alone cannot determine the optimal proof standard. For example, scholars have identified the strength of evidence, the accuracy of adjudicators, and the merits distribution as additional factors that must be considered.⁷ This article puts each of these considerations (and others) into a new framework for selecting proof standards.

Critics of the probabilistic approach will be much less happy with my treatment of the first issue (framing the standard in probabilistic terms). I adopt the probabilistic approach without a formal response to the many criticisms that have been leveled against it. My rationale is three-fold: (1) appellate courts have adopted the probabilistic approach; (2) the probabilistic approach allows for easier quantification than alternatives (like the explanatory approach⁸); and, most important, (3) I believe the implications of my analysis apply whichever approach to proof standards is adopted. Even the most vociferous critics of the probabilistic approach would concede that the utilities of trial outcomes are at least relevant to selecting a proof standard. So too should such critics recognize the relevance and directions of impact of the other factors this article identifies. Operationalizing my framework non-probabilistically would be difficult, but not impossible. Indeed, proponents of other approaches arguably must do so to take into account all the relevant factors.

⁶ E.g., Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 LAW & PHILO. 223 (2008); Michael DeKay, *The Difference between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 LAW & SOC. INQUIRY 95 (1996); Ronald J. Allen, *The Restoration of In re Winship*, 76 MICH. L. REV. 30, 47 n. (1977).

⁷ LARRY LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW: AN ESSAY IN LEGAL EPISTEMOLOGY 73 (2006).

⁸ E.g., Michael S. Pardo, *Second-Order Proof Rules*, 61 FLA. L. REV. 1083, 1104 (2009) (explanatory approach) (“A fact is proven by clear and convincing evidence when the explanation of the evidence and events in dispute that includes this fact is clearly and convincingly better than explanations that do not.”). A parallel probabilistic instruction might read: “A fact is proven by clear and convincing evidence when you’re at least 75% sure of its truth.”

The goal of this article is to outline a systematic and complete justification for selecting one probabilistic standard of proof over another as a model for future inquiry. Perhaps the most important conclusion of this article is that selecting a proof standard, absent data not realistically attainable, is necessarily tentative. Some of the factors that go into selecting a proof standard are affected by the choice of standard in ways that pure theory cannot predict. One might throw up one's hands at this point, but that would be a mistake. Courts and legislators must select standards under conditions of imperfect information. The outcomes of real cases hang in the balance. The analysis set forth in this article ought to guide the selection of proof standards, even though it may not fully determine the optimal choice. At a minimum, the systematic and complete approach set forth here helps to identify flawed reasoning on proof standards by prominent legal theorists like Richard Posner.

In particular, this article examines will contests alleging that the testator lacked the mental capacity to execute the will⁹---then illustrates how the analysis has general applications. Part I introduces the will contest context, asking a deceptively simple question: which standard of proof---preponderance or clear and convincing evidence---is better? Part II provides what I believe to be the first systematic outline of a mathematical answer to the question of deciding between two proof standards (in any context). Part III discusses at length the assumptions behind that outline and examines whether each assumption is valid and how it affects the conclusion, if at all. Part IV combines the insights of Parts II and III to discuss whether a revised

⁹ A definition is in order: "To have mental capacity to make a will, the law requires the testator to be 'of sound mind.' To be 'of sound mind,' the testator must, when executing a will, be capable of knowing and understanding in a general way the nature and extent of his or her property, the natural objects of his or her bounty, and the disposition that he or she is making of that property, and must also be capable of relating these elements to one another and forming an orderly desire regarding the disposition of the property." RESTATEMENT (THIRD) OF PROPERTY (WILLS & DON. TRANS.) § 8.1 cmt.c (2003) (hereinafter "RESTATEMENT").

answer is warranted and whether a definitive answer is achievable. Part V considers extensions of the analysis to proof standards in three other contexts.

I. QUESTION

At a dinner party Mrs. [Brooke] Astor had at her apartment in January 2002 for Kofi Annan, Dr. [Henry] Kissinger testified, Mrs. Astor leaned toward him and asked, "Who is the black fellow who is sitting on the other side of me?"

"Kofi Annan," he said he responded. And when she later asked if Mr. Annan was distinguished, Mr. Kissinger said he told her: "He is a very distinguished man. He is secretary general of the United Nations."¹⁰

This testimony came in a criminal case against Mrs. Astor's only son, Anthony Marshall, in which the state alleged, among other things, that Mr. Marshall took advantage of Mrs. Astor's "diminished mental capacity" by having her amend her will in his favor.¹¹ The Astor case was unusual not just in the star power of the witnesses. It is very rare for a will to be challenged in criminal court. In taking this approach, the state arguably had to prove incapacity beyond a reasonable doubt. In the simultaneously and still pending will contest,¹² the burden is on the proponent of the will to prove capacity by a preponderance of the evidence.¹³

New York's approach to the burden in capacity will contests is traditional, but the modern approach, followed by a majority of jurisdictions, is to place the burden to prove incapacity on the contestant.¹⁴ That is where the burden rested in the criminal case against Mr.

¹⁰ John Eligon, *At Astor Trial, Famous Faces Turn Heads, and Testify*, N.Y. TIMES A23 (May 22, 2009).

¹¹ *People v. Marshall*, Indictment, at <http://news.findlaw.com/cnn/docs/brookeastor/nymrshl11107ind.html> (visited June 19, 2009) (hereinafter "Indictment"). Mr. Marshall was convicted. John Eligon, *Mrs. Astor's Son Guilty of Taking Tens of Millions*, N.Y. TIMES A1 (Oct. 9, 2009).

¹² *In re Astor*, 2009 WL 1415846 (N.Y. App. Div. 2d Dept. 2009).

¹³ *In re Will of Buckten*, 578 N.Y.S.2d 754, 755 (App. Div. 4th Dept. 1991); accord A. G. Sulzberger, *Fate of the Astor Fortune Is Uncertain, Despite a Verdict*, N.Y. TIMES A25 (Oct. 9, 2009).

¹⁴ THOMAS E. ATKINSON, HANDBOOK OF THE LAW OF WILLS § 196, at 508 (1937) ("According to the orthodox view the proponent has the burden of proof as to mental capacity, though a substantial number of courts place this burden on contestant."); EUNICE L. ROSS & THOMAS J. REED, WILL CONTESTS § 6:14 (2d ed. 2008); see also U.P.C. § 3-407 ("Contestants of a will have the burden of establishing lack of testamentary intent or capacity")

Marshall. But will contests require a lower standard of proof than criminal cases.¹⁵ In will contests alleging testator incapacity, most jurisdictions require proof by a preponderance of the evidence.¹⁶ That standard, however, is not universal.

Some courts lay down a rule, which, in form at least, requires more than a preponderance. It has been said that the evidence, especially of insanity, need not be more than to a reasonable satisfaction, or that insanity must be established by the manifest weight of the evidence, or that the evidence must be clear, or the evidence must be strong, or that the evidence must be clear, definite, and weighty, or that there must be a great preponderance of the evidence of incapacity, or that the evidence must be clear, convincing and satisfactory, or the greatest and most satisfactory evidence of incapacity, or that the evidence must be conclusive, or cogent and convincing, or evidence like that which is required in a criminal case to rebut and to overcome the presumption of innocence.¹⁷

To this list one could add that in Kentucky the presumption of capacity “can only be rebutted by the strongest showing of incapacity.”¹⁸ Similarly, although California courts frequently reiterate that the standard is preponderance, the state Supreme Court has declared that “there is a *strong* presumption of competency.”¹⁹ So too Illinois seems to adhere to the preponderance test, but the Supreme Court reversed a jury verdict because the instruction required only a “slight

¹⁵ Louisiana for a time required proof beyond a reasonable doubt of testamentary incapacity. Succession of Lyons, 452 So. 2d 1161, 1164 (La. 1984).

¹⁶ 3 WILLIAM J. BOWE & DOUGLAS H. PARKER, PAGE ON THE LAW OF WILLS § 29.35, at 574 (2004) (“To sustain the burden of proof, a preponderance of the evidence is necessary.”) (citing case law from 33 jurisdictions); 95 C.J.S. § 40 (2008). *E.g.*, Looney v. Estate of Wade, 839 S.W.2d 531, 533 (Ark. 1992); Edwards v. Edwards, 520 So.2d 1370, 1372 (Miss.1988); In re Estate of Dion, 623 N.W.2d 720, 729 (N.D. 2001); Matter of Estate of Roosa, 753 P.2d 1028, 1032 (Wyo. 1988).

¹⁷ 3 BOWE & PARKER, *supra* note 16, § 29.35, at 577 (citing case law for a heightened general standard from nine jurisdictions) (footnotes omitted). Medical clinicians interpret the presumption of capacity as requiring “clear evidence” to the contrary. Thomas G. Gutheil, *Common Pitfalls in the Evaluation of Testamentary Capacity*, 35 J. AM. ACAD. PSYCHIATRY & LAW 514, 515 (2007).

¹⁸ Bye v. Mattingly, 975 S.W.2d 451, 455 (Ky. 1998); *see also* Keasler v. Estate of Keasler, 973 S.W.2d 213, 217 (Tenn. App. 1997) (“strong evidence”).

¹⁹ In re Flatau's Estate, 76 P.2d 506, 508 (Cal. 1938) (emphasis added).

preponderance”²⁰ and has elsewhere explained that “the evidence of incapacity must *clearly* preponderate to authorize the setting aside of the will.”²¹

A treatise describes allocation of the burden of proof in incapacity will contests as “a crazy quilt of apparently conflicting and confusing maxims and principles which vary from state to state in an astounding variety of verbal formulae.”²² Much the same can be said of the minority views of the requisite level of proof. However, the modal heightened standard appears to be clear and convincing evidence.²³ That, with slight variations in wording, is the standard in Louisiana,²⁴ New Jersey,²⁵ Pennsylvania,²⁶ Washington,²⁷ and Wisconsin.²⁸ Is the clear and convincing evidence standard better in this context than the preponderance test?

II. ANSWER OUTLINE

The first step in determining which proof standard is superior is to pick a criterion. Implementing the testator’s intent is the fundamental premise of the law of wills and the basis for the mental capacity requirement.²⁹ Thus, the question becomes which proof standard better effectuates the testator’s intent.

²⁰ Norton v. Clark, 97 N.E. 1079, 1083 (Ill. 1912).

²¹ Down v. Comstock, 149 N.E. 507, 513 (Ill. 1925) (emphasis added).

²² ROSS & REED, *supra* note 14, § 6:14 (2d ed. 2008).

²³ This standard applies in many other contexts as well. 2 MCCORMICK ON EVIDENCE § 339, at 488 (Kenneth S. Brown ed., 6th ed. 2006).

²⁴ Succession of Lyons, 452 So. 2d 1161, 1165-66 (La. 1984) (codified in La. C. C. Art. 1482(A)) (rejecting beyond a reasonable doubt standard).

²⁵ In re Hoover's Estate, 91 A.2d 155, 156 (N.J. Super. A.D. 1952).

²⁶ In re Cohen's Estate, 284 A.2d 754, 755 (Pa. 1971) (“clear, strong, and compelling evidence”).

²⁷ Pond's Estate v. Faust, 163 P. 753, 753 (Wash. 1917).

²⁸ Matter of Sorenson's Estate, 274 N.W.2d 694, 696 (Wis. 1979) (“clear, convincing and satisfactory evidence”).

One commentator has described imposition of the clear and convincing standard as “the trend,” Laurie Dearman Clark, Comment, *Louisiana's New Law On Capacity To Make And Receive Donations: “Unduly Influenced” By The Common Law?*, 67 TUL. L. REV. 183, 212 (1992), but I see no support for this assertion. Most of the jurisdictions applying the standard have done so for many years. E.g., *In re Hoover's Estate*, 91 A.2d at 156 (1952) (N.J.); *In re Lowrey's Estate*, 13 Pa. D. & C. 532, 535 (1930); *Pond's Estate*, 163 P. at 753 (1917) (Wash.); *In re Emerson's Will*, 198 N.W. 441, 443-44 (Wis. 1924); *see also In re Flatau's Estate*, 76 P.2d at 508 (suggestion of heightened standard from 1938) (Cal.); *Norton*, 97 N.E. at 1083 (same from 1912) (Ill.).

²⁹ RESTATEMENT, *supra* note 9, § 8.1 cmt.b.

The next step in selecting a proof standard is to define the two standards in common terms.

All burdens of persuasion deal with probabilities. The preponderance standard is a more-likely-than-not rule, under which the trier of fact rules for the plaintiff if it thinks the chance greater than 0.5 that the plaintiff is in the right. The reasonable doubt standard is much higher, perhaps 0.9 or better. The clear-and-convincing standard is somewhere in between.³⁰

But where exactly? A survey of 170 federal judges generated a mean, median, and mode of 0.75 for the clear and convincing standard.³¹

To quantify the effect of different standards, one must recognize that there are four possible outcomes of a fully litigated will contest alleging incapacity. The contestant can either win or lose, and the testator either had or did not have capacity. The following 2 x 2 box summarizes.

Table 1. Possible Incapacity Will Contest Outcomes

	<i>Will Contestant</i>	
Incapacity	<i>Loses</i>	<i>Wins</i>
Yes	False Negative	True Positive
No	True Negative	False Positive

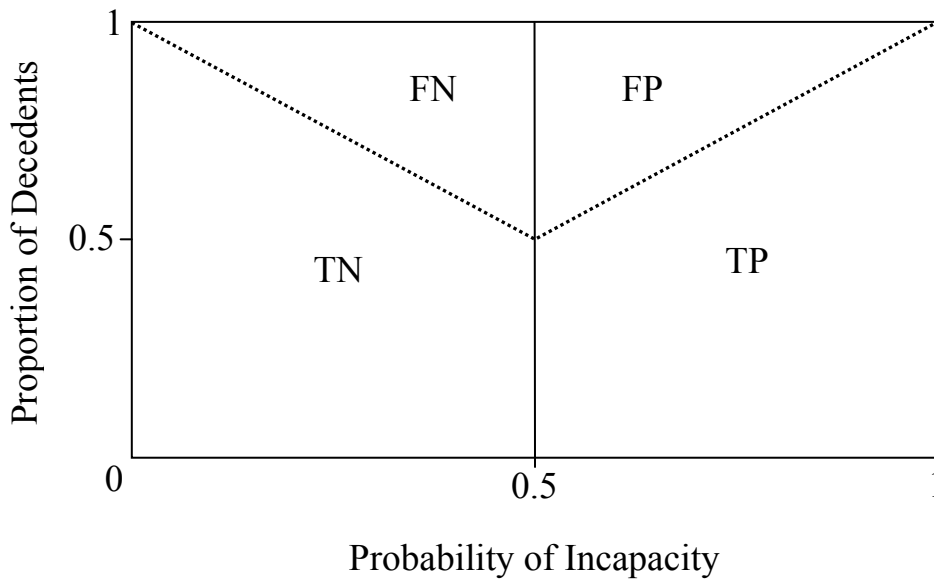
Raising the proof standard---making it more difficult to prove incapacity---will obviously lead to fewer findings of incapacity, thus generating more negative outcomes and fewer positive outcomes. In order to quantify this effect, we need to make further assumptions. Assume that

³⁰ *Brown v. Bowen*, 847 F.2d 342, 345-46 (7th Cir. 1988).

³¹ C.M.A. McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence or Constitutional Guarantees?*, 35 VAND. L. REV. 1293, 1328 tbl.5 (1982); see also *United States v. Fatico*, 458 F. Supp. 388, 410 tbl. (E.D.N.Y. 1978) (reporting a range of 0.6 to 0.75 in survey of eight federal district judges).

each testator falls somewhere on a continuum from 0 to 1, where the number reflects the probability based on admissible evidence that the testator lacked capacity. In the absence of data about the underlying rate of incapacity, assume a uniform distribution---in other words, that exactly the same number of testators fall at each spot along the continuum. Figure 1 represents the universe of decedents with wills under this distribution. Assuming trial accurately assesses the likelihood of incapacity and applies the preponderance standard, the vertical line at 0.5 separates successful and unsuccessful will contests and the V-shaped lines represent the number of correct contest outcomes. The resulting four regions correspond to the four possible outcomes depicted in Table 1.

Figure 1. Outcomes of Incapacity Will Contests Assuming Preponderance Standard and Uniform Distribution

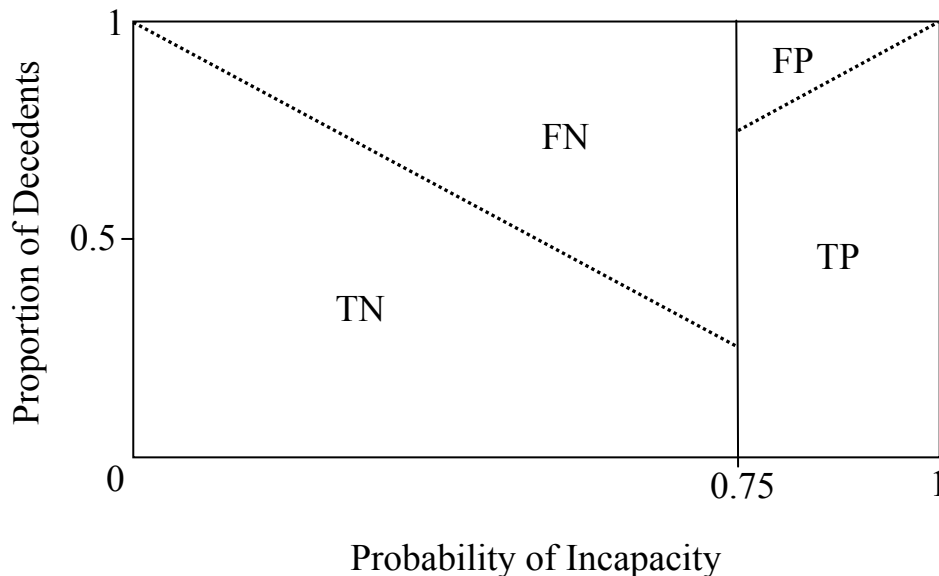


Why are the dotted lines V-shaped? Consider first the case where the probability of incapacity is zero (the left extreme of the figure). Because the probability of incapacity is less than 0.5, all such wills are probated and this was the correct result (a true negative) every time because every testator at this level had capacity. Hence, the dotted line begins in the upper left corner where probability of incapacity is zero and the proportion of decedents is one. Look

second at the break-point, where the probability of incapacity is 0.5. Half of decedents at this level lacked capacity. (Note that by assumption a trier of fact cannot tell which half---it can only observe the probability of incapacity.) Because there is a presumption of capacity, the wills of these testators on the bubble are upheld. Half of the time that was right (true negative); half of the time wrong (false negative). Thus, the dotted correct outcome line goes through the center of the figure, the 0.5/0.5 point. Finally, at the level where everyone lacks capacity, all wills will be thrown out, explaining the dotted line endpoint at the upper right corner of the figure (all true positives). The dotted line is linear because, by assumption, the error rate rises linearly up to the standard and falls linearly thereafter.

Applying the clear and convincing evidence test instead pushes the proof standard to 0.75. This leads to fewer findings of incapacity and therefore more true and false negatives. Figure 2 depicts the shift graphically.

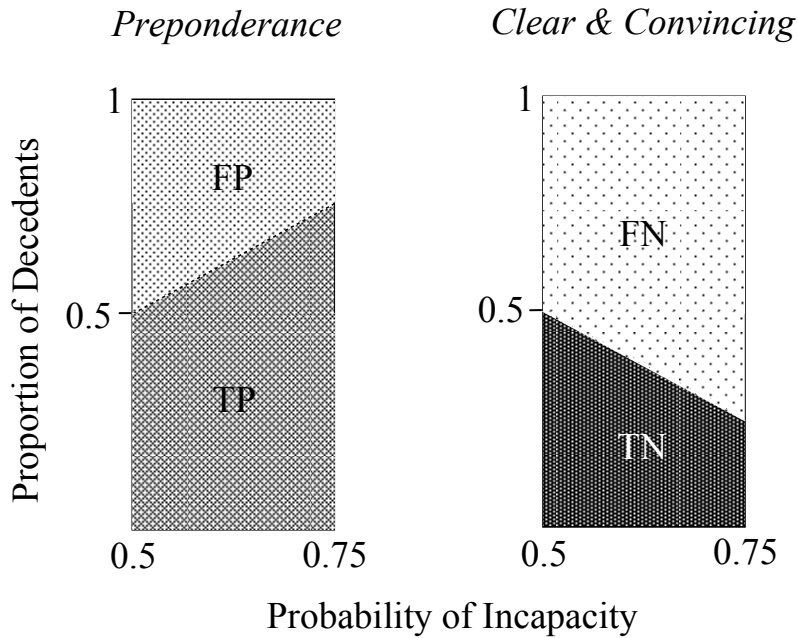
Figure 2. Outcomes of Incapacity Will Contests Assuming Clear and Convincing Standard and Uniform Distribution



From Figures 1 and 2 it should be apparent that the only difference in outcomes between the preponderance and clear and convincing evidence standards is in the region between 0.5 and

0.75 likelihood of incapacity. Wills by such testators are thrown out under the preponderance standard, but survive under the clear and convincing standard. A mix of true and false positives under the preponderance standard is replaced by a mix of true and false negatives (see Figure 3).

Figure 3. Preponderance and Clear and Convincing Evidence Standards Compared



If the goal were minimizing the number of errors, the preponderance standard would clearly prevail. There are far fewer false positives under the preponderance standard than there are false negatives under the clear and convincing evidence test. Geometry shows that the total error rate under preponderance is 0.25; under clear and convincing, it is 0.31. Error minimization, however, is *not* the goal; as stated earlier, advancing testator intent is. The question then is whether the mix of true and false negatives under the clear and convincing evidence standard better advances testator intent than would the true and false positives the preponderance standard would impose. If P is the proportion of decedents and U is utility (i.e.,

the probability that the distribution matches testator intent), then the clear and convincing evidence standard dominates if

$$(1) \quad P_{TN} \times U_{TN} + P_{FN} \times U_{FN} > P_{TP} \times U_{TP} + P_{FP} \times U_{FP}$$

This equation follows Figure 3 by including only the region between 0.5 and 0.75 probability of incapacity.

The right-hand side of the equation simplifies. Both true positives and false positives result in the estate being distributed via the fallback scheme, generally intestacy. Let I equal the likelihood that intestacy matches testator intent. The intestate distribution is independent of what the will says or whether the testator lacked capacity. It follows that $I = U_{TP} = U_{FP}$, so the right side reduces to $I \times (P_{TP} + P_{FP})$. It is not necessary to calculate the individual values of P_{TP} and P_{FP} , because geometry shows the sum to be 0.25.

The left-hand side of the equation is more difficult, because there is good reason to think $U_{TN} \neq U_{FN}$. A true negative affirms a will that was executed by a testator who had capacity; a false negative results in probate of a will that was executed by a testator who lacked capacity. For ease of exposition, let G (for “good” will) = U_{TN} , and B (for “bad” will) = U_{FN} . Substituting new variable names, doing mathematics relegated to a footnote,³² and dividing both sides of the equation by 0.25 lead to the following revised condition for clear and convincing superiority:

$$(2) \quad 0.375 \times G + 0.625 \times B > I$$

This equation shows that the values of G , B , and I determine which standard dominates. What do we know about each value? Barring other defects, it seems safe to assume that a will executed by a testator with capacity accurately reflects the testator’s intent. In symbols, $G = 1$. B is difficult to estimate. Bad wills can be divided into two categories: honest mistakes and

³² $P_{TN} = \int_{x=0.5}^{0.75} (1-x)dx$. $P_{FN} = 0.25 - P_{TN}$.

intentional overreaching. In the latter category it seems safe to assume that the resulting will never (or almost never) effectuate testator's intent. On the other hand, at least some portion of honest mistakes may stumble upon the testator's true plan---by, for example, chance, adherence to prior statements, or the guiding hand of a knowledgeable lawyer or friend. Thus, a best guess for B is that it is quite low, but greater than zero.

Fortunately, there are extensive data on I . Four studies of probated wills estimate I at 0.28,³³ 0.31,³⁴ 0.23,³⁵ and 0.20.³⁶ These low values decide the question. Even if bad wills never effectuate testators' intent ($B = 0$), the effect of the clear and convincing standard upholding more good wills (0.375) is greater than the accuracy of intestacy under the preponderance standard (highest estimate of 0.31). The five states that have adopted the clear and convincing evidence standard have apparently made the right choice.

III. ASSUMPTIONS

A. Uniform Distribution

The true distribution of incapacity among testators is almost certainly not uniform. Such a distribution implies an overall incapacity rate of 0.50. Common sense dictates that the true rate must be much lower. In other words, the distribution is skewed to the left. But can we be more specific?

One small study of incapacity will contests found the most common comorbid medical conditions to be dementia syndrome (0.40), alcohol related (0.28), neurological or psychiatric

³³ Frederick R. Schneider, *A Kentucky Study of Will Provisions: Implications for Intestate Succession Law*, 13 N. KY. L. REV. 409 (1987). Due to incomplete reporting in the article, the 0.28 figure is a best estimate. The possible range of values is 0.25 to 0.35.

³⁴ MARVIN B. SUSSMAN, JUDITH N. CATES, & DAVID T. SMITH, *THE FAMILY AND INHERITANCE* (1970).

³⁵ Allison Dunham, *The Method, Process and Frequency of Wealth Transmission at Death*, 30 U. CHI. L. REV. 242, 252 n.20 (1963).

³⁶ Edward H. Ward & J.H. Beuscher, *The Inheritance Process in Wisconsin*, 1950 WIS. L. REV. 393, 414 tbl.10.

disorder (0.28), personality disorder (0.20), and suicide (0.12).³⁷ Based on prevalence estimates and testation and mortality statistics,³⁸ I estimate that a maximum of 45% of testate deaths involved suicide or a testator with one of these conditions.³⁹ This is a nearly absolute upper bound on incapacity among testators. The true number must be vastly smaller. Many people who have mild dementia,⁴⁰ an alcohol problem, or an anxiety disorder, for example, are perfectly capable of executing a will. Furthermore, wills are generally drafted by attorneys and witnessed by at least two people.⁴¹ Gross cases of incapacity are unlikely to go unnoticed by all of these participants.

Another way to get at testamentary incapacity is by analogy to other areas. A great deal of research has been done on capacity to consent to medical treatment. That work “may be applicable to other complex capacities such as testamentary capacity.”⁴² The levels of incapacity

³⁷ Kenneth I. Shulman, Carole A. Cohen, & Ian Hull, *Psychiatric Issues in Retrospective Challenges of Testamentary Capacity*, 20 INT’L J. GERIATRIC PSYCHIATRY 63, 67 tbl.1 (2005) (n = 25; apparently Canadian data). See generally Aleksandar A. Jovanović et al., *Medical Reasons for Retrospective Challenges of Testamentary Capacity*, 20 PSYCHIATRIA DANUBINA 485, 490 (2008) (Serbian sample of contests finding statistically significant association between diagnoses and testamentary incapacity).

³⁸ Brenda L. Plassman et al., *Prevalence of Dementia in the United States: The Aging, Demographics, and Memory Study*, 29 NEUROEPIDEMIOLOGY 125, 125 (2007) (0.139 prevalence of dementia for population 71 years and older); Ronald C. Kessler et al., *Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication*, 62 ARCH. GEN. PSYCHIATRY 617, 620 tbl.1 (2005) (0.156 prevalence of severe and moderate DSM-IV disorders, including alcohol abuse and dependence); B.F. Grant et al., *Prevalence, Correlates, and Disability of Personality Disorders in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions*, 65 J. CLIN. PSYCHIATRY 948, 948 (2004) (0.148 prevalence of personality disorders); Suicide, Facts at a Glance, Summer 2008, at www.cdc.gov/ncipc/dvp/suicide/suicide_data_sheet.pdf (visited June 16, 2009) (0.013 suicide rate) (hereinafter “Suicide”); Robert A. Stein & Ian G. Fierstein, *The Demography of Probate Administration*, 15 U. BALT. L. REV. 54 (1985) (“On the average, testate decedents executed a will approximately five to seven years prior to their deaths.”); Dunham, *supra* note 35, at 279 tbl.15 (average age of will was 4.9 years). Based on these last two statistics, I estimated dementia rates five years before death.

³⁹ This estimate assumes that these conditions are independent. They are not. E.g., Jane E. Brody, *An Emotional Hair Trigger, Often Misread*, N.Y. TIMES D7 (June 16, 2009) (reporting that ten percent of borderline personality disorder patients die by suicide; recall that the overall suicide rate is 1.3%, see Suicide, *supra* note 38). This is another reason, see text *infra*, that the true number must be lower than 0.45.

⁴⁰ See Fayaz Roked & Abdul Patel, *Which Aspects of Cognitive Function Are Best Associated with Testamentary Capacity in Patients with Alzheimer’s Disease?*, 23 INT’L J. GERIATRIC PSYCHIATRY 552, 553 tbl.1 (2007) (finding that 92.6% of Alzheimer’s patients with mild, and 51.9% of those with moderate, cognitive impairment had testamentary capacity).

⁴¹ Cf. Jovanović et al., *supra* note 37, at 490 (contested wills were primarily holographic).

⁴² Shulman, Cohen, & Hull, *supra* note 37, at 67 tbl.1 (n=25; apparently Canadian data).

to make medical decisions observed among control groups range from 0 to 0.18, and are generally less than 0.09.⁴³ It is tempting to conclude that testamentary capacity may be at similarly low levels. However, as the 0 to 0.18 range reflects, capacity is task specific, so it is impossible to draw strong conclusions about testamentary capacity from medical capacity findings. Still, this research is broadly suggestive that the true level of testamentary incapacity is quite low.

Another imperfect approach to the question is to look at studies of will contests and settlements. In one large wills study, there were will contests in 0.01 cases and redistributions in 0.14.⁴⁴ Of course, there are many reasons why parties would decline to file a meritorious will contest and why settlement negotiations would break down, but even if we assume every contest and every redistribution was motivated by the testator's incapacity, incapacity affected the distribution in at most 0.15 of the testate cases studied.

To examine whether a skew in the distribution affects the conclusion that the clear and convincing standard dominates, assume a simple linear distribution ($y = 1 - x$). This distribution implies a 16.7% incapacity rate. Calculus relegated to a footnote⁴⁵ shows how Equation 2 is impacted. The result is that the clear and convincing standard remains superior if

$$(3) \quad 0.389 \times G + 0.611 \times B > I$$

⁴³ Dilip V. Jeste et al., *Magnitude of Impairment in Decisional Capacity in People with Schizophrenia Compared to Normal Subjects: An Overview*, 32 SCHIZOPHRENIA BULL. 121, 126 (2005).

⁴⁴ SUSSMAN, CATES, & SMITH, *supra* note 34, at 125, 184.

⁴⁵ $P_{TN} = \frac{\int_{x=0.5}^{0.75} (1-x)(1-x) dx}{\int_{x=0.5} (1-x) dx} = 0.389$

In other words, the case for the heightened proof standard becomes stronger as the distribution is skewed toward capacity---in general, as the likelihood of incapacity declines. This is consistent with the reasoning of some courts.⁴⁶

Still, this result may be surprising to those familiar with the standard statistical approach to selecting the optimal proof threshold. According to this approach, the optimal proof level does not depend on the underlying distribution. Rather, it is purely a function of the utilities of the four possible outcomes. Specifically,

$$(4) \quad P^* = \frac{1}{\left(\frac{U_{TP} - U_{FN}}{U_{TN} - U_{FP}}\right) + 1} \quad 47$$

⁴⁶ See, e.g., *General Motors Corp. v. Toyota Motor Co.*, 467 F. Supp. 1142, 1173 (D. Ohio 1979) (“[I]f it is unlikely that a type of allegation can be supported, clear and convincing evidence will be required to meet the burden of persuasion.”), *rev’d in part on other grounds*, 667 F.2d 504 (6th Cir.); see also 2 MCCORMICK ON EVIDENCE, *supra* note 23, at 488 n.25 (suggesting disapproval of this approach).

⁴⁷ Michael L. DeKay, *The Difference between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 LAW & SOCIAL INQUIRY 95, 111 eq.17 (1996); Alan D. Cullison, *Probability Analysis of Judicial Fact-Finding: A Preliminary Outline of the Subjective Approach*, 1 U. TOL. L. REV. 539, 569 eq.10b (1969); John Kaplan, *Decision Theory and Reasonable Doubt*, in COMMUNICATION SCIENCES AND THE LAW: REFLECTIONS FROM THE JURIMETRICS CONFERENCE (Layman E. Allen & Mary E. Caldwell eds., 1965).

Some wrongly accuse Professor Cullison of failing to consider the utilities of accurate results. Erik Lillquist, *Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability*, 36 U.C. DAVIS L. REV. 85, 107-08 (2002). But Cullison expressly measures the costs of Type I and Type II errors relative to accurate results.

Professor Tribe levels a more potent but still unpersuasive attack on the statistical approach and, by implication, the mathematical method of the present article. Laurence H. Tribe, *Trial By Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971). First, Tribe argues that the four utilities depend on the particular case. “[T]he trier might justly regard as worse the erroneous conviction of a man to whose guilt he had attached a probability [just above the decision standard] than the erroneous conviction of one whose guilt had seemed to be virtually certain.” *Id.* at 1382. That accords with human psychology, but misses the point. The question is not how the trier evaluates utilities, but how society should. The strength of the case against him makes no difference to a falsely convicted individual. More important for the purposes of this paper, testator intent is a more objective and reliable yardstick than the swirl of competing values at stake in the criminal justice system. See *infra* text accompanying notes 73-76. Second, Tribe contends that the trier’s utilities will cloud her perceptions of the evidence. *Id.* at 1383-84. As with the first criticism, this is an argument against allowing the trier to set the proof standard in each case, not an argument against having one. Indeed, Tribe recognizes this and argues that the standard should be set at the institutional level. *Id.* at 1384-85. But that does not satisfy Tribe. The formula itself fails to account for variables that Tribe thinks are essential, including how many false convictions and erroneous acquittals there are likely to be. This is precisely why the statistical model is useful: it generates an answer without needing information that is nearly impossible to gather. It can do so because it focuses on the marginal effects of each type of outcome. See *infra* text accompanying note 48.

Substituting into this equation 1 for U_{TN} (G), 0.31 for U_{TP} and U_{FP} (I), and 0 for U_{FN} (B) generates an optimal proof standard of 0.69. By selecting the high estimate of I and low boundary of B , 0.69 represents a low-end estimate.

If the optimal proof standard is independent of the merits distribution, why isn't the choice between the preponderance and clear and convincing standards similarly independent? The short answer is that the optimal standard formula considers only changes at the margin, whereas comparing two standards requires calculating areas between the two standards.⁴⁸ This article has already demonstrated that skewing the distribution toward capacity makes the clear and convincing standard relatively more appealing. Alternative distributions that tilt the scale in favor of preponderance are theoretically possible. Such a distribution would have to be skewed toward incapacity in the 0.5 to 0.75 range. Common sense and the data considered above, however, suggest that this is extremely unlikely. The percentage of testators who are incapacitated is almost certainly less than fifty and there is no reason to think this leftward skew would somehow be reversed in the 0.5 to 0.75 range.

In sum, the uniform distribution assumption is unlikely to be true, but deviation in the overwhelmingly more likely direction actually strengthens the case for the clear and convincing evidence standard.

B. Distribution Is Independent of Proof Standard

The simple answer assumes that the area under the total curve between 0.5 and 0.75 is not affected by selection of the proof standard. Economic theory suggests that this assumption may

⁴⁸ D.H. Kaye has argued that the standard statistical approach minimizes *expected* errors rather than *actual* errors. D.H. Kaye, *Clarifying the Burden of Persuasion: What Bayesian Decision Rules Do and Do Not Do*, 3 INT'L J. EVID. & PROOF 1 (1999). Kaye is correct only if the expected merits distribution is symmetrical around the proof standard, which case selection theory would predict, *see infra*, but which is not necessarily true. *Cf.* RICHARD A. EPSTEIN, *FORBIDDEN GROUNDS: THE CASE AGAINST EMPLOYMENT DISCRIMINATION LAWS* 224 (1992) ("It is critical to know not only the *probability* of Type I and Type II error but also the severity of loss associated with each type.").

be false. Switching to the clear and convincing standard would make wills more robust and therefore more attractive. More people would execute wills as a result, shifting the distribution up throughout the range. But there is a countervailing effect: the greater security a will provides against incapacity would lead some people to delay executing a will until later in life. This is because the risk of dementia would loom less large under the clear and convincing standard than it did under the preponderance standard.⁴⁹ Theory cannot predict which of these two effects would dominate. It would seem very unlikely that either effect would be substantial. Few people considering making a will know about the proof standard for incapacity will contests.

An evidentiary effect may be more potent. The simple model classifies individuals based on the admissible evidence of their capacity. Raising the proof threshold will lead contestants to produce more evidence of incapacity. This will tend to shift rightward the distribution between 0.5 and 0.75. Because the incentive to produce evidence is strongest near the decision standard, there would also likely be a skew toward incapacity. As discussed above (Section III.A *supra*), such a skew would favor the preponderance standard and, if substantial enough, could even tip the scale in that direction. But is it plausible to think such a substantial skew is likely?

The evidentiary effect assumes will contestants under a preponderance regime are not presenting as much evidence of incapacity as they could. But strategic and practical considerations weigh against this. Setting to one side the burden of proof, the preponderance standard has just the reverse effect on will proponents. They will need to produce more evidence to rebut a showing of incapacity.⁵⁰ The strategic contestant anticipates this incentive and responds by increasing her own evidentiary showing. In other words, as long as the parties do

⁴⁹ Cf. Daniel L. Rubinfeld & David E.M. Sappington, *Efficient Awards and Standards of Proof in Judicial Proceedings*, 18 RAND J. ECON. 308, 314 (1987) (explaining that while reducing the penalty for crime can increase accuracy, such a reduction may encourage crimes to be committed).

⁵⁰ Cf. Vern R. Walker, *Preponderance, Probability and Warranted Factfinding*, 62 BROOK. L. REV. 1075, 1115 (1996).

not know the outcome in advance, both will have strong incentives to produce a great deal of evidence under either proof standard. The distribution ultimately rests on the complete package of evidence, not just the evidence presented by the contestant.⁵¹ Defendants in will contests cannot just sit on their hands for another reason as well: they must be concerned with the possibility of judgment as a matter of law.⁵² And, practically speaking, attorneys concerned about malpractice liability and paid either by the hour or on contingency will be highly motivated to root out most probative evidence. Still, the evidentiary effect cannot be wholly discounted and pushes toward the preponderance standard.

C. Preponderance = 0.5; Clear and Convincing Evidence = 0.75

This assumption has both descriptive and normative components. Descriptively, as reported above, a large survey of judges found a mean, median, and mode of 0.75 for the clear and convincing evidence standard.⁵³ This is strong evidence, but it obscures the fact that 65% of judges picked a level other than 0.75 and the responses ranged from 0.5 to 1 (see Figure 4).⁵⁴

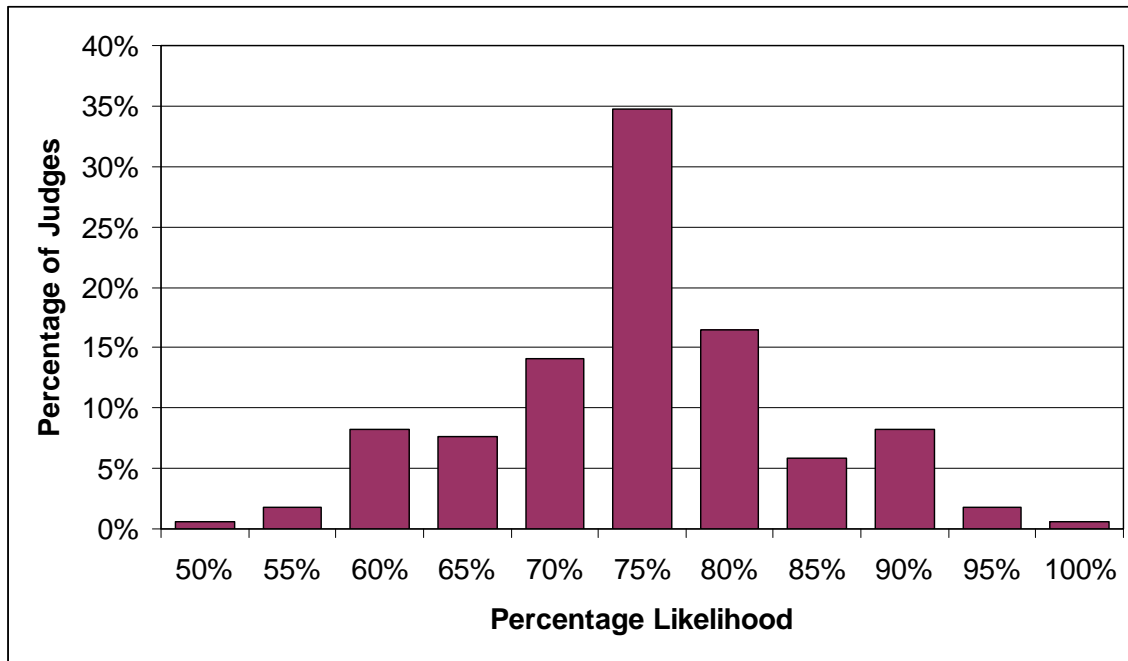
⁵¹ See Mike Redmayne, *Standards of Proof in Civil Litigation*, 62 MODERN L. REV. 167, 181-82 (1999).

⁵² See Walker, *supra* note 50, at 1115 (“The three standards of proof employed in litigation all help to create an incentive to produce more evidence than is sufficient to avoid suffering judgment as a matter of law.”).

⁵³ McCauliff, *supra* note 31, at 1328 tbl.5. Many courts, it should be noted, resist quantification. *E.g.*, Massachusetts Inst. of Tech. v. Harman Int’l Indus., 584 F. Supp. 2d 297, 308 (D. Mass. 2008) (describing clear and convincing standard as “vague and impressionistic”); Lockard v. Carson, 287 N.W.2d 871, 874 (Iowa 1980) (“‘clear and satisfactory’ refers to the Character or nature of the evidence, whereas ‘preponderance’ of the evidence is a Quantitative measure”).

⁵⁴ McCauliff, *supra* note 31, at 1328 tbl.5.

Figure 4. Distribution of Judges' Percentages Associated with Clear and Convincing Evidence



Judges were much more consistent in describing preponderance as at or just above 0.5: 88% picked either 0.5 or 0.6.⁵⁵ One problem with these numbers is that juries, not judges, decide many will contests at the trial level.⁵⁶ (Judges, of course, decide motions and appeals, so their perception of the standard is significant even in jury trials.⁵⁷)

Laypeople do not share judges' consensus regarding the preponderance standard. One study reported that students' and jurors' means and medians hover around 0.75, not 0.5.⁵⁸ This perception arguably translates into juror confusion, at least in deciding hypothetical cases. Studies have shown that the standard of proof described in traditional legal terms---preponderance, clear and convincing, or beyond a reasonable doubt---does not affect the

⁵⁵ *Id.* at 1331 & tbl.7. Whether judges actually apply the standards faithfully has been questioned. “We doubt very much that an experienced trial judge is much bothered by euphemisms such as clear and convincing or preponderance of the evidence.” *Matter of Estate of Bennett*, 865 P.2d 1062, 1067 (Kan. App. 1993). But even the *Bennett* court stated that citing the wrong standard would be reversible error. *Id.*

⁵⁶ Eleven states and the District of Columbia do not permit jury trials in will contests. ROSS & REED, *supra* note 14, § 4:9 n.67.

⁵⁷ *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 253-54 (1986).

⁵⁸ Rita James Simon & Linda Mahan, *Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom*, 5 LAW & SOC'Y REV. 319, 325 (1971).

percentage plaintiff win rate.⁵⁹ Some of the same studies, however, demonstrate that reframing the standards in numerical terms---51%, 71%, and 91% confidence---eliminates the apparent confusion and leads plaintiff win rates to decline as the required confidence level increases.⁶⁰ In other words, the 0.5 and 0.75 probability levels may not accurately capture present-day reality (at least for jurors and some judges), but numerical standards and jury instructions could remedy that situation.⁶¹

Are such remedies normatively appealing? The affirmative case is straightforward. It should be harder to deprive an individual⁶² of his liberty than of his inheritance.⁶³ In some cases, the costs of one type of error are much higher than the costs of the other. The verbal formulations---beyond a reasonable doubt and clear and convincing evidence---attempt to reflect these policy judgments, but in practice do not. The studies cited above show that people simply do not understand the language: the plaintiff win rate does not vary with the standard.

⁵⁹ Dorothy K. Kagehiro & W. Clark Stanton, *Legal vs. Quantified Definitions of Standards of Proof*, 9 L. & HUMAN BEHAV. 159, 164, 168 (1985); James R.P. Ogloff, *A Comparison of Insanity Defense Standards on Juror Decision Making*, 15 L. & HUMAN BEHAV. 509, 516, 519 (1991); *see also id.* at 521 (only “51.6% of participants identified the correct standard of proof”). *Cf.* David U. Strawn & Raymond W. Buchanan, *Jury Confusion: A Threat to Justice*, 59 JUDICATURE 478, 481 (1976) (“After seeing and hearing the video [standard criminal] instructions, only 50 per cent of the instructed jurors understood that the defendant did not have to present any evidence of his innocence, and that the state had to establish his guilt, with evidence, beyond any reasonable doubt.”). *See generally* Murray Levine, *Do Standards of Proof Affect Decision Making in Child Protection Investigations?*, 22 L. & HUMAN BEHAV. 341, 345 tbl.1 (1998). *But see* SANFORD H. KADISH ET AL., CRIMINAL LAW AND ITS PROCESSES 31-32 (8th ed. 2007) (“Empirical studies confirm that jurors convict more readily when instructed under a more-likely-than-not standard than when instructed under the reasonable doubt standard.”).

⁶⁰ Kagehiro & Stanton, *supra* note 59, at 164-65.

⁶¹ More research is needed, but verbal clarification may suffice. *See, e.g.*, Kagehiro & Stanton, *supra* note 59, at 173 (“Our results suggest that legal definitions may have their intended effect on verdicts if they are presented in some form of comparative context.”); *see also* J.P. McBaine, *Burden of Proof: Degrees of Belief*, 32 CAL. L. REV. 242, 246-47 (1944) (“The only sound and defensible hypotheses are that the trier, or triers, of facts can find what (a) probably has happened, or (b) what highly probably has happened, or (c) what almost certainly has happened.”).

⁶² Like Brooke Astor’s son, Anthony Marshall.

⁶³ *See* Addington v. Texas, 441 U.S. 418, 425 (1979) (“In cases involving individual rights, whether criminal or civil, the standard of proof at a minimum reflects the value society places on individual liberty.”) (internal quotation marks, text alteration notations, and citation omitted).

One counter-argument is that some research suggests that people make systematically better predictions in verbal terms than in numerical ones.⁶⁴ Unlike the survey research described above, however, the cognitive psychology literature on this subject does not examine jurors' ability to follow instructions on hypothetical cases.⁶⁵ Still, this is an important point: if jurors think better in words than in numbers, then the solution may be better verbal formulations of proof standards, not quantification.

Perhaps the strongest argument against quantification⁶⁶ is that quantified standards are insufficiently flexible: the trier of fact should balance the costs of false positives and false negatives in the particular case and adjust the standard of proof accordingly.⁶⁷ Verbal instructions allow such flexibility; numerical standards would not.⁶⁸ The studies showing no variance across legal standards could be cited as evidence that jurors in fact reason this way. Only the legal standard varied across treatments; the fact pattern was the same. The implicit

⁶⁴ Alf C. Zimmer, *Verbal Vs. Numerical Processing of Subjective Probabilities*, in DECISION MAKING UNDER UNCERTAINTY (Roland W. Scholz ed., 1983).

⁶⁵ *E.g.*, *id.* 163-64, 178-79, 180 (memory, college performance predictions, and exchange rate predictions)

⁶⁶ For a persuasive response to a catalogue of other counter-arguments, see Dorothy K. Kagehiro, *Defining the Standard of Proof in Jury Instructions*, 1 PSYCH. SCI. 194, 198 (1990). See also Peter Tillers & Jonathan Gottfried, *Case Comment--United States v. Copeland, 369 F. Supp. 2d 275 (E.D.N.Y. 2005): A Collateral Attack on the Legal Maxim that Proof Beyond a Reasonable Doubt Is Unquantifiable?*, 5 LAW, PROBABILITY & RISK 135, 141-51 (2006).

Professor Clermont argues that quantification would reduce accuracy because hard variables mesh poorly with soft or unquantified variables, dehumanize legal procedure, mask complexity with illusory precision, and fail to "accord with our ingrained way of thinking." Kevin M. Clermont, *Procedure's Magical Number Three: Psychological Bases for Standards of Decision*, 72 CORNELL L. REV. 1115, 1147-48 (1987). None of these arguments is persuasive. Hard and soft variables mesh just as poorly under the current system, but the problem is hidden in ambiguity. If quantified standards can lead to better, more consistent results, then a little "dehumanization" and changing ingrained ways of thinking are prices worth paying. Numerical standards incorporating a modest range could mitigate the illusory precision concern. See *infra*. Most important, Clermont's premise that the present legal standards meaningfully convey information, *id.* at 1148, is apparently false in light of the research findings reported above.

⁶⁷ See, e.g., Elisabeth Stoffelmayr & Shari Seidman Diamond, *The Conflict Between Precision and Flexibility in Explaining "Beyond a Reasonable Doubt,"* 6 PSYCHOL. PUB. POL'Y & L. 769, 784 (2000) ("A single uniform standard across cases is not an optimal resolution when the decisions to which the standard is being applied carry different costs.").

⁶⁸ Tillers & Gottfried, *supra* note 66, at 156. Again, clearer verbal standards are another possible solution that might eliminate confusion and still retain some flexibility.

relative costs of error therefore also remained constant, so the lack of variation in outcomes is consistent with case-by-case balancing.

One response is that quantified standards could be expressed loosely or as a permissible range.⁶⁹ For example, the clear and convincing evidence instruction could require confidence between 70% and 80%. This would in effect divide responsibility for setting the standard between legislatures and appellate judges, on the one hand, and triers of fact, on the other. As long as the standard were not too loose nor the range too wide, such an approach would not substantially impact the analysis of this article.

There are strong arguments against giving jurors too much flexibility in setting the proof standard. One pair of commentators puts the case for quantified standards as follows: (1) triers of fact should not be permitted “to strike a balance that is wildly at variance with the values of society at large”; (2) society’s authorized lawmakers---legislators and appellate judges---should select a proof standard that best accommodates the competing interests; and (3) society can most effectively communicate to triers of fact and enforce its standards by using numbers.⁷⁰ Moreover, under the flexible approach, if costs are nonlinear, which there is good reason to suppose, “efficiency might require that [the trial court] treat similarly situated defendants

⁶⁹ See James Franklin, *Case Comment—United States v. Copeland*, 369 *F. Supp. 2d* 275 (E.D.N.Y. 2005): *Quantification of the “Proof Beyond Reasonable Doubt” Standard*, 5 *LAW, PROBABILITY & RISK* 159, 165 (2006) (advocating “well above a probability of 0.8” for beyond a reasonable doubt); Stoffelmayr & Diamond, *supra* note 67, at 782 (giving example of .87 to .92 range). Cf. Neil B. Cohen, *Confidence in Probability: Burdens of Persuasion in a World of Imperfect Knowledge*, 60 *N.Y.U. L. REV.* 385 (1985) (arguing for confidence interval approach to proof).

⁷⁰ Tillers & Gottfried, *supra* note 66, at 155; see also Redmayne, *supra* note 51, at 183 (weighing of utilities is policy decision “better made by the legislature than by judges”). Cf. Josef Athanas, *Comment: The Pros and Cons of Jury Trials in Will Contests*, 1990 *U. CHI. LEGAL F.* 529, 551 (arguing against jury trials in will contests in part because “the legislature should change the law rather than having juries misapply it”).

differently.”⁷¹ For example, prison overcrowding could ratchet up the criminal standard of proof.⁷²

The case for relatively strict quantification is especially strong in incapacity will contests. Arguments for flexibility are generally directed toward the criminal standard of proof beyond a reasonable doubt.⁷³ Two commentators illustrate some of the policy tradeoffs in criminal cases:

Should society be willing to risk 10 guilty defendants going free rather than one innocent person convicted? Or is the proper ratio 100 to one? Should we be willing to accept lower risks in a “spitting on the sidewalk” case than in a capital homicide case? In a bomb-terrorism case, should the risks be inverted with a preference for convicting 10 innocents rather than letting one guilty go free?⁷⁴

Will contests for incapacity are less diverse and have fewer direct impacts on non-parties. As shown above, the key variable in setting the proof standard is the probability that intestacy reflects testator’s intent if the will were invalid. Some evidence of incapacity may go to this question---e.g., “she told me she wanted me to inherit, so she must have been incapacitated when she signed the will disinheriting me”---but such evidence is not required. Nor will the jury have access to the aforementioned will studies and survey research estimating *I* in different situations.⁷⁵ In short, the case for flexibility is much weaker in will contests than in criminal

⁷¹ Michael L. Davis, *The Value of Truth and the Optimal Standard of Proof in Legal Disputes*, 10 J. L. ECON. & ORG. 343, 351 (1994).

⁷² *Id.*

⁷³ E.g., Lillquist, *supra* note 47.

⁷⁴ Jack B. Weinstein & Ian Dewsbury, *Comment on the Meaning of “Proof Beyond a Reasonable Doubt,”* 5 LAW, PROBABILITY & RISK 167, 168-69 (2006).

It has also been argued that quantifying the reasonable doubt standard would damagingly make explicit society’s acceptance of false convictions. Tribe, *supra* note 47, at 1375. In light of the publicity surrounding recent exonerations, this worry seems almost quaint. Erik Lillquist, *Absolute Certainty and the Death Penalty*, 42 AM. CRIM. L. REV. 45, 90 (2005).

⁷⁵ Cf. Davis, *supra* note 71, at 350-51 (“If the marginal cost of error varies with the total number of errors, a trial court that wished to set the optimal standard would need information about facts not directly before it.”).

Two commentators recommend telling jurors about likely criminal sentences so as to allow better weighing of costs in setting the proof standard. Stoffelmayr & Diamond, *supra* note 67, at 783.

cases. The countervailing advantage of consistency across cases and decision-makers strongly favors quantification (or relatively precise verbal formulation).⁷⁶

D. Courts Correctly Gauge Probability of Incapacity

The vertical lines in the figures imply that courts can accurately discern even the smallest differences in probability of incapacity. For example, under the preponderance standard, every will executed by a testator with a 49% probability of incapacity is upheld, and every will executed by a testator with a 51% probability of incapacity is struck down. While I am not the first to make this simplifying assumption,⁷⁷ it is obviously unrealistic. (Reversals in actual will contests alleging incapacity confirm the obvious.⁷⁸)

How does error impact the choice of standards? The error rate (like parties' uncertainty, *see infra* Section III.J) is almost certainly highest around the decision standard. Assume juries can identify cases within 0.1 of the standard, but can do no better than a coin toss deciding such cases.⁷⁹ Introducing this error profile reduces the percentage point advantage of the clear and convincing evidence standard from 6.5 (Equation 2) to 1.6, assuming a uniform distribution.⁸⁰ Thus, erroneous verdicts and judgments tilt the scale toward the preponderance standard, but by an amount that theory alone cannot determine. Given that only a tiny fraction of cases are

⁷⁶ See Redmayne, *supra* note 51, at 171 (“[C]an the right to consistent weighing of moral harm be protected if the determination of the relevant utilities is left to each individual fact-finder?”).

⁷⁷ George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 8 fig.1 (1984).

⁷⁸ *E.g.*, Paskvan v. Mesich, 455 P.2d 229, 240 (Alaska 1969); Brown v. Brown, 287 P.2d 913 (Okla. 1955).

⁷⁹ Data from two non-representative samples suggest that the jury verdict error rate is at least 11-13%. Bruce D. Spencer, *Estimating the Accuracy of Jury Verdicts*, 4 J. EMPIRICAL LEGAL STUD. 305, 326 (2007).

⁸⁰ Computations on file with author.

See Emily Sherwin, *Clear and Convincing Evidence of Testamentary Intent: The Search for a Compromise Between Formality and Adjudicative Justice*, 34 CONN. L. REV. 453, 463 n.47 (2002):

Under any standard of proof, there will be a certain number of inaccurate estimates of probability, wrongly placing the probability of the required fact on one or the other side of the prescribed line. Some of the erroneous estimates of probability under a clear and convincing standard—those that wrongly conclude that the required fact is highly probable when in actuality it is merely more probable than not—will now produce correct outcomes from the standpoint of truth. But the number of outcomes that fit this description will be overshadowed by the number of wrong outcomes that result from the skewed standard.

litigated to judgment, *see infra* Section III.J, the effect of decision errors is almost certainly very small.

E. Wills Executed by Testators with Capacity Are Accurate (G = 1)

This assumption actually appears twice in the model: once explicitly and once implicitly in the estimates of *I*. The *I* values come from studies comparing the dispositive schemes of probated wills with the pattern that would otherwise have been imposed through intestacy. Using the will as the gold standard assumes that it accurately reflects testators' intent (i.e., that $G = 1$).⁸¹ Commentators have questioned this assumption, suggesting that "cultural standards"⁸² or "customary"⁸³ arrangements may trump testators' true intent. To my knowledge, no one has ever advocated disregarding a will on such grounds. But just because the unchallenged will is the gold standard by default does not mean it perfectly reflects what testators want.

One way to test whether wills generally match testator intent is to compare preferences in wills with preferences expressed elsewhere. In one will study, 85.8% of decedents gave everything to their surviving spouse when also survived by lineal kin.⁸⁴ By comparison, 70.8% of survey respondents said they would leave everything to their spouse if their mother were the only other survivor.⁸⁵ When survived instead by a spouse and children, the percentage of survey

⁸¹ Actually, it goes farther and assumes that all wills, not just good wills, accurately reflect testator's intent. Obviously, wills infected by undue influence or testamentary incapacity are almost certainly inconsistent with this assumption. There is, however, no way from the studies to isolate good wills, but there is reason to think the vast majority of wills are good.

⁸² Ronald J. Scalise Jr., *Honor Thy Father and Mother?: How Intestacy Law Goes Too Far In Protecting Parents*, 37 SETON HALL L. REV. 171, 183 n.64 (2006) ("The drafters of wills are subject to influences by lawyers and other advisors that may be motivated by cultural standards, rather than an individual's actual desires.").

⁸³ Mary Louise Fellows et al., *An Empirical Study of the Illinois Statutory Estate Plan*, 1976 U. ILL. L. F. 717, 722 (1976) ("The attorney who typically drafts the will may have influenced the testate's distributive pattern by suggesting that certain will provisions are customary.").

⁸⁴ SUSSMAN, CATES, & SMITH, *supra* note 34, at 89 tbl.5-1.

⁸⁵ Mary Louise Fellows, Rita J. Simon, & William Rau, *Public Attitudes About Property Distribution at Death and Intestate Succession Laws in the United States*, 1978 AM. BAR FOUND. RES. J. 319, 351 tbl.7. The Sussman study involved exclusively Ohio wills. Fellows' survey also included Ohio, along with four other states. The Ohio respondents in Fellows' study were particularly generous to the surviving spouse: 82.8% gave the spouse the entire estate, *id.* at 352 tbl. 8, which is consistent with Sussman's findings.

respondents giving the entire estate to the spouse dropped to 51.6%-58.3% depending on the age of the children.⁸⁶ In one sense, these broadly consistent results are reassuring: majorities in will studies and telephone surveys alike favor the surviving spouse over all other kin. The direction of the difference suggests that culture and custom are not more powerful in influencing a will than a survey response. The intestacy statute reflects society's judgment about how an estate should be distributed, generally forcing the surviving spouse to share with lineal kin. That people deviate further from that judgment in their wills than in surveys supports the view that wills are not often driven by culture and custom.

To culture and custom, one must add old-fashioned mistake as a source of deviation from testator intent. A classic example is *Mahoney v. Granger*,⁸⁷ in which the testator instructed her lawyer that she wanted her twenty-five cousins to share her estate equally. Apparently not realizing that the testator also had a living aunt, the lawyer drafted a residuary clause in favor her "heirs at law." Because the aunt was the only "heir at law" under Massachusetts law, the cousins took nothing, despite the testator's clear statement that she wanted them to inherit. There is a modern trend to correct such errors, but the majority of jurisdictions still follow the traditional rules of no extrinsic evidence and no reformation.⁸⁸

How much worse than perfection would good wills have to be to favor the preponderance standard? Using the uniform distribution, G would have to be less than 0.827.⁸⁹ In other words, a will executed by a person with capacity and free of any other defects would have to reflect the testator's intent less than 82.7% of the time to tip the scales toward the preponderance standard.

⁸⁶ *Id.* at 359 tbls.11, 12. Again, Ohioans were most generous to their spouses. *See id.* at 361 tbls.13, 14.

⁸⁷ 186 N.E. 86 (Mass. 1933).

⁸⁸ JESSE DUKEMINIER ET AL., WILLS, TRUSTS, AND ESTATES 365-66 (7th ed. 2005).

⁸⁹ 79.7% for the $y = 1 - x$ distribution.

Custom and culture may be powerful and mistakes numerous, but it is hard to believe they introduce a 17.3% error rate into wills.⁹⁰

F. Will Studies Show How Well Intestacy Matches Testator Intent ($I \leq 0.31$)

This estimate for I is derived from studies of wills probated in 1982 and earlier.⁹¹ Obviously, much has changed since 1982, but three potential sources of error stand out: (1) changes in intestacy law; (2) changes in family structure and hence survivorship; and (3) changes in testator preferences.

By far the biggest reason for the disparity between testator preferences and intestacy is that testators usually want their surviving spouse to take the entire estate whereas intestacy statutes often divide the estate. For example, in the Cleveland wills study, 86% of testators devised the entire estate to their surviving spouse when they were also survived by lineal kin (generally parents and/or children); the intestacy statute divided the estate.⁹² After wills studies and survey research⁹³ pointed out this incongruity, some states modified their intestacy statutes to give surviving spouses a larger share.⁹⁴ This would tend to raise I , to be sure, but a clear majority of states still require a split.⁹⁵ In most of the country, therefore, it does not appear that intestacy statutes have substantially improved in advancing decedents' intent regarding surviving spouses.⁹⁶

⁹⁰ One could argue that these aren't really errors at all. The lawyer who convinces a testator to dispose of property in a different way most likely acts in good faith. And the testator's changed mind may reflect their new intent, not a custom-induced error.

⁹¹ Schneider, *supra* note 33, at 412 (1987); SUSSMAN, CATES, & SMITH, *supra* note 34 (1970); Dunham, *supra* note 35 (1963); Ward & Beuscher, *supra* note 36 (1950).

⁹² SUSSMAN, CATES, & SMITH, *supra* note 34, at 89.

⁹³ *E.g.*, Fellows, Simon, & Rau, *supra* note 87.

⁹⁴ Susan N. Gary, *Adapting Intestacy Laws to Changing Families*, 18 LAW & INEQ. 1, 2 n.8 (2000).

⁹⁵ See Laura A. Rosenbury, *Two Ways To End a Marriage: Divorce or Death*, 2005 UTAH L. REV. 1227, 1266 (2005) ("In . . . thirty-four states, the surviving spouse never receives the deceased spouse's entire estate in any situation where the deceased spouse left children, even if the surviving spouse is the other parent of those children.").

⁹⁶ At least where there are no children from a previous marriage, *see infra*.

The case of the surviving spouse may still be problematic, but perhaps due to increasing divorce rates there are fewer of them. It is true that the total percentage of Americans married declined from 67% in 1950 to 54% in 2008,⁹⁷ but the percentage of persons 65 and older who were married actually increased from 55% in 1980 to 58% in 2007.⁹⁸ And persons 65 and older account for 73% of deaths annually.⁹⁹ Of course, with divorce more common today, more of these married older adults may be in second marriages. Survey research suggests that individuals with a surviving spouse and child from a previous marriage do favor splitting the estate.¹⁰⁰ Twenty-nine states “now give the surviving spouse a reduced share if she is not the other parent of the deceased spouse’s children.”¹⁰¹ How common the second marriage case is, and how well modern intestacy statutes approximate testator intent are questions that would need to be answered before updating estimates of *I*.¹⁰²

Notably, the same demographic change exerts force in the opposite direction. Remarriage often creates stepchild and stepparent relationships. It has been estimated that “one out of every three Americans is a member of a stepfamily.”¹⁰³ Intestacy law generally excludes step relationships. Some have argued that this thwarts many stepparents’ intent.¹⁰⁴

⁹⁷ “MS-1. Marital Status of the Population 15 Years Old and Over, by Sex and Race: 1950 to Present,” at <http://www.census.gov/population/socdemo/hh-fam/ms1.xls> (visited June 22, 2009).

⁹⁸ U.S. Census Bureau, Statistical Abstract of the United States: 2009, at 37 tbl.33; U.S. Census Bureau, Statistical Abstract of the United States: 2000, at 42 tbl.40.

⁹⁹ Melonie Heron et al., Deaths: Final Data for 2006, 57 Nat’l Vital Statistics Reports No. 14, tbl.3 (Apr. 2009), at ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/NVSR/57_14/Table03.xls (visited June 22, 2009).

¹⁰⁰ Fellows, Simon, & Rau, *supra* note 87, at 366 tbl.18 (23.0% wanted spouse to take entire estate; 28.9% wanted spouse to take between 51-99%; 37.2% wanted an equal split between spouse and child; and 11.0% wanted spouse to take less than 50%).

¹⁰¹ Rosenbury, *supra* note 95, at 1268.

¹⁰² Note that (assuming a uniform distribution) if $I > 0.375$, that would not necessarily tilt the scale in favor of the preponderance standard. One would need to know *B*, the probability of a bad will reflecting intent.

¹⁰³ Gary, *supra* note 94, at 30.

¹⁰⁴ RALPH C. BRASHIER, INHERITANCE LAW AND THE EVOLVING FAMILY 157 (2004) (“The sweeping but easily administered objective rule that denies stepchildren an intestate share under the laws of most states undoubtedly is at odds with the wishes of many stepparents who dies intestate.”).

Other demographic changes similarly push down *I*. “Approximately 7% of the nation’s couples are in unmarried committed relationships, including roughly 1.7 million gay and lesbian couples.”¹⁰⁵ (This is consistent with the overall marital status numbers reported above.) Because same-sex couples are permitted to marry in only a handful of states, intestacy law’s exclusion of unmarried partners raises antidiscrimination concerns as well. Unmarried, committed partners overwhelmingly prefer that the surviving partner share in the estate,¹⁰⁶ but intestacy statutes give unmarried partners nothing.

A final concern with relying on relatively old research findings is that people’s dispositive preferences may have changed over time. Short of undertaking a new study, there is little to say in response to this concern. However, given the stability of observed *I* over studies from diverse areas over a 47-year period, perhaps the burden should be on the proponent of change to offer a persuasive theory for why preferences would have changed in the direction of intestacy statutes.

G. Fallback Is Intestacy

Not every successful will contest results in the estate being divided according to the intestacy statute. Sometimes the court gives effect to a prior will. The Brooke Astor case described above illustrates. There, one allegation was that her son wrongfully induced her to

¹⁰⁵ Jennifer K. Robbennolt & Monica Kirkpatrick Johnson, *Legal Planning for Unmarried Committed Partners: Empirical Lessons for a Preventive and Therapeutic Approach*, 41 ARIZ. L. REV. 417, 418 (1999); see also Jason Fields, *America’s Families and Living Arrangements: 2003*, at 16, Current Population Reports, U.S. Dep’t of Commerce (Nov. 2004) (“The proportion of all households that were unmarried-partner households [opposite-sex only] has been steadily increasing, from 2.9 percent of all households in 1996 to 4.2 percent in 2003.”); BRASHIER, *supra* note 104, at 2 (arguing that census numbers are “probably underrepresentative because some unmarried cohabitants are reluctant to identify themselves as such and because the number excludes unmarried couples who reside in a household headed by someone else”).

¹⁰⁶ Mary Louise Fellows et al., *Committed Partners and Inheritance: An Empirical Study*, 16 L. & INEQ. 1, 38, 47, 50 (1998); accord Robbennolt & Johnson, *supra* note 105, at 441.

execute a codicil giving him her residuary estate, which was to go to charity under her prior will.¹⁰⁷ What is the effect of a prior will on the proof standard question?

The accuracy of intestacy does not depend on testator's capacity. The likelihood that a prior will reflects testator's intent does. If the testator had capacity when executing a codicil or new will, then the old will almost necessarily fails to reflect intent ($U_{FP} = 0$). If, on the other hand, the testator lacked capacity, then there would seem to be a very high probability that the prior will was accurate ($U_{TP} = P$).¹⁰⁸ Plugging these values into Equation (3) and assuming a uniform distribution, the clear and convincing standard wins only if $B > P - 0.6$.¹⁰⁹ This seems unlikely given that we expect B to be very close to zero and P to be very close to one. Jurisdictions currently requiring clear and convincing evidence of incapacity should consider instead requiring only a preponderance when evaluating a codicil or replacement will.¹¹⁰

H. Promoting Testator Intent Is Dichotomous

A closely related assumption for all three parameters (G , B , and I) is that a distribution either matches testator's intent perfectly (equal to one) or it fails utterly (equal to zero). Of course, some distributions would be more offensive to the testator than others. The intestacy statute may attempt to achieve distributions that are least offensive to most people rather than achieving perfect results in the maximum number of cases.¹¹¹ Failing to award partial credit, as it were, may be a heavy thumb on the scale against intestacy and therefore in favor of the clear and convincing standard.

¹⁰⁷ Indictment, *supra* note 11, at 6-7 (Count 9).

¹⁰⁸ *Cf.* Sherwin, *supra* note 80, at 463 ("an informal document offered as a will often reflects a disposition that the testator at least considered, even if the disposition was ultimately rejected").

¹⁰⁹ Skewing the distribution toward the left (toward capacity) favors the clear and convincing standard, but probably not by enough to change the bottom line. Applying the $y = 1 - x$ distribution, the clear and convincing standard dominates if $B > P - 0.64$, which still seems very unlikely.

¹¹⁰ *Cf.* Stoffelmayr & Diamond, *supra* note 67, at 782 (suggesting that "a legislature might set different quantitative [proof] levels for different offenses").

¹¹¹ See Lawrence W. Waggoner, *Marital Property Rights in Transition*, 59 MO. L. REV. 21, 29 (1994) ("No intestacy regime can hope to be 'suitable' for every person who dies intestate.").

The case against partial credit relies on formalism and realism. Formally, a distribution either is or is not what the testator wanted. A solution that gets it only partially right is not what the testator wanted. The argument from realism is based on the difficulty in awarding partial credit. Assume the testator wanted to give everything to his surviving spouse, but the intestacy statute awards only half to his spouse and the rest to his children. Does this advance testator's intent by 0.5, more, or less? Would it matter if the testator were estranged from his children? Even this simple case illustrates that non-arbitrary intermediate values would be difficult to estimate.

Note too that the net effect of relaxing the assumption is ambiguous in theory. No doubt I would go up, but so too would B . Even wills executed by a testator without capacity probably include some provisions that are not wholly objectionable. For example, the codicil in the Brook Astor case changed the residuary clause, but left the rest intact. Which effect dominates? To tip the scale in favor of preponderance (assuming a uniform distribution), the change in I (ΔI) would have to be greater than $0.065 + 0.625 \times \Delta B$. Using one will study,¹¹² for example, if the intestate splitting arrangement advanced the intent by 0.5 of testators who were survived by a spouse and lineal kin and who willed everything to their spouse, I would rise by 0.22, from 0.31 to 0.53. B would have to rise by 0.25 to outweigh I 's increase and save the clear and convincing evidence standard. It seems unlikely that incapacitated testators are on average 25% satisfied with their wills. Then again, assuming that all testators who will everything to their spouse give 50% approval to splitting is probably an overestimate. If one assumes instead that only half of such testators give 50% partial credit and half zero, then I would rise by 0.11 and B would need to rise by only 0.07 for the clear and convincing standard to prevail. Data are needed to resolve the question, but awarding partial credit plainly has the potential to flip the result.

¹¹² SUSSMAN, CATES, & SMITH, *supra* note 34.

I. Promoting Testator Intent Is the Only Objective

It has been suggested that the mental capacity requirement does not promote testamentary freedom at all. One commentator in a classic wills handbook argued that eliminating the mental capacity requirement “would be carrying the conception of freedom of testation to its fullest extent.”¹¹³ In a different context, I have argued against this type of reasoning.¹¹⁴ And in fairness to the handbook author, a subsequent edition adds the correct statement that “an attitude of consent is essential and this presupposes some degree of mental capacity.”¹¹⁵

It is now widely accepted that promoting testator intent is at least one goal of the capacity requirement.¹¹⁶ Other reasons include: (1) to protect the family; (2) an incompetent man or woman is not a “person”; (3) to preserve an appearance of legitimacy; (4) to protect the sane testator who later loses capacity; (5) to protect society from irrational acts; and (6) to protect vulnerable testators from exploitation.¹¹⁷ The “personhood” line of argument appears to devolve into testator intent: it is the inability to formulate intent, to reason, that deprives the insane of personhood.¹¹⁸ Similarly, the fourth and six reasons are just specific manifestations of the concern for testator intent: we protect the sane will because we think it reflects testator’s intent and we worry about exploitation because we fear it might overcome a vulnerable person’s intent. The third and fifth reasons both address externalities; neither seems as important as what happens to the interested parties: the testator and the family.

¹¹³ ATKINSON, *supra* note 14, § 77, at 186 (1937); *see also* George Alexander & Thomas Szasz, *From Contract to Status via Psychiatry*, 13 SANTA CLARA L. REV. 537 (1973) (similar argument for contracts).

¹¹⁴ *See* Fredrick E. Vars, *Illusory Consent: When an Incapacitated Patient Agrees to Treatment*, 87 OR. L. REV. 353, 355 (2008) (“Assent without capacity is not an expression of autonomy; it is at best an illusion of autonomy.”).

¹¹⁵ THOMAS E. ATKINSON, *HANDBOOK OF THE LAW OF WILLS* § 51, at 233 (2d ed. 1953).

¹¹⁶ DUKEMINIER ET AL., *supra* note 85, at 146; Mary Louise Fellows, *In Search of Donative Intent*, 73 IOWA L. REV. 611, 621 (1988).

¹¹⁷ DUKEMINIER ET AL., *supra* note 85, at 146-48.

¹¹⁸ *Cf.* Bailey H. Kuklin, *The Asymmetrical Conditions of Legal Responsibility in the Marketplace*, 44 U. MIAMI L. REV. 893, 925 (1990) (“Autonomous choices entail knowledge (information and foreseeability) and capability (reason and sense).”).

When the testator's and family's interests conflict, who wins? The general answer in the law of wills is that the testator wins.¹¹⁹ If there is a valid will, it will control, subject to very few exceptions.¹²⁰ But that is not quite the right question when the issue is setting the proof standard for an incapacity contest. Instead, one needs to ask what influence, if any, family protection should have on the values of *G*, *B*, and *I*. The relative weight of *G* and *B* would seem independent of family concerns: a will will be enforced with or without any provision for family. That leaves *I*. One goal of intestacy is clearly family protection.¹²¹ Should *I* as a result be assigned a value higher than its percentage correspondence with testator intent?

Some have argued that family protection is the primary goal of the capacity requirement.¹²² That seems plainly wrong. Despite suggestions to the contrary,¹²³ the capacity doctrine is not an open-ended invitation for juries to strike down dispositions that do not adequately provide for family.¹²⁴ Favoring intestacy over testation in this way would be a blunt instrument to protect the family.¹²⁵ The law could do so more directly with an elective share for spouses, forced heirs for issue, or a family maintenance system like the one in England, Australia, or Canada.¹²⁶ These would represent express policy judgments that certain dispositive

¹¹⁹ See Pamela Champine, *Expertise and Instinct in the Assessment of Testamentary Capacity*, 51 VILL. L. REV. 25, 55 (2006) ("As important as the family preference is, it takes second place to the overriding policy of wills law, which is effectuating the testator's intent."); see also *id.* at 54 ("the family's interest, in all cases, derives from the testator's presumed intention to benefit them"). But see Jeffrey M. Alden, *Testamentary Capacity in a Nutshell: A Psychiatric Reevaluation*, 18 STAN. L. REV. 1119, 1123-24 (1966) (the primary concern of testamentary capacity is protection of the family).

¹²⁰ E.g., U.P.C. § 2-202 (elective share of surviving spouse).

¹²¹ Milton D. Green, *Public Policies Underlying the Law of Mental Incompetency*, 38 MICH. L. REV. 1189, 1216 (1940).

¹²² *Id.* at 1218-19.

¹²³ Milton D. Green, *Proof of Mental Incompetency and the Unexpressed Major Premise*, 53 YALE L.J. 271 (1944).

¹²⁴ Alexander M. Meiklejohn, *Contractual and Donative Capacity*, 39 CASE W. RES. L. REV. 307 (1988-89). Fairly applied, the capacity doctrine could strike down a will giving the bulk of an estate to family (e.g., Brooke Astor's son) in favor of a prior will giving the lion's share to charity.

¹²⁵ Jane B. Baron, *Empathy, Subjectivity, and Testamentary Capacity*, 24 SAN DIEGO L. REV. 1043, 1051 (1987).

¹²⁶ See Ronald Chester, *Should American Children Be Protected Against Disinheritance?*, 32 REAL PROP. PROB. & TR. J. 405 (1997) (advocating adoption of British Columbia family maintenance system). Louisiana promotes testator intent through a high incapacity standard and the family through forced heirship.

schemes should not be allowed. Most important, they leave intact a testator's other dispositive wishes, unlike a finding of incapacity.¹²⁷

In sum, advancing testator's intent should be the only goal of the capacity requirement, but there is certainly fair room for policy debate on this issue.

J. Every Case Is Litigated to Judgment

The simple answer, by applying the decision rule to every testator, implicitly assumes that all cases are litigated to judgment and hence subjected to the rule. That simply isn't true. About one percent of testate cases involve will contests,¹²⁸ although a high fraction of those contain an allegation of testamentary incapacity.¹²⁹ Hence, the overwhelming majority of wills are either enforced as written or informally revised through settlement by interested parties. Which cases are actually litigated?

A leading model posits that uncertainty over the outcome drives litigation.¹³⁰ Uncertainty is plainly greatest around the decision standard, so that is where the bulk of litigated cases should be.¹³¹ To take an extreme case, assume that incapacity will contests are litigated to judgment only where the testator's likelihood of incapacity is within 0.1 of the decision standard (and all such cases are in fact litigated to judgment). How does each standard of proof advance testator

¹²⁷ See Baron, *supra* note 125, at 1051 ("regulatory effects of holding a testator incompetent on mental incapacity grounds are considerably more draconian than the explicit regulations set forth in elective share statutes and the like").

¹²⁸ Jeffrey A. Schoenblum, *Will Contests---An Empirical Study*, 22 REAL PROP. PROB. & TR. J. 607, 613 chart 3 (1987) (0.86%, 66/7638 filed wills); Jeffrey P. Rosenfeld, *Will Contests: Legacies of Aging and Social Change*, in INHERITANCE AND WEALTH IN AMERICA 185 tbl.8.3 (Robert K. Miller, Jr. & Stephen J. McNamee, eds. 1998) (1.14%, about 90/7817 probates); SUSSMAN, CATES, & SMITH, *supra* note 34, at 184 (1.32%, 6/453 testate cases). Will contest rates were apparently higher historically in some jurisdictions. See Kristine S. Knaplund, *The Evolution of Women's Rights in Inheritance*, 19 HASTINGS WOMEN'S L.J. 3, 30 & n.183 (2008).

¹²⁹ Schoenblum, *supra* note 128, at 648 chart 22 (1987) (73%, 48/66 will contests).

¹³⁰ Priest & Klein, *supra* note 77. See Bruce L. Hay & Kathryn E. Spier, *Settlement of Litigation*, in 3 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW 443 (Peter Newman, ed. 1998) (describing divergent party expectations as "the most influential account of why cases may fail to settle").

¹³¹ Priest & Klein, *supra* note 77, at 20 ("According to the model, given some error in the parties' estimates of Y', an interval will exist around the decision standard which will contain some large fraction of the set of litigated disputes.").

intent in the litigated cases? The clear and convincing evidence standard wins only where $B > 1$, which is to say never.¹³² Settlement, in other words, has the power to flip the simple answer.

Professor Emily Sherwin has suggested in another context that greater ambiguity about the meaning of the clear and convincing standard (*see supra* Section III.C) would lead to less settlement and more litigation.¹³³ This is an excellent theoretical point, but ignores the underlying distribution and the possibility of more precise jury instructions. First, as explained above (*see supra* Section III.A), there is good reason to think the merits distribution is skewed toward capacity. This means fewer cases near the clear and convincing standard than near preponderance. Therefore, even if a higher percentage of cases near the heightened proof threshold fail to settle, the overall settlement rate might increase. Second, and more important, even with a uniform distribution, quantifying the clear and convincing proof standard (*see supra* Section III.C) could reduce uncertainty and thereby facilitate settlement.

What can we say about cases that are not litigated? The standard model of litigation assumes that a plaintiff will sue when the expected benefits outweigh the expected costs.¹³⁴ Expected benefits include the probability of prevailing multiplied by the expected amount of an award. Litigation costs money, so if the parties accurately assess trial prospects there are almost always gains to be achieved through settlement.¹³⁵ That is why the case selection model above is premised on the notion that parties cannot foresee trial outcomes with perfect accuracy. We assumed above a zone of uncertainty of plus or minus 0.1. Below that threshold, no will contest should ever succeed; above that threshold, the proponent of the will should always agree to a

¹³² The text reports the result for the uniform distribution. The dominance of the preponderance standard is even stronger with a leftward skew (i.e., $y = 1 - x$).

¹³³ Sherwin, *supra* note 80, at 470-73.

¹³⁴ Robert D. Cooter & Daniel L. Rubinfeld, *Economic Analysis of Legal Disputes and Their Resolution*, 27 J. ECON. LIT. 1067, 1075-76 (1989).

¹³⁵ “Almost” because a party may value trial as an opportunity for self-expression or publicity, for example. Hay & Spier, *supra* note 130, at 442. *See infra* text accompanying note 141 (“strategic litigation”).

settlement favoring the contestant.¹³⁶ If that happens, the simple answer holds, even though only a tiny fraction of cases are litigated.

But do settlements replicate trial outcomes? Economic theory predicts that they will if “(1) the plaintiff and the defendant have the same expectations about the trial, and (2) the plaintiff and defendant bear the same transaction costs to resolve the dispute.”¹³⁷ The second condition will not be satisfied with respect to some will contests. The costs of defending a will contest are generally borne by the estate.¹³⁸ This means that an interested beneficiary if successful in the will contest may pay only a portion of the cost of litigation. Even more important, the unsuccessful will beneficiary pays nothing out of pocket. The beneficiary is therefore likely to be more aggressive and successful in settlement negotiations than at trial. But the beneficiary (unless also the executor) is not the one who controls the litigation in such cases. The executor may have less incentive to aggressively defend the will, which would tend to tilt the scales back toward pro-plaintiff settlements. That may explain why some beneficiaries choose to intervene in will contests.¹³⁹

So the estate paying the costs of will contests pushes settlements to favor will proponents. But how does this affect the selection of proof standard? In effect, it elevates the effective proof standard in settlement. For example, the decision rule at trial may be 0.50, but only contestants with cases above 0.60 will succeed in settlement. Because the vast majority of cases are not

¹³⁶ Cf. Hay & Spier, *supra* note 130, at 445 (“As an initial approximation, . . . we might anticipate that cases will generally settle for an amount roughly equal to the expected judgment at trial.”).

¹³⁷ ROBERT COOTER & THOMAS ULEN, *LAW & ECONOMICS* 445 (5th ed. 2008). See generally Janet Cooper Alexander, *Do the Merits Matter? A Study of Settlements in Securities Class Actions*, 43 *STAN. L. REV.* 497, 499 (1991) (“Moreover, the available theoretical models of settlement behavior predict or imply that settlement outcomes will approximate trial outcomes.”).

¹³⁸ E.g., Tex. Probate Code § 243 (Thomson Reuters 2009).

¹³⁹ See *Matter of Estate of Roosa*, 753 P.2d 1028, 1031 (Wyo. 1988) (sole beneficiary allowed to intervene in will contest). But see *Peery v. Swafford*, No. 03A01-9803-CV-00087, 1998 WL 744109, at 1 (Tenn. Ct. App. Oct. 2, 1998) (rejecting beneficiary’s intervention motion because it was too late and because his interests were adequately represented by the proponents of the will).

litigated to judgment, the simple answer's analysis should be applied to the settlement threshold, not the litigation threshold. Depending upon its magnitude, this effect could flip the answer. The presumptive optimal threshold, again, is 0.69. Shifting the standards upward moves the preponderance threshold closer to that optimum and the clear and convincing standard farther from it.

Even where the assumptions hold, empirical evidence has called into question economic theory's prediction that settlements will approximate trial outcomes. One analysis of securities class actions data concluded that the merits did not affect settlements, citing an agency problem akin to the one discussed above, along with a host of other factors.¹⁴⁰ The researcher speculated that a similar disconnect might be found in "strategic litigation, where the ultimate outcome of the litigation is not its real object but the mere filing of a lawsuit achieves the litigant's objective."¹⁴¹ Will contests do not seem a likely candidate for strategic litigation.

Something other than money, however, is plainly motivating will contestants. One researcher concludes: "there is no correlation between size of the estate and the likelihood of contest."¹⁴² If will contestants are deviating from the economic model in this respect,¹⁴³ are they also indifferent to the merits? That seems unlikely. A failed will contest would seem to vindicate non-monetary interests (e.g., fairness, respect, etc.) little more than monetary ones. But settlements can deviate from the merits for other reasons.

¹⁴⁰ Alexander, *supra* note 137, at 597.

¹⁴¹ *Id.*

¹⁴² Schoenblum, *supra* note 128, at 617. *Accord* Rosenfeld, *supra* note 125, at 185 tbl.8.3.

¹⁴³ Hay & Spier, *supra* note 130, at 444 ("the likelihood of settlement decreases as the amount at stake in the case increases"). The asymmetric cost effect described *supra* text accompanying notes 138-39 is mitigated to the extent the parties are motivated by factors other than money.

Economic theory has developed an explanation for settlements in weak cases, a further divergence from the view that settlements replicate trial outcomes.¹⁴⁴ The basic idea is that it is costly for defendants to assess the merits of a plaintiff's case. It would seem less costly the farther away one is from the proof standard. This suggests that the clear and convincing standard should reduce nuisance suits, because the distribution is quite probably skewed toward capacity. Fewer weak suits are close enough to the proof standard to masquerade as strong ones. It has been asserted that "[a] great percentage of will contests are merely for nuisance value."¹⁴⁵ But do the data support this assertion?

One study of 66 filed will contests found that 24 (36%) settled out of court.¹⁴⁶ This settlement rate is substantially lower than in general civil litigation,¹⁴⁷ suggesting relatively few nuisance suits.¹⁴⁸ But perhaps more probative is the fact that "the contestant rarely obtained less than a substantial portion of the estate."¹⁴⁹ To be sure, large settlement amounts are not inconsistent with the economic theory of nuisance suits,¹⁵⁰ but settlements for small amounts are more likely to indicate nuisance suits. Fourteen additional contests (21%) were either voluntarily terminated or dismissed,¹⁵¹ which may represent nuisance suits that were unsuccessful. The

¹⁴⁴ E.g., Avery Katz, *The Effect of Frivolous Lawsuits on the Settlement of Litigation*, 10 INT'L REV. L. & ECON. 3 (1990).

¹⁴⁵ Walter T. Burke, *Avoiding Pitfalls When Counseling the Elderly*, 15 NO. 2 COMPLEAT LAW. 13, 18 (1998).

¹⁴⁶ Schoenblum, *supra* note 128, at 620.

¹⁴⁷ Hay & Spier, *supra* note 130, at 442 ("Of filed lawsuits in America, over 90 percent settle before trial . . ."); RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* § 21.4, at 597 (7th ed. 2007) ("only 4 percent of all civil cases in state courts are tried").

¹⁴⁸ Alternative interpretations are possible. See, e.g., Ray D. Madoff, *Lurking in the Shadow: The Unseen Hand of Doctrine in Dispute Resolution*, 76 S. CAL. L. REV. 161, 177-82 (2002) (arguing that will contests are less likely to settle because testator is dead, testator's intent is malleable, doctrine has moral dimension, and the remedy is all-or-nothing).

¹⁴⁹ Schoenblum, *supra* note 128, at 621.

¹⁵⁰ Katz, *supra* note 144, at 4.

¹⁵¹ Schoenblum, *supra* note 128, at 619 chart 6.

remaining 28 cases (42%) were either adjudicated or still pending.¹⁵² Seventeen out of the 25 adjudicated cases (68%) were decided in favor of the proponent.¹⁵³ Without an independent measure of the merits of settled and dismissed cases, it is impossible to know how common nuisance will contests are, but the data do not seem to indicate an unusually high percentage. Nor do the data rule out a substantial number, which, again, would tilt in favor of the clear and convincing standard.

By focusing on filed will contests this study may have missed settlements that avoid a court filing. Another study of wills found a redistribution of the estate in 50 out of 360 (14%) testate cases.¹⁵⁴ Twenty-one redistributions involved a car,¹⁵⁵ which seems an unlikely way to resolve a potential will contest. In 17 cases, property was redistributed from the surviving spouse to lineal descendants or vice versa.¹⁵⁶ Because the spousal intestate share varied from one-third to three-quarters,¹⁵⁷ these redistributions could have preempted will contests. Of the remaining 12 redistributions, only 3 appear at all likely to have been settlements of threatened will contests.¹⁵⁸ Overall, then, as many as 20 out of 360 testate cases (5.6%) may have involved a pre-filing settlement. While this represents a high-end estimate, it shows that examining only filed contests may overlook a large number of settlements. But again, without knowing the merits of the potential claims, it impossible to say whether these settlements mirror trial outcomes or simply avoid nuisance actions.

¹⁵² Another will study suggests that even fully litigated will contests often include settlements: five out of five contests litigated to jury verdict involved an out-of-court settlement. SUSSMAN, CATES, & SMITH, *supra* note 34, at 184.

¹⁵³ Schoenblum, *supra* note 128, at 626 chart 8. This finding of “no pronounced bias toward the proponent or contestant,” *id.* at 627, is broadly consistent with the Priest-Klein case selection model described above. *See supra* note 130. *Cf.* Mark A. Lemley & Colleen V. Chien, *Are the U.S. Patent Priority Rules Really Necessary?*, 54 HASTINGS L.J. 1299, 1312 n.51 (2003) (finding that win rates for claims requiring proof by preponderance were no different than for claims requiring clear and convincing evidence).

¹⁵⁴ SUSSMAN, CATES, & SMITH, *supra* note 34, at 122-25.

¹⁵⁵ *Id.* at 122.

¹⁵⁶ *Id.* at 122-23.

¹⁵⁷ *Id.* at 89.

¹⁵⁸ *Id.* at 123 (reporting that three cases involved disinheritance or near disinheritance).

To summarize, the greater uncertainty of the clear and convincing evidence standard might lead to greater litigation, but quantifying the standard could avoid that problem. The asymmetric cost structure of will contests raises the effective proof standard higher than the decision rule at trial. This pushes in favor of the preponderance standard. Pushing in the other direction is the desire to avoid nuisance suits. Nuisance suits do not appear to be particularly pronounced in will contests, but the data do not rule out the possibility that they occur in significant numbers. In short, settlement complicates substantially the choice of proof standard.

IV. A BETTER ANSWER?

Laypeople estimate the probability required for the preponderance standard at the same level as judges place the clear and convincing evidence standard.¹⁵⁹ And mock jury research shows no significant difference in plaintiff win rates between the two standards.¹⁶⁰ The entire civil standard of proof question is meaningless in the absence of jury instruction reform. Quantified or comparative standards appear to be effective. If legislatures and appellate courts want any policymaking role in setting the standard of proof, they should institute reform in that direction.

Assuming the two standards are meaningfully differentiated, which standard is better for incapacity will contests? The simple answer presented here mirrored the standard approach of comparing the utilities of the possible outcomes and concluded that the clear and convincing evidence standard dominated. Examination of the assumptions underlying this answer, however, call it into some doubt.

Relaxing the assumption that the fallback is intestacy strongly suggests that the preponderance standard is superior when there is a prior will. That conclusion seems relatively

¹⁵⁹ See *supra* notes 31, 58 and accompanying text.

¹⁶⁰ See *supra* note 59 and accompanying text.

robust even in light of uncertainties created by the other assumptions. Several other factors push in favor of the preponderance standard more broadly but with unknown force: (1) testator intent is arguably not dichotomous, which would tend to increase the utility of distribution via intestacy; (2) testator intent is arguably not the only goal of the capacity requirement, which would similarly make intestacy more attractive; (3) decision error favors the preponderance standard by a small amount; (4) so too the asymmetric cost structure of will contests tilts toward preponderance; and (5) the preponderance standard creates marginally better incentives to produce evidence.

Against these factors, the clear and convincing evidence standard (1) gets better in achieving testator intent if the distribution of testators, as seems almost certain, is skewed toward capacity, and (2) probably reduces nuisance suits more than the preponderance standard. A mystery factor is whether old will studies reflect the accuracy of intestacy today, given changes in statutes, family structures, and testators.

If forced to choose based on presently existing information (as courts and legislators are), the case for the clear and convincing evidence standard seems stronger on balance. The simple answer weighs heavily in its favor. Using conservative estimates of the utilities of the four possible outcomes generates an optimal standard of 0.69,¹⁶¹ not far from the 0.75 level associated with the clear and convincing standard. The likely distributional skew toward capacity strengthens this finding, as does concern for reducing nuisance suits. On the other hand, asymmetric costs, evidentiary incentives, and decision error push toward the preponderance standard. As noted above, asymmetric costs are potentially very important given how few cases

¹⁶¹ If one accepts the argument for quantification, the question becomes why not simply set the standard at 0.69. Indeed, more refined proof levels are already at work in this context. Recall the dizzying array of verbal formulations. *See supra* text accompanying notes 16-21. Some have argued that three is a magic number. Clermont, *supra* note 66; *see also* McBaine, *supra* note 61. Whether more refined standards are appropriate is outside the scope of this article.

are litigated, but they are mitigated in this context because will litigants appear to be motivated by factors other than money. Evidentiary incentives do not seem particularly profound given uncertainty, risk aversion, and the other incentives at work. Decision error applies only to that tiny fraction of cases litigated to judgment.

That leaves three more difficult considerations. First, the will studies underlying the simple answer are dated, but that should not disqualify them in the absence of new data or persuasive theory about the direction of change. Second, awarding partial credit for distributions that deviate from the testator's preferred scheme would be arbitrary without evidence that doing so is what testators want.¹⁶² Finally, if family protection is given weight in setting the incapacity proof standard, then that could tilt the scales in favor of the preponderance standard. However, as argued above, testator intent should be the only relevant goal.¹⁶³

In light of all the relevant factors, a definitive answer to the question is not possible based on currently available data. There are simply too many unknowns, pushing in opposite directions. Should partial credit be awarded and, if so, how much? Do old will studies accurately reflect the current situation? Are other goals relevant? One commentator has argued that because the preponderance standard minimizes the total number of errors, “[a]ny deviation from this rule requires unequivocal justification, which is hard to provide in cases not featuring

¹⁶² Researchers could ask respondents not just how they would prefer their estate to be distributed under different factual scenarios, *e.g.*, Fellows, Simon, & Rau, *supra* note 87, but also how they feel about the intestate distribution. For example, the researcher could ask a follow-up question to respondents who deviate from the intestate pattern: “If the distribution you just selected is given a score of 100 on a 100-point scale where 100 is ‘exactly what you want’ and zero is ‘totally unacceptable,’ where does the following [intestate] distribution fall on that scale?” Answers could help generate a more refined estimate of *I*.

¹⁶³ To these three considerations, one could add consistency. If the case for the preponderance standard is relatively strong for replacement wills and codicils, consistency argues in favor of applying it generally. *See Pierce v. Pierce*, 127 S.W.2d 791, 794 (Tenn. 1939) (rejecting clear and convincing standard in favor of preponderance: “Recognition and application of exceptions to the fixed rules governing the proof required in civil and criminal cases, respectively, would give rise to great confusion.”). This argument should not prevail: the existence of a prior will is a bright-line rule and lawyers who practice in this area already have to be very sensitive to standards of proof. *See infra* text accompanying note 167 (undue influence split).

asymmetric risks of error.”¹⁶⁴ This article demonstrates that “unequivocal justification” is probably impossible to provide even in cases, like incapacity will contests, that feature asymmetric costs of error.

The conclusion should not be application of the preponderance standard across the board. Rather, unequivocal justification should not be required. Minimizing errors is an attractive goal, but only because doing so will generally minimize the cost of errors. It is the cost, not the error itself that matters.¹⁶⁵ Professor Stein would probably not disagree. He suggests that asymmetric error costs will at least frequently qualify as an unequivocal justification. Perhaps a better statement of the preference for the preponderance standard is as a default rule when error costs are nearly equal or unknown. Once asymmetric costs are shown, the factors outlined in this article should be weighed, with no strong thumb on either side of the scale.

V. EXTENSIONS

Three examples illustrate some broader implications of the foregoing analysis: undue influence, fraud, and criminal law. The closest analog to testamentary incapacity is undue influence, which along with incapacity is one of the two most common grounds for will contests.¹⁶⁶ Here too there is a split between jurisdictions requiring contestants to prove undue influence by a preponderance of evidence and other jurisdictions requiring clear and convincing

¹⁶⁴ ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 153 (2005). *Accord* Redmayne, *supra* note 51, at 174 (arguing that preponderance standard “should only be abandoned where the reasons for doing so are compelling”). Professor Stein elsewhere appears to back away somewhat from this statement. Richard A. Bierschbach & Alex Stein, *Overenforcement*, 93 GEO. L.J. 1743, 1758 (2005) (“[O]ur theory offers no concrete prescriptions for a perfectly tailored counterbalancing approach [like a heightened evidentiary requirement] that would do exactly right on a case-by-case basis. As a practical matter, this is unattainable.”).

¹⁶⁵ *Cf.* EPSTEIN, *supra* note 48, at 224.

¹⁶⁶ Schoenblum, *supra* note 128, at 648 chart 22 (49 out of 66 contests [74%] involved undue influence claim; 48 [73%] involved incapacity; the next most common claim, lack of testamentary formalities, was alleged in only 9 cases [14%]). Undue influence is a slippery concept, but has been defined as follows: “A donative transfer is procured by undue influence if the wrongdoer exerted such influence over the donor that it overcame the donor's free will and caused the donor to make a donative transfer that the donor would not otherwise have made.” RESTATEMENT, *supra* note 9, § 8.3(b).

evidence.¹⁶⁷ Substitute “probability of undue influence” for “probability of incapacity” and almost the entire analysis above applies. The most notable exception is that with undue influence, there is a wrongdoer who manipulates the content of the will. Recall how we tried to get a sense of how often wills executed by incapacitated testators reflected their intent by subdividing such “bad” wills into innocent mistakes and intentional malfeasance. With undue influence there are no innocent mistakes, so there is good reason to think that *B* is very, very close to zero. This weakens the case for the clear and convincing standard. It is therefore somewhat surprising that almost three times as many states (fourteen¹⁶⁸) apply the heightened standard to undue influence claims as to incapacity claims (five). In light of the overlap of issues, however, it is not surprising that all five of the states requiring clear and convincing evidence of incapacity require the same of undue influence.¹⁶⁹

A second example may shed light on the apparent anomaly that many states make it more difficult to prove undue influence than incapacity. Professor Alex Stein justifies application of the clear and convincing evidence standard to claims of fraud on the ground that an individual found to have committed fraud suffers a social sanction in addition to a legal one.¹⁷⁰ The heightened proof standard avoids overdeterrence by reducing the probability of a fraud finding.¹⁷¹ In our terms, Stein is arguing that the utility of true and false positives is lower for fraud cases than for other civil actions, because a fraud finding has the negative effect of overdeterrence. Equation (4) shows that the optimal proof standard necessarily rises as the

¹⁶⁷ ROSS & REED, *supra* note 14, § 7:12 (listing authority from 29 states following the preponderance rule; listing authority from 16 states requiring clear and convincing evidence).

¹⁶⁸ ROSS & REED, *supra* note 14, § 7:12, place Iowa and Ohio on both sides of the divide. The authority they cite puts both states in the preponderance camp. *Matter of Estate of Todd*, 585 N.W.2d 273 (Iowa 1998); *West v. Lucas*, 139 N.E. 859 (Ohio 1922).

¹⁶⁹ ROSS & REED, *supra* note 14, § 7:12, list Louisiana citations in support of both standards, but the controlling statute indicates that the state applies the clear and convincing evidence standard. La. Civ. Code Ann. art. 1483.

¹⁷⁰ STEIN, *supra* note 164, at 153. *Accord Addington v. Texas*, 441 U.S. 418, 424 (1979).

¹⁷¹ *Id.*

utility of positive findings declines, so Stein is on the right track. However, his account falls far short of a justification for the fraud standard.

Even within the standard statistical approach, one needs to know how large the overdeterrence effect is. If the utility of positive findings goes down only slightly, the preponderance standard may still do better than the clear and convincing evidence standard.¹⁷² Depending on the size of the overdeterrence effect, the total amount of fraud may be relevant. The greater the court's error rate, the bigger the effect would have to be to justify deviation from the preponderance standard. And all of this assumes an agreed-upon metric that measures deterrence along with all other purposes of the law of fraud. My own tentative view is that a heightened standard of proof for fraud is probably not justified.

But the fraud case sheds light on undue influence. Being found guilty of undue influence probably carries a social sanction much like being found guilty of fraud. With undue influence, however, avoiding overdeterrence is even more important since the boundary between good and bad behavior is blurrier. It is not wrong for a child to ask his parent to be remembered in her will, and some testators may need and want assistance from individuals who are also beneficiaries. The same concern with overdeterrence that arises in fraud cases is therefore magnified and weighs heavily in favor of a heightened standard for undue influence.¹⁷³

A criminal conviction could similarly be said to inflict social as well as legal harm, but the emphasis of those who defend the reasonable doubt standard is usually different: the cost of convicting an innocent person is greater than the cost of acquitting a guilty person.¹⁷⁴ Judge

¹⁷² In this sense, Stein fails his own “unequivocal evidence” test. *Id.* at 153. In fairness, Stein may have intended to offer an explanation rather than a full justification.

¹⁷³ Even without that concern, however, this article suggests that there is a reasonably strong case for the clear and convincing evidence standard for undue influence.

¹⁷⁴ Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 STAN. L. REV. 1477, 1504 (1999). Even those who criticize this rationale cite the costs of false convictions in favor of the high standard. See Richard Lempert, *The Economic Analysis of Evidence Law: Common Sense on Stilts*, 87 VA. L. REV. 1619, 1665 (2001)

Richard Posner attributes this disparity in part to the high cost of imprisonment.¹⁷⁵ This is obviously a direct application of the standard statistical approach. Posner (like Stein) makes no attempt to quantify the relative costs of errors. Some have argued that in the criminal context quantification of the utilities and of the standard itself are not possible.¹⁷⁶

Posner offers a second strand of argument in favor of the reasonable doubt standard: inequality of resources.¹⁷⁷ Because the government generally has such a huge advantage in resources, the high burden is needed to even up the playing field. Recall the discussion above of asymmetric costs in will contests (Section III.J). The proponent of the will has a resource advantage because the costs of litigation come out of the estate instead of out-of-pocket. This argues for a lower standard of proof to make the contestant's life easier. In criminal law, the situation is reversed: it is the prosecution that has superior resources, and this, argues Posner, militates in favor of requiring proof beyond a reasonable doubt.

And while the government's resources in any particular case will vastly outstrip those of an individual defendant, prosecutorial resources are insufficient to cover every crime. Prosecutors will therefore draw defendants from among the suspects most likely to be convicted if tried---in other words, most likely to be guilty.¹⁷⁸ In making this argument, Posner demonstrates sensitivity to importance of the underlying distribution and case selection. But his argument justifies a lower, not higher, standard of proof. This article demonstrated above that a leftward merits skew generally favors a higher standard of proof. All else held equal, a

(relying on "an aspiration to never convict innocent people and a moral judgment about the wrongfulness of inflicting the pain of criminal conviction on people who are not guilty of crimes").

¹⁷⁵ RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* § 22.4, at 648 (7th ed. 2007). Note, however, that the reasonable doubt standard applies to all crimes, not just those leading to imprisonment.

¹⁷⁶ *See, e.g.,* McCullough v. State, 657 P.2d 1157, 1159 (Nev. 1983) (Reasonable doubt is "inherently qualitative. Any attempt to quantify may impermissibly lower the prosecution's burden of proof and is likely to confuse rather than clarify."); Lillquist, *supra* note 47, at 91 ("the traditional understanding wrongly assumes that we know *how* to weigh these costs and benefits").

¹⁷⁷ Posner, *supra* note 174, at 1505.

¹⁷⁸ POSNER, *supra* note 175, § 22.4, at 649.

rightward skew like the one Posner identifies should have the opposite impact on selection of the proof standard.

Posner does not discuss the potential for a high proof standard to reduce the probability of conviction and thus the disincentive to commit crime. This would be analogous to the incentive effects identified above: greater security from the clear and convincing evidence standard should mean more wills and wills executed later in life. Posner's omission is particularly remarkable because in the same section he argues that pro-defendant Supreme Court precedent on other issues "would have reduced the expected cost of punishment, and so driven crime rates even higher."¹⁷⁹ In fairness, Posner may believe that the resource constraint is so severe that it drives the conviction rate independent of the standard of proof.

Conclusion

This article posed a very specific question: in will contests alleging incapacity, which standard of proof---preponderance or clear and convincing evidence---is better? There were reasons to think a definitive answer was attainable: (1) there is basic agreement that advancing testator intent is the primary goal of the law of wills; (2) studies have examined how well the fallback system of intestacy matched testators' intent; and (3) the base rate of incapacity is almost certainly low. But these facts were not enough. The final answer was necessarily equivocal, but exposed previously unrecognized factors that are relevant in selecting any standard of proof.

Perhaps most important is sensitivity to the effects of settlement. One cannot simply assume that settlements will mirror trial outcomes. The asymmetric cost structure of will contests is just one example of why a myopic focus on trial may be inadequate. So too some understanding of the merits distribution can affect choice of proof standard. As in will contests,

¹⁷⁹ *Id.* § 22.4, at 650.

it may not be necessary to know the precise distribution, but simply to know the direction of the likely skew. And while there were difficulties measuring the utilities of the four outcomes, this article suggests that that basic problem may not be intractable at least in some contexts. A final, important thread of the argument of this article is that proof standard instructions should be clarified. The whole question of which standard to apply is almost meaningless if jurors do not differentiate among the standards, as appears presently to be the case. More precise language may suffice, but a numerical approach deserves serious consideration.

Perhaps the biggest contribution of this article is to present a more complete framework in which to evaluate proof standards. Leading commentators have sometimes recognized various of the factors, but more often have failed to examine critical assumptions or even cited factors in favor of one proof standard that actually push for another. To be fair, precision of result may necessarily be elusive in some areas, but precision of analysis is not. And while the framework is presented in probabilistic terms, the lessons learned are applicable to other approaches to proof standards.