Reasonable doubt revisited

ELIAS TSAKAS†

Maastricht University

First draft: March, 2016
This draft: February, 2017

Abstract

Choice rules based on probability thresholds are common in several disciplines. The most well-known application of such a threshold rule is the standard of reasonable doubt. Accordingly, a rational juror prefers to convict a defendant if and only if the probability that she attaches to the defendant being guilty is above a given threshold. In this paper we prove that generically such a threshold exists if and only if the juror reasons only about two events, viz., the defendant’s guilt and innocence. This result implies that threshold rules are usually inconsistent with individual rationality. Thus, if we insist on using a threshold choice rule, we will have to accept some irrational convictions (false negatives) or some irrational acquittals (false positives) or both. We subsequently characterize each probability threshold in terms of the irrationalities that it induces. Finally, we discuss the empirical implications of our theory.

1. Introduction

Choice rules based on probability thresholds are widely used in various disciplines, including law (e.g., Kaplan, 1968; Tribe, 1971; Schauer and Zeckhauser, 1996; Kaplow, 2012; Talley, 2013), medicine (e.g., Pauker and Kassirer, 1975, 1980), economics (e.g., Shavell, 1985; Andreoni, 1991; Kaplow, 2011), statistics (e.g., Neyman and Pearson, 1933) and finance (e.g., Roy, 1952; Telser, 1955-56). For example in law, a juror should convict the defendant if and only if the probability that she assigns to him being guilty is above a certain threshold (Kaplan, 1968); in medicine, a doctor administers a treatment to the patient if and only if the probability that she attaches to him suffering from a specific disease is above a given threshold (Pauker and Kassirer, 1975); in finance, an investment is admissible if and only if the probability of the returns being below some fixed level (e.g., the bankruptcy level) does not exceed a certain threshold (Telser, 1955-56). A natural question arises then in each of the previous contexts: how do we choose the threshold?

In this paper, we address this question within the context of law. The reason for focusing primarily on legal decisions is twofold. Firstly, the foundations of probability thresholds have been mostly

---

*I am indebted to Jing Chen, Judge Georgios Alexopoulos, Franz Dietrich, Frederik Herzberg, Bram van Hofstraeten, Anqi Li, Arkadi Predtetchinski, Philipp Strack, Peter Wakker, Chris Woolnough, Dimitrios Xeftoris and the seminar participants in the University of Amsterdam, UC Berkeley, University of Cyprus and Maastricht University for valuable comments and helpful discussions.

†Department of Economics (AE1), Maastricht University, P.O. Box 616, 6200 MD, Maastricht, The Netherlands; Homepage: http://www.elias-tsakas.com/home.html; E-mail: e.tsakas@maastrichtuniversity.nl

1For a discussion of the relation between threshold rules in different disciplines, see Saks and Neufeld (2011, 2012).
discussed within the law literature. Secondly, in law the use of probability thresholds is normatively postulated – by the lawmaker – contrary to other disciplines where threshold rules are simply strategies willingly chosen by the decision maker. In this sense, answering the previous question has important practical implications for the interpretation of the law. In either case, our conclusions hold across the aforementioned fields, thus constituting a general theory of threshold choice rules.

Let us begin by noticing that while the law suggests the use of some probability threshold, it does not specify a precise value and instead leaves it to the juror’s discretion. That is, the law implicitly postulates a threshold rule which is consistent with the juror’s individual rationality, i.e., ideally, the juror prefers to convict the defendant if and only if the probability she attaches to guilt is above this threshold. Such a threshold is called the standard of reasonable doubt (e.g., Kaplan, 1968). Obviously, if it exists, the standard of reasonable doubt is the answer to our previous question.

Thus, we now ask: does the standard of reasonable doubt always exist? In other words, is there always a rational threshold rule? As it turns out, this is the case only under very stringent conditions. In particular, we prove that generically the standard of reasonable doubt exists if and only if the only event that the juror reasons about is the defendant’s guilt (Theorem 1). That is, if the juror reasons about the circumstances under which the crime was committed – and not just about whether the defendant committed it or not – her preferred verdict will in fact also depend on her beliefs about these circumstances, rather than solely on the probability she attaches to the defendant being guilty. Therefore, threshold choice rules are in general not rational.

Before moving forward, let us first elaborate on the main idea behind the previous result. We consider an underlying set of states, each corresponding to a different configuration of the world. An event is identified by a subset of the state space, and the juror reasons about certain events while ignoring the remaining ones, either because she is unaware of them or because she consciously disregards them. Mathematically, the events that the juror reasons about form an algebra, called the juror’s frame. Naturally, the juror assigns probabilities only to events in her frame. Thus, her set of all possible beliefs is represented by an n-dimensional simplex, with n being the cardinality of the partition that generates her frame. Notice that the set of beliefs that make the juror prefer a conviction are identified by a half-space, and therefore the only way a rational choice can be determined by the probability attached to a single event (viz., “guilt”) is if the dimension of the aforementioned simplex is $n = 2$, i.e., if the juror’s frame contains only two events (viz., “guilt” and “innocence”).

The previous impossibility result has important implications for all fields where probability thresholds are employed. In disciplines where threshold choice rules are simply strategies willingly chosen by the decision maker – e.g., in medicine or statistics or finance – new, more complicated, strategies need to be introduced, if our priority is to maintain rationality on the decision-maker’s part. On the other hand, in disciplines where the use of a probability threshold is normatively postulated – e.g., in law – we will have to accept the possibility of some irrationalities. In particular, every threshold will lead either to some irrational convictions (false negatives), or to some irrational acquittals (false positives), or to both. Hence, the selection of a probability threshold essentially depends on our attitude towards irrationalities, e.g., extreme aversion to irrational acquittals would prompt us to choose a low threshold,

---

2 The study of probability thresholds within the law theory and practice dates all the way back to Locke (1690) (for an overview see Hemmens et al., 1997). To the best of our knowledge, the first one to formalize this idea using decision-theoretic tools was Professor Kaplan (1968).

3 The practical importance of understanding probability thresholds in law is illustrated by the fact that approximately 29% of adult Americans have served as trial jurors at least once in their lifetime, while 71% of the criminal defendants in the U.S. are convicted by a jury trial (U.S. Department of State, 2009).

4 Later in the paper we become more precise on the notion of genericity that we employ. In fact, we prove a slightly stronger version of the result, which provides necessary and sufficient conditions for the existence of the standard of reasonable doubt also non-generically.

5 Notice that our definition of false positives and false negatives is in expectation (see Remark 2).
whereas extreme aversion to irrational convictions would prompt us to choose a high standard. While
in this paper we do not formally model such preferences, we nonetheless informally assume aversion for
irrationalities in general. That is, if one threshold leads to strictly fewer irrationalities than another
one, we say that the former dominates the latter.\(^6\) Then, aversion to irrationalities implies that the
juror will never pick a dominated threshold, and the thresholds that are not dominated are called weak
standards of reasonable doubt.

Subsequently, we characterize the weak standards of reasonable doubt, showing that they form a
specific subinterval in \([0, 1]\) (Theorem 2). The lower bound of the interval (lower standard of reasonable
doubt) is always strictly larger than 0, and corresponds to the only undominated threshold inducing
extreme aversion to irrational acquittals (see Proposition 2). On the other hand, an undominated
threshold inducing extreme aversion to irrational convictions exists if and only if the juror prefers to
convict the defendant at all states where the defendant is guilty, irrespective of the circumstances under
which the crime was committed. This is the case when the upper bound of the interval (upper standard
of reasonable doubt) is lower than 1 (see Proposition 3). In fact, the interval collapses to a single point
– with the upper and the lower standard coinciding – if and only if the standard of reasonable doubt
exists, i.e., if and only if the juror reasons only about the defendant’s guilt/innocence, in which case
no irrationalities are induced (see Proposition 1).

The main idea behind introducing two distinct standards – an upper and a lower one – resembles one
that first appeared in medicine (Pauker and Kassirer, 1980). Accordingly, there are two probability
thresholds: the doctor administers the treatment if the probability of the patient suffering from a
certain disease is above the upper threshold, she does not administer it if the probability is below the
lower threshold, and she runs additional tests if the probability is in between. While the interpretation
of the respective extreme thresholds is different, the common denominator of the two models is that
the upper standard yields strong aversion to false negatives whereas the lower standard yields strong
aversion to false positives (see Remark 3).

Overall in the light of our results, the notion of certainty beyond reasonable doubt can now be
reinterpreted. In particular, one can replace certainty in the defendant’s guilt (viz., \(ex\ post\) extreme
aversion to false negatives) with certainty of avoiding irrational convictions (viz., \(ex\ ante\) extreme
aversion to false negatives), i.e., formally instead of setting the probability threshold to 1, we set
it equal to the upper standard of reasonable doubt. This interpretation maintains the conventional
wisdom that the burden of proof is high without requiring absolute certainty of guilt (for an overview
of this discussion, see Hemmens et al., 1997, and references therein). At the same time, our new
interpretation is consistent empirical evidence (see discussion in Section 5).

The paper is structured as follows: In Section 2 we introduce our framework and define the standard
of reasonable doubt. Section 3 contains our main impossibility result. In Section 4 we define the weak
standards of reasonable doubt and we prove our positive results. In Section 5 we discuss the empirical
implications of our theory. Section 6 contains a concluding discussion. All proofs are relegated to the
Appendix.

2. The standard of reasonable doubt

There are two agents, a (female) juror and a (male) defendant. Let \(\Omega\) be a finite state space.\(^7\) Each
state \(\omega \in \Omega\) is a full description of all the relevant aspects of the world. Let \(G \subseteq \Omega\) be the event
that the defendant is guilty of the crime he is accused for, with the complement \(I := \neg G\) denoting
the event that he is innocent. Note that \(G\) is a coarse description of the world, in the sense that it

---

\(^6\)In Section 4 we precisely define what it means for a threshold to induce “fewer irrationalities” than another one.
\(^7\)Our analysis can be directly extended to any measurable state space.
contains different specifications of how the crime could have taken place, e.g., there are different ways
of committing a crime, differing for instance in the defendant’s intentions or in the degree of cruelty
involved. In either case, we assume that the law is detailed enough to clearly specify at which states
the defendant is considered guilty (resp., innocent). Thus, the events $G$ and $I$ have a well-defined
interpretation in the natural language.

An algebra $\mathcal{R}$ of subsets of $\Omega$ is called a (reasoning) frame, and contains the events that the juror
reasons about at the time of her decision.\(^8\) Events outside $\mathcal{R}$ are not even considered by the juror,
either because she is unaware of them or because she consciously disregards them. In this sense, the
frame $\mathcal{R}$ can also be seen as the juror’s working language. We naturally assume that $G$ is always $\mathcal{R}$-
measurable, and a fortiori so is $I$, i.e., the juror always reasons about the defendant’s guilt/innocence.
In fact, if $G$ and $I$ are the only events that she reasons about, her frame $\mathcal{R}$ collapses to the coarsest
possible frame $\mathcal{G} := \{\Omega, G, I, \emptyset\}$, henceforth called the trivial frame.\(^9\)

**Example 1.** Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $G = \{\omega_1, \omega_2\}$. Assume that the juror reasons only about the
defendant’s guilt/innocence, i.e., $\mathcal{R} = \mathcal{G} = \{\Omega, \{\omega_1, \omega_2\}, \{\omega_3\}, \emptyset\}$. In particular, the two states in $G$
differ in the defendant’s intention to commit the crime, viz., at $\omega_1$ the defendant committed the crime
unintentionally, whereas at $\omega_2$ he did it intentionally. However, the juror does not reason about the
defendant’s intentions, i.e., $\{\omega_1\} \notin \mathcal{R}$ and $\{\omega_2\} \notin \mathcal{R}$. This is exactly the framework considered by
most papers in the literature (e.g., Kaplan, 1968; Andreoni, 1991).

Let $X$ be the set of all possible verdicts that are allowed by the law. For convenience and without
loss of generality we take a finite $X \subseteq [0, \infty]$, with $0 \in X$ being the verdict that acquits the defendant
and $X_+ := X \setminus \{0\}$ being the set of verdicts that convict him. In fact each verdict in $X$ can be formally
viewed as a constant degenerate act in $\mathcal{F} := (\Delta(X))^\Omega$.

Given a frame $\mathcal{R}$, we let $U^\mathcal{R} : X \times \Omega \to \mathbb{R}$ be the juror’s state-dependent utility function, where
the random variable $U^\mathcal{R}_x := U^\mathcal{R}(x, \cdot)$ is assumed to be $\mathcal{R}$-measurable for every $x \in X$. Throughout the
paper, for notation simplicity, we omit the superscript $\mathcal{R}$, thus simply writing $U$ and $U_x$, respectively.
Moreover, the juror forms a subjective belief $\pi \in \Delta(\Omega, \mathcal{R})$, where as usual $\Delta(\Omega, \mathcal{R})$ denotes the set of
all probability measures over the measurable space $(\Omega, \mathcal{R})$. Then, her preferences over $X$ (given her
frame $\mathcal{R}$) are represented by the state-dependent expected utility (SDEU) function,

\[
\mathbb{E}_\pi U_x = \int_\Omega U_x d\pi. \quad (1)
\]

There are various axiomatizations of SDEU functions in the literature, both within the Savage and
the Anscombe-Aumann framework (Fishburn, 1973; Karni et al., 1983; Karni, 1993a,b).\(^10\)

For an arbitrary $x \in X_+$, we introduce the auxiliary $\mathcal{R}$-measurable random variable

\[
V_x := U_x - U_0, \quad (2)
\]

and we impose the following natural assumption:

\[ (A_0) \text{ For every } x \in X_+, \text{ we let } V_x(\omega) < 0 \text{ for every } \omega \in I. \]

\(^8\)In this paper we are not interested in the juror’s reasoning process, and therefore we do not model how her frame
evolves throughout the trial before eventually converging to $\mathcal{R}$.

\(^9\)Our concept of the trivial frame should not be confused with the trivial algebra $\{\Omega, \emptyset\}$.

\(^10\)In general, in SDEU models, the beliefs are not identified uniquely from the preferences over $\mathcal{F}$. Thus, some
additional structure needs to be imposed in order to be able to identify the beliefs. In fact, the different axiomatizations
in the literature differ from each other in the additional structure they impose.
The interpretation is straightforward, viz., the juror prefers acquitting an innocent defendant over convicting him, irrespective of the circumstances or the magnitude of the sentence. Notice that we do not require the juror to necessarily prefer to convict a guilty defendant, as this may depend on the precise circumstances or on the magnitude of \( x \). The sentence \( x \in X_+ \) is said to be trivial if either \( V_x \geq 0 \) or \( V_x \leq 0 \). It is nontrivial otherwise. Obviously, by \( (A_0) \), it cannot be the case that \( V_x \geq 0 \), and therefore \( x \in X_+ \) is nontrivial if and only if there exists some \( \omega \in G \) such that \( V_x(\omega) > 0 \). We find it uninteresting to study trivial sentences and thus we focus exclusively on nontrivial ones.

**Remark 1.** While our work is related to several papers on frame-dependent preferences (e.g., Ahn and Ergin, 2010; Karni and Vierø, 2013; Schipper, 2013), it should not be seen as part of this literature. The reason is that these representations simultaneously look at the juror’s preferences across the different frames (ex ante stage) as well as given some fixed frame (interim stage), whereas we only look at the interim stage.\(^{12,13}\) Indeed, as we have already mentioned, we do not formally model the frame-formation process, but rather we fix the juror’s frame to be the one she has at the time of the decision. In this sense a standard SDEU model suffices to represent the underlying preferences given the fixed frame \( \mathcal{R} \).

A *decision problem* is a nonempty set of verdicts \( \Gamma \subseteq X \) with \( 0 \in \Gamma \), among which the juror chooses one. Throughout this paper we mostly focus on binary decision problems, thus implicitly assuming that – in case of a guilty verdict – the sentence has been determined exogenously, e.g., by the lawmaker.\(^{14}\) A *choice rule* in the decision problem \( \Gamma \subseteq X \) is a strategy \( \sigma : \Delta(\Omega, \mathcal{R}) \rightarrow \Gamma \) that prescribes a choice for each of the juror’s beliefs.

For a binary decision problem \( \Gamma = \{0, x\} \) and an arbitrary probability threshold \( p \in [0, 1] \), the *threshold (choice) rule* \( \sigma_p : \Delta(\Omega, \mathcal{R}) \rightarrow \Gamma \) prescribes conviction if the juror’s belief belongs to

\[
D_p := \{ \pi \in \Delta(\Omega, \mathcal{R}) : \pi(G) \geq p \},
\]

and prescribes acquittal otherwise. That is formally, \( \sigma_p \) is defined by

\[
\sigma_p(\pi) := \begin{cases} 
  x & \text{if } \pi(G) \geq p, \\
  0 & \text{if } \pi(G) < p.
\end{cases}
\]

As we have already mentioned, threshold rules have been extensively studied in the context of law (e.g., Kaplan, 1968; Kaplow, 2012), economics (e.g., Andreoni, 1991; Kaplow, 2011), statistics (e.g., Neyman and Pearson, 1933), medicine (e.g., Pauker and Kassirer, 1975, 1980) and finance (e.g., Roy, 1952; Telser, 1955-56). Moreover, they have been used as an auxiliary tool in applications and examples within elsewhere-focused papers (e.g., Feddersen and Pesendorfer, 1998; Kamenica and Gentzkow, 2011). Sometimes a threshold rule is viewed as a suggestion postulated by an institution (e.g., by the law), whereas in other cases it is merely a strategy consciously chosen by the decision maker (e.g., in medicine or in finance). Here we should stress that whenever we say that the *use of a probability threshold is postulated by the law*, we do not mean that the lawmaker necessarily pins down a threshold, but rather that the lawmaker requires a threshold to be specified (e.g., by the juror herself).

11 As usual, for an arbitrary random variable \( Y : \Omega \rightarrow \mathbb{R} \), we write \( Y \geq 0 \) whenever \( Y(\omega) \geq 0 \) for all \( \omega \in \Omega \).
12 All these papers start from a collection of preference relations, one for each frame \( \mathcal{R} \), and they first obtain an expected utility representation conditional on each given frame. Subsequently, they study the relationship between the different frame-dependent representations.
13 A second difference – of relatively minor importance – between our setting the ones in all the aforementioned papers is that, we allow for a state-dependent utility function (given each frame).
14 This last assumption is removed in Section 6.2, where we consider decision problems \( \Gamma \) with \( |\Gamma| > 2 \).
A choice rule is said to be *rational* if it prescribes a rational choice to every belief, i.e., formally \( \sigma \) is rational in \( \Gamma \), if \( \sigma(\pi) \in \arg \max_{y \in \Gamma} E_{\pi} U_y \) for all \( \pi \in \Delta(\Omega, R) \). Specifically in a binary decision problem \( \Gamma = \{0, x\} \), the threshold rule \( \sigma_p \) is rational if it prescribes conviction to every belief in \( C_x := \{ \pi \in \Delta(\Omega, R) : E_{\pi} V_x \geq 0 \} \), \(^{(5)}\) and prescribes acquittal otherwise, i.e., formally \( \sigma_p(\pi) = x \) if and only if \( \pi \in C_x \).

**Definition 1.** We say that \( p_x \in [0, 1] \) is the *standard of reasonable doubt* for \( x \in X_+ \), if

\[
C_x = D_{p_x}. \tag{6}
\]

Notice that the standard of reasonable doubt is not a choice rule, but rather a probability threshold such that, *the juror prefers to convict the defendant* \( (E_{\pi} V_x \geq 0) \) *if and only if the probability that she attaches to him being guilty is above this threshold* \( (\pi(G) \geq p_x) \).

**Example 1 (continued).** Recall our example with \( \Omega = \{\omega_1, \omega_2, \omega_3\} \), \( G = \{\omega_1, \omega_2\} \) and \( R = G \). Moreover, let the decision problem be \( \Gamma = \{0, 1\} \) and the utility function be

\[
U_x(\omega) = \begin{cases} 
-x^2 + 2x - 1 & \text{if } \omega \in G, \\
-x + 1 & \text{if } \omega \in I,
\end{cases}
\]

which is obviously \( R \)-measurable. Then, observe that \( C_1 = D_{1/2} \), thus implying that \( p_1 = 1/2 \) is the standard of reasonable doubt for \( x = 1 \). Hence, if the juror follows the threshold rule \( \sigma_{1/2} \), her choice will always be rational, irrespective of her belief. \( \triangle \)

Obviously, if the standard of reasonable doubt exists, then it is unique.\(^{15}\) On the other hand, if it does not exist, no threshold rule is consistent with rationality. This last case is particularly interesting when the use of a probability threshold is postulated by some institution, e.g., by the law. Then, non-existence of the standard of reasonable doubt directly implies that the juror’s choice will definitely be irrational, at least for some beliefs.

### 3. Existence of the standard of reasonable doubt

So far, we have defined the standard of reasonable doubt and we have introduced the corresponding choice rule on the basis of it, but we have not answered the most fundamental question, viz., *does \( p_x \) always exist?* In other words, *is there some threshold \( p \) so that the choice rule \( \sigma_p \) is rational?* As it turns out, such threshold exists only under very stringent conditions.

**Theorem 1.** The standard of reasonable doubt \( p_x \) exists if and only if \( V_x \) is \( G \)-measurable.

Let us now discuss the implications of the previous theorem, focusing on two fundamental questions that our result answers. First, *when does the standard of reasonable doubt actually exist?* And second, *when it does not exist, what happens if we still use a threshold choice rule?*

We begin with the first question. Notice that Theorem 1 essentially says that the standard of reasonable doubt does not exist, unless either (i) the juror reasons only about the defendant’s guilt/innocence and nothing else, or (ii) she reasons about additional events which however she finds irrelevant for her

\(^{15}\)Moreover, it is straightforward to show that, if \( p_x \) exists then it is necessarily the case that \( p_x \in (0, 1) \) (see bottom of the proof of Theorem 1 in the Appendix).
decision. In fact, the previous two conditions cannot be identified from the juror’s preferences (Schipper, 2013).\textsuperscript{16} In either case, when the first condition is violated (i.e., when the juror reasons about events outside the trivial frame $\mathcal{G}$), the second condition is non-generic (i.e., the set of utility functions that yield a $\mathcal{G}$-measurable $V_x$ is Lebesgue-null). Therefore, \textit{generically, the standard of reasonable doubt exists if and only if the juror reasons only about events in $\mathcal{G}$}. In this sense – and given that in real life we expect most jurors to reason about events outside $\mathcal{G}$ – it is justified to call Theorem 1 an \textit{impossibility result}.

Now, let us turn to our second question. Recall from the previous section that, when the standard of reasonable doubt does not exist, every threshold rule is irrational for at least some beliefs. Thus, one of the following three scenarios necessarily holds: the defendant is irrationally acquitted for some beliefs (\textit{false positive}), or the defendant is irrationally convicted for some beliefs (\textit{false negative}), or both. Which is the case, depends on the choice of the threshold. In either case, the bottom line is that the juror will probably find herself in a difficult situation, where her own preferences will be in conflict with the threshold rule.

\textbf{Remark 2.} Our notions of “\textit{false positive}” and “\textit{false negative}” are slightly different from the usual type I and type II errors. In particular, our terms refer to (ex ante) mistakes made by the juror \textit{in expectation}, and not to mistakes in the sense of (ex post) wrongful choices. \hfill \textdegree

\textbf{Example 2.} Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$, with $G = \{\omega_1, \omega_2\}$ and $F = \{\omega_2\}$ denoting the event that the defendant is guilty and the event that he intended to commit the crime, respectively. Furthermore, assume that the juror reasons about every event in the power set of $\Omega$, i.e., $\mathcal{R} = \sigma(\{F, G\}) = 2^\Omega$. Moreover, let the decision problem be $\Gamma = \{0, 4\}$ and the utility function be

$$U_x(\omega) = \begin{cases} -2x^2 + 10x & \text{if } \omega \in \{\omega_1\}, \\ 10x^2 - 2x & \text{if } \omega \in \{\omega_2\}, \\ -x & \text{if } \omega \in \{\omega_3\}, \end{cases}$$ \hfill (7)

Then, observe that $C_4 \neq D_p$ for every $p \in [0, 1]$, thus implying that there is no standard of reasonable doubt. Of course, this is obvious given our Theorem 1, as it is straightforward to verify that $V_4$ is not $\mathcal{G}$-measurable. Let us also illustrate it graphically (on Figure 1). First, note that the shaded area contains the beliefs in $C_4$, whereas the area above the dashed line contains the beliefs in $D_p$, for each $p \in [0, 1]$. Now, observe that the straight lines that are associated with $C_4$ and $D_p$ are not parallel to each other and therefore the two areas will not coincide for any $p \in [0, 1]$. As a consequence, the juror prefers to convict the defendant under the belief $\pi_1$ and acquit him under $\pi_2$, even though $\pi_2(G) > \pi_1(G)$.\textsuperscript{17} Hence, there is no threshold rule which is always rational. Indeed, for every $p \in [0, 1]$ the threshold rule $\sigma_p$ will either induce an irrational acquittal (e.g., under $\pi_1 \in C_4 \setminus D_p$ when $p = 0.87$), or an irrational conviction (e.g., under $\pi_2 \in D_p \setminus C_4$ when $p = 0.21$), or even both (under $\pi_1 \in C_4 \setminus D_p$ and under $\pi_2 \in D_p \setminus C_4$ when $p = 0.50$). That is, high thresholds lead to false positives, low thresholds to false negatives, and intermediate thresholds to both. \hfill \textdegree

\textsuperscript{16}Of course, Schipper (2013) considers a framework with state-independent preferences, but nevertheless the idea is the same, i.e., the preferences over acts would be the same in the two aforementioned cases, and therefore by simply looking at the juror’s choices we cannot tell which of the two holds when the standard of reasonable doubt exists, while if the standard does not exist both are violated.

\textsuperscript{17}Intuitively, this is because under $\pi_1$ she deems much more likely (than under $\pi_2$) that the defendant committed the crime intentionally, conditional on him being guilty, i.e., $\pi_1(\omega_2|G)$ is much larger than $\pi_2(\omega_2|G)$. 


Figure 1: Nonexistence of the standard of reasonable doubt.

4. Towards resolving the impossibility result

The main implication of our Theorem 1 is that, generically, if the juror reasons about events outside $\mathcal{G}$, no threshold rule is rational. Thus, the obvious problem arises in cases where the use of probability thresholds is postulated by an institution, like it is for instance the case in legal systems. Namely, we ask: if we must pick one of the irrational threshold rules, which one shall we choose?\(^{18}\)

4.1. Weak standards of reasonable doubt

The answer to the previous question is far from obvious, as it depends on the society’s preferences for irrationalities (i.e., for false positives and false negatives), which is not modelled in our case. Nevertheless, even without explicitly introducing such a preference relation, if we simply assume that the society has aversion to irrationalities, we can rule out some clearly undesirable and counter-intuitive thresholds, using a dominance type of argument. Intuitively, if one threshold rule leads to fewer irrationalities than another one, then the former rule dominates the latter, which is then eliminated. The thresholds that survive elimination are henceforth called weak standards of reasonable doubt and are formally defined below.

**Definition 2.** We say that $p_x^w \in [0,1]$ is a weak standard of reasonable doubt for $x \in X_+$, if

$$\max\{0,p_x^\ell\} \geq p_x^w \geq \min\{p_x^u,1\},$$

where

$$p_x^u := \min\{p \in [0,1] : C_x \supseteq D_p\}$$

is the upper (weak) standard of reasonable doubt and

$$p_x^\ell := \max\{p \in [0,1] : C_x \subseteq D_p\}$$

is the lower (weak) standard of reasonable doubt.\(^{19}\)

\(^{18}\)No such question arises when we are not restricted to using threshold rules. In such case, we can replace threshold rules with more complicated choice rules, designed to minimize – or even avoid – irrationalities.

\(^{19}\)The reason we use the minimum (resp., maximum) instead of the infimum (resp., supremum) is that in our case the two coincide, due to the fact that $\{p \in [0,1] : C_x \supseteq D_p\}$ (resp., $\{p \in [0,1] : C_x \subseteq D_p\}$) is closed.
Before moving forward, let us first point out that whenever the upper standard and the lower standard exist, it will always be the case that \( p_x^u \geq p_x^l \), with equality holding if and only if \( p_x \) exists (see Proposition 1 below). Later in the paper, we also provide necessary and sufficient conditions for the existence of our two extreme standards, \( p_x^l \) and \( p_x^u \) (see Propositions 2 and 3 respectively).

Now, the first obvious task is to formally describe the – so far informally stated – dominance concept, and to show that our weak standards are indeed exactly those that survive elimination. For an arbitrary \( x \in X_+ \), and an arbitrary \( p \in [0, 1] \) we define the set of beliefs for which the threshold rule \( \sigma_p \) prescribes a rational choice, by

\[
R_x^p := \{ \pi \in \Delta(\Omega, \mathcal{R}) : \sigma_p(\pi) \in \mathop{\arg\max}_{y \in (0, x]} \mathbb{E}_\pi U_y \}. \tag{11}
\]

Obviously, the set of beliefs in \(-R_x^p\), for which \( \sigma_p \) prescribes an irrational choice, can be partitioned into those inducing a false positive \( P_x^p := C_x \setminus D_p \) and those inducing a false negative \( N_x^p := D_p \setminus C_x \) (e.g., see Example 2 in the previous section).

We say that the threshold rule \( \sigma_p \) dominates the threshold rule \( \sigma_{p'} \) (or simply, \( p \) dominates \( p' \)) whenever \( R_x^p \supseteq R_x^{p'} \), i.e., whenever \( \sigma_p \) induces fewer irrationalities than \( \sigma_{p'} \). In these cases \( \sigma_{p'} \) is called dominated, and it is therefore eliminated.

**Theorem 2.** The threshold \( p \) is a weak standard of reasonable doubt if and only if \( \sigma_p \) is not dominated.

Now, let us provide some additional intuition for our weak standards. The upper standard of reasonable doubt is the only undominated threshold that never induces irrational convictions but may induce irrational acquittals. In this sense, \( p_x^u \) implicitly postulates extreme aversion to false negatives. On the other hand, the lower standard of reasonable doubt is the only undominated threshold that never induces irrational acquittals but may induce irrational convictions. Thus, \( p_x^l \) implicitly postulates extreme aversion to false positives. All other weak standards will induce both false negatives and false positives. In fact, while moving our probability threshold upwards along the interval \((p_x^l, p_x^u)\), we will obtain fewer irrational convictions and more irrational acquittals. Thus, in order to choose one of the weak standards of reasonable doubt, we need to impose additional structure on the society’s preferences for irrationalities, i.e., we need to make the trade-off between false positives and false negatives explicit, which we do not formally do in this paper.

**Example 2 (continued).** Recall the example from the previous section, now depicted on Figure 2. As we have already discussed, the standard of reasonable doubt does not exist, and therefore no threshold rule is rational. It is rather straightforward to see that \( p_x^u = 0.83 \) and \( p_x^l = 0.21 \) are the upper and the lower standard of reasonable doubt respectively. Indeed, \( \sigma_{p_x^u} \) induces irrational acquittals (viz., for beliefs in the shaded area below the upper dashed line), but no irrational convictions (viz., all beliefs in the blank area lead to an acquittal). Likewise, \( \sigma_{p_x^l} \) induces irrational convictions (viz., for beliefs in the blank area above the lower dashed line), but no irrational acquittals (viz., all beliefs in the shaded area lead to a conviction). Obviously, probability thresholds \( p > p_x^u \) still lead to no irrational convictions, but lead to strictly more irrational acquittals, and therefore \( p \) is dominated by \( p_x^u \). Likewise, probability thresholds \( p < p_x^l \) still lead to no irrational acquittals, but lead to strictly more irrational convictions, and therefore \( p \) is dominated by \( p_x^l \). On the other hand, every \( p_x^u \in (p_x^l, p_x^u) \) (e.g., \( p_x^u = 0.50 \)) leads to fewer irrational acquittals and more irrational convictions than \( p_x^l \), and it also leads to fewer irrational convictions and more irrational acquittals than \( p_x^l \), thus illustrating the trade-off between false positives and false negatives that emerges when we compare undominated probability thresholds. \( \triangledown \)

**Remark 3.** The underlying idea behind introducing our weak standards is similar to one that first appeared in medicine (Pauker and Kassirer, 1980). Accordingly, instead of defining a single probability threshold, Pauker and Kassirer (1980) define two distinct thresholds, a lower one and an upper one.
Then, the treatment is administered if the probability of the patient suffering from the disease is above the upper threshold, it is not administered if the probability is below the lower threshold, while further tests are run if it lies between the two. Of course, the precise interpretation of our upper and lower standard is different in the context of law, where there is no possibility of the juror requesting additional evidence once the trial process has been completed, but nevertheless the main idea is strikingly similar. In particular, Pauker and Kassirer’s (1980) upper threshold exhibits strong aversion to false negatives (similarly to our \( p^u_x \)), while their lower threshold induces strong aversion to false positives (similarly to our \( p^l_x \)). Thus “running further tests” in their framework can be seen as a process of evaluating the trade-off between false positives and false negatives.

Now, going back to our original motivation for introducing the weak standards, the main idea was to identify an irrational threshold rule when the standard of reasonable doubt does not exist. However, our definition of the weak standards is also valid when the standard of reasonable doubt does exist. In this last case, it is natural to ask what the relationship between the different standards of reasonable doubt is. Indeed, the following result shows that in this case all standards coincide, i.e., \( p^u_x = p^l_x = p_x \).

**Proposition 1.** The standard of reasonable doubt \( p_x \) exists if and only if \( p^u_x = p^l_x \).

**Remark 4.** The previous analysis provides an answer to a long-standing debate within the law literature, between those in favor of explicitly specifying the standard of reasonable doubt (Kaplan, 1968) and those against it (Tribe, 1971).\(^{20}\) In particular, according to the previous proposition, there exist multiple (undominated) thresholds, whenever the juror reasons about events outside \( \mathcal{G} \). Thus, it is unclear which one the lawmaker had in mind, thus making it virtually impossible to pin down a unique threshold. The irony in our proposed resolution is that we reach the same conclusion as Professor Tribe using decision-theoretic tools, even though his argument (against specifying the standard) is partly based on the overall non-suitability of decision theory for court decisions.

### 4.2. Existence of the weak standards of reasonable doubt

Finally, let us turn our attention to existence, focusing on the two extreme standards. Starting with the lower standard, the following result proves that it always exists, i.e., we can always find an undominated threshold rule that postulates extreme aversion to irrational acquittals (false positives).

\(^{20}\)For a detailed comparative analysis of the two views, see Milanich (1981). For a more recent discussion on whether and how the standard of reasonable doubt should be quantified, see Newman (2006) and references therein.
Proposition 2. The lower standard of reasonable doubt \( p^l \) always exists.

Now let us turn our attention to the upper standard. The following result shows that the upper standard exists if and only if the juror prefers to always convict a guilty defendant, irrespective of the circumstances under which the crime could have been committed, i.e., in order for an undominated threshold rule that postulates extreme aversion to irrational convictions (false negatives) to exist, it must necessarily be the case that the juror always prefers to convict a guilty defendant (e.g., see Example 2 in the previous section).

Proposition 3. The upper standard of reasonable doubt \( p^u \) exists if and only if \( V_x(\omega) \geq 0 \) for all \( \omega \in G \).

Example 2 (continued). Recall the example from the previous section, now letting the underlying decision problem be \( \Gamma = \{0, 6\} \) instead (henceforth see Figure 3). Notice that there exists some \( \omega \in G \) such that \( V_6(\omega) < 0 \). Indeed, for severe sentences the juror prefers to acquit a guilty defendant if he committed the crime unintentionally, i.e., \( V_6(\omega_1) = -1 \). For starters, it is easy to see that \( p^l_6 = 0.15 \) is the lower standard of reasonable doubt. However, notice that there is no upper standard of reasonable doubt, i.e., there is no \( p \in [0, 1] \) such that \( C_6 \supseteq D_p \). Indeed, even if we set \( p = 1 \), the juror may still prefer to acquit the defendant, viz., \( \pi_0 \) attaches probability 1 to \( G \), and still \( \mathbb{E}_{\pi_0} V_6 < 0 \). This is because \( \pi_0 \) assigns sufficiently high probability to the defendant having committed the crime unintentionally, conditionally on him being guilty, in which case she prefers to acquit him. Finally, observe that – even though \( p^u_6 \) does not exist – every \( p^w_6 \in [p^l_6, 1] \) is a weak standard.

\[\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{Figure3}
\caption{Nonexistence of the upper standard of reasonable doubt.}
\end{figure}\]

5. Empirical implications

There is a large volume of applied research on the standard of reasonable doubt (e.g., see Simon and Mahan, 1971; Nagel, 1979; Dane, 1985; Connolly, 1987; Dhani, 2008, and references therein). One striking finding is that the estimates of the standard of reasonable doubt induced by the different elicitation approaches significantly differ from each other.\(^{21}\) In particular, direct questioning relies on explicitly asking subjects – e.g., students or mock jurors or actual judges – to state their probability threshold for convicting a defendant, and it typically yields an estimate of approximately 0.90 (e.g., see Simon and Mahan, 1971). On the other hand, the decision-theoretic approach relies on asking subjects

\(^{21}\)For an early overview of the different elicitation methods, see Dane (1985).
– from similar pools as above – to evaluate the four basic scenarios, i.e., “convict the guilty”, “convict the innocent”, “acquit the guilty” and “acquit the innocent”. Then, using these values as utilities, the respective probability threshold is computed, which typically falls in the range of 0.50 – 0.60 (e.g., see Nagel, 1979; Dane, 1985). Furthermore, while the decision-theoretic approach does better at predicting actual behavior than direct questioning, none of the two is a particularly good predictor (e.g., Dane, 1985; Connolly, 1987). There are different proposed explanations for this discrepancy, based on either operational or conceptual or psychological arguments, such as for instance the vagueness of the instructions or the framing of the questions (for an overview, see Connolly, 1987, pp. 110–111). Here, using our theory, we propose a new one.

For starters observe that the decision-theoretic approach essentially restricts the experimental subjects to reason only about the events in $G$. Thus, the standard of reasonable doubt $p_x$ always exists, which is probably what we (indirectly) elicit with this method.

On the other hand, direct questioning is not as explicit in specifying the subject’s frame $R$, which could very well be finer than $G$, especially when the question is based on a specific – actual or simulated – case, like for instance in Dane (1985). In this case, the standard of reasonable doubt does not exist, meaning that we probably elicit some weak standard of reasonable doubt $p_x^w \in [p^l_x, p^u_x]$. Of course in this case, the elicited threshold should be interpreted as the subject’s most preferred threshold among the irrational ones that are available, i.e., it would reflect not only the subject’s preference over outcomes, but also the subject’s preference for rationality from a normative point of view. This may also explain the low predictive capacity of direct questioning, in the sense that the subjects report an irrational threshold rule – either their most preferred one or perhaps the one that they think that society postulates – but in the end they choose rationally when the actual decision needs to be taken, and therefore their actual choice may easily contradict their stated threshold (e.g., see Dane, 1985; Connolly, 1987). In fact, since direct questioning yields consistently high estimates, we conjecture that the subjects report a threshold closer to the upper standard. This would suggest that subjects prefer irrational acquittals (false positives) over irrational convictions (false negative). The latter is also consistent with the conventional wisdom that the threshold postulated by the law is quite high.

In either case, our previous analysis is quite informal. While our theory provides a plausible explanation for the aforementioned discrepancy between the different elicitation methods, it still needs to be formally tested, both normatively (i.e., in terms of whether it captures the jurors’ interpretation of the law) and descriptively (i.e., in terms of whether it predicts well the jurors’ actual decision).

6. Discussion

6.1. Belief restrictions

Throughout the paper, we have followed a belief-free approach in the sense that a choice rule is deemed rational if it prescribes a rational choice for every belief in $\Delta(\Omega, R)$. However, it is often the case that during the trial some beliefs are ruled out based on the evidence presented to the court. For instance, if the defendant admits that he killed the victim, the juror rules out all the beliefs that put positive probability to the defendant not having done it. Thus we ask the following natural question: **how should we adapt our theory when such an exogenous restriction is imposed on the juror’s beliefs?**

Let us focus on cases where the belief restriction captures the idea that an event in $R$ has been proven in court, like for instance in the example above. In particular, let $K \in R$ be the event that the defendant killed the victim, thus implying that only beliefs in $\Delta(K, R) := \{\pi \in \Delta(\Omega, R) : \pi(K) = 1\}$ are considered by the juror. In this case, all our definitions should be adjusted conditional on the
beliefs in $\Delta(K, R)$. In particular, $C_x$ and $D_p$ are respectively replaced by

$$
C^K_x := \{ \pi \in \Delta(K, R) : E_x V_x \geq 0 \} \subseteq C_x, \\
D^K_p := \{ \pi \in \Delta(K, R) : \pi(G) \geq p \} \subseteq D_p,
$$

and all our results hold verbatim.

Indeed, the conditional standard of reasonable doubt exists if $C^K_x = D^K_p$ for some $p \in [0, 1]$. It is quite straightforward to verify that if the standard of reasonable doubt exists unconditionally, then it also exists conditional on any $K \in R$. The converse is not necessarily true, e.g., in Example 1 the standard of reasonable doubt exists conditionally on $E = \{ \omega_2, \omega_3 \}$, but not unconditionally. Likewise, the conditional upper standard of reasonable doubt exists if and only if $K \cap \{ \omega \in G : V_x(\omega) < 0 \} = \emptyset$, i.e., if and only if the prosecution has proven that states of the world where the juror prefer to acquit a guilty defendant have not occurred. In fact, if the juror actually uses the upper standard of reasonable doubt as her threshold for conviction (as suggested in Section 5), it is not even enough for the prosecutor to prove that $\pi(G) = 1$, unless he has also proven that $\pi(\{ \omega \in G : V_x(\omega) < 0 \}) = 0$. In this sense, the burden of proof may become quite heavy.

### 6.2. Multinomial choice

So far throughout the paper, our analysis has focused exclusively on binary decision problems $\Gamma = \{0, x\}$. Now suppose instead that the juror does not just choose whether to acquit or convict the defendant, but also picks the sentence to be issued in case she convicts him, i.e., formally, $|\Gamma| > 2$ (e.g., see Lundberg, 2016). This case also resembles the one in the safety-first principle from finance, where the investor can usually choose from a large set of assets (Telser, 1955-56). Then, the juror will prefer to convict the defendant if her beliefs belong to

$$
C_\Gamma := \{ \pi \in \Delta(\Omega, R) : \max_{x \in \Gamma} E_x V_x \geq 0 \}.
$$

Obviously this set of beliefs is not necessarily associated with a hyperplane, as illustrated in Figure 4, which depicts the adaptation of Example 2 with $\Gamma = \{0, 2, 4\}$, i.e., more specifically $C_\Gamma \neq D_p$ for all $p \in [0, 1]$. Therefore, a version of Theorem 1 will still hold. In fact, it will be even harder to find a rational threshold rule.

![Figure 4: The standards of reasonable doubt with $|\Gamma| > 2$.](image)

Now, let us switch attention to the weak standards of reasonable doubt, defining the upper and the lowered standard by $p^u_\Gamma := \min \{ p \in [0, 1] : C_\Gamma \supseteq D_p \}$ and $p^l_\Gamma := \max \{ p \in [0, 1] : C_\Gamma \subseteq D_p \}$ respectively.
The two extreme standards are graphically illustrated in Figure 4. Notice that whenever the two extreme standards exist, it is the case that $p_u^\Gamma \leq p_u^x$ and $p_l^\Gamma \leq p_l^x$ for all $x \in \Gamma$. This last conclusion is consistent with the one of Lundberg (2016), who showed that leaving the sentence at the juror’s discretion leads to a lower probability threshold.

Finally, note that $p_l^\Gamma$ always exists, as $p_l^\Gamma = \min_{x \in \Gamma} p_l^x$. On the other hand, $p_u^\Gamma$ does not always exist. Nevertheless, it is “easier” for an upper standard of reasonable doubt to exist when the decision problem contains more choices, e.g., if $p_u^x$ exists for at least one $x \in \Gamma$, so does $p_u^\Gamma$.

6.3. Subjective perception

Throughout the paper we have assumed that the juror shares the same interpretation of the events in $\mathcal{R}$ as the lawmaker. Now, consider the possibility that the juror perceives certain events differently, e.g., suppose that according to the law self-defense implies guilt, whereas according to the juror it implies innocence. Formally speaking – borrowing terminology from logic – the juror and the lawmaker would attach different semantics (viz., a different set of states) to the same syntactic proposition (viz., guilt in the previous example). This distinction is particularly relevant when the interpretation of certain relevant events is not transparent, e.g., in antitrust cases. In either case, all our results will still hold, by taking $\mathcal{R}$ to be the juror’s perceived frame, even if the latter differs from the lawmaker’s intended frame.

A. Proofs

Let us first introduce some additional machinery that we will need to prove some of our results. Let $\mathcal{P}$ be the (finite) partition of $\Omega$ that induces the frame $\mathcal{R}$, i.e., formally, $\mathcal{P}$ is a subcollection of nonempty events in $\mathcal{R}$ such that (i) $P_1 \cap P_2 = \emptyset$ for all $P_1, P_2 \in \mathcal{P}$, and (ii) for every $E \in \mathcal{R}$ there exist $P_1, \ldots, P_n \in \mathcal{P}$ such that $E = P_1 \cup \cdots \cup P_n$. It is well known that such a partition always exists. Since every event in $\mathcal{P}$ is $\mathcal{R}$-measurable, we define $V_x(P) := V_x(\omega)$ for an arbitrary $\omega \in P$. Then, notice that $\Delta(\Omega, \mathcal{R})$ is identified with the unit simplex over the finite set $\mathcal{P}$, which spans the hyperplane

$$H = \{ q \in \mathbb{R}^P : \sum_{P \in \mathcal{P}} q(P) = 1 \}$$

of the $|\mathcal{P}|$-dimensional euclidean space. Moreover, for any convex set $M \subseteq H$, let $\delta^* (\cdot | M) : H \to \mathbb{R}$ be the support function of $M$, defined by

$$\delta^*(q|M) := \sup \left\{ \sum_{P \in \mathcal{P}} \pi(P)q(P) \bigg| \pi \in M \right\}$$

for each $q \in H \subseteq \mathbb{R}^P$. Then, it follows from Rockafellar (1970, p.112) that, for every $q \in H$,

$$M \subseteq \left\{ \pi \in H : \sum_{P \in \mathcal{P}} \pi(P)q(P) \leq p \right\} \iff p \geq \delta^*(q|M). \tag{A.1}$$

Proof of Theorem 1. Necessity ($\Rightarrow$). Let $p_x$ be the standard of reasonable doubt. Consider the following

22Observe that this not an if-and-only-if statement, i.e., there may exist some $\Gamma = \{0, x, y\}$ with $p_u^x$ existing, even though neither $p_u^x$ nor $p_u^y$ exists.
hyperplanes of $H$,

\[
H_e := \{ q \in H : \sum_{P \in \mathcal{P}} q(P) V_x(P) = 0 \}
\]

\[
H_d := \{ q \in H : \sum_{P \in \mathcal{P}} q(P) W_x(P) = p_x \}
\]

\[
= \{ q \in H : \sum_{P \in \mathcal{P}} q(P) W_x(P) = \sum_{P \in \mathcal{P}} q(P) p_x \}
\]

\[
= \{ q \in H : \sum_{P \in \mathcal{P}} q(P) (W_x(P) - p_x) = 0 \},
\]

where $W_x(P) := 1$ if $P \subseteq G$ and $W_x(P) := 0$ otherwise. Notice that $H_e$ and $H_d$ are the hyperplanes associated with the half-spaces $C_x$ and $D_p$, respectively. Hence, by the fact that $C_x = D_p$, it follows that $H_e = H_d$. Thus, there exists some $\lambda \in \mathbb{R}$ such that $V_x(P) = \lambda(W_x(P) - p_x)$ for all $P \in \mathcal{P}$. That is, $V_x(P) = \lambda(1 - p_x)$ for each $P \subseteq G$ and $V_x(P) = -\lambda p_x$ for each $P \subseteq I$. Therefore $V_x$ is $\mathcal{G}$-measurable.

**Sufficiency ($\Leftarrow$).** Let $V_x$ be $\mathcal{G}$-measurable. Define $V_x(G) := V_x(P)$ for $P \subseteq G$ and $V_x(I) := V_x(P)$ for $P \subseteq I$. Notice that by $(A_0)$, it is the case that $V_x(I) < 0$. Moreover, since $x$ is nontrivial, we also obtain $V_x(G) > 0$. Thus,

\[
C_x = \{ \pi \in \Delta(\Omega, \mathcal{R}) : \sum_{P \in \mathcal{P}} \pi(P) V_x(P) \geq 0 \}
\]

\[
= \{ \pi \in \Delta(\Omega, \mathcal{R}) : \pi(G) V_x(G) + (1 - \pi(G)) V_x(I) \geq 0 \}
\]

\[
= \{ \pi \in \Delta(\Omega, \mathcal{R}) : \pi(G) (V_x(G) - V_x(I)) \geq -V_x(I) \}
\]

\[
= \{ \pi \in \Delta(\Omega, \mathcal{R}) : \pi(G) \geq V_x(I)/(V_x(I) - V_x(G)) \}
\]

\[
= D_{p_x},
\]

where $p_x := V_x(I)/(V_x(I) - V_x(G))$. And obviously $p_x \in (0, 1)$, since $V_x(I) < 0$ and $V_x(G) > 0$.

**Proof of Theorem 2. Necessity ($\Rightarrow$).** Suppose that $p_x^u$ is a weak standard of reasonable doubt, i.e., let $p_x^w \geq p_x^w \geq p_x^l$. Now, unless $p_x^w = 1$, consider an arbitrary $p > p_x^u \geq p_x^l$, thus obtaining $D_p \subseteq D_{p_x}$ and $C_x \not\subseteq D_p$. Hence, it will necessarily be the case that $P_x^u \subseteq P_x^l$, and therefore $R_x^{p_x^u} \not\subseteq R_x^{p_x^l}$, i.e., $\sigma_{p_x^u}$ is not dominated by $\sigma_{p_x^l}$. Likewise, unless $p_x^w = 0$, consider an arbitrary $p < p_x^w \leq p_x^l$, in which case we get $D_p \supseteq D_{p_x}$ and $C_x \not\subseteq D_p$. Then, similarly to the previous case, it will necessarily be the case that $N_x^{p_x^w} \subseteq N_x^{p_x^l}$, and therefore $R_x^{p_x^w} \not\subseteq R_x^{p_x^l}$, i.e., $\sigma_{p_x^w}$ is again not dominated by $\sigma_{p_x^l}$, which completes the proof.

**Sufficiency ($\Leftarrow$).** Suppose that $p$ is not a weak standard of reasonable doubt, i.e., let either $p > p_x^u$ or $p < p_x^l$. Of course, $p > p_x^u$ (resp., $p < p_x^l$) can hold only if the upper standard (viz., the lower standard) exists and is strictly less than 1 (resp., strictly greater than 0). So let us start by taking $p > p_x^u$ where $p_x^u < 1$. Then, by (9), we obtain $C_x \supseteq D_{p_x^l} \supseteq D_p$. Hence, it is the case that $N_x^{p_x^u} = N_x^{p_x^l} = \emptyset$ and $P_x^{p_x^u} \not\subseteq P_x^{p_x^l}$, thus implying $R_x^{p_x^u} \not\subseteq R_x^{p_x^l}$, i.e., $\sigma_p$ is dominated by $\sigma_{p_x^u}$. Likewise, if we take $p < p_x^l$ with $p_x^l > 0$, then by (10), we obtain $C_x \subseteq D_p \not\subseteq D_{p_x^l}$. Therefore, $N_x^{p_x^u} \subseteq N_x^{p_x^l}$ and $P_x^{p_x^u} = P_x^{p_x^l} = \emptyset$, thus implying $R_x^{p_x^u} \supseteq R_x^{p_x^l}$, i.e., $\sigma_p$ is dominated by $\sigma_{p_x^l}$, which completes the proof.

**Proof of Proposition 1. Necessity ($\Rightarrow$).** Let $p_x$ be the standard of reasonable doubt, i.e., it is the case that $C_x = D_{p_x}$. Then, it will necessarily be the case that $R_x^p = \Delta(\Omega, \mathcal{R})$, i.e., the choice prescribed by $\sigma_{p_x}(\pi)$ is rational for every belief $\pi \in \Delta(\Omega, \mathcal{R})$. Hence, for every $p \in [0, 1]$ it is the case that $R_x^p \supseteq R_x^p$. Indeed, pick an arbitrary $p < p_x$, which always exists (see Footnote 15). In this case, $D_p \supseteq D_{p_x}$, and therefore $D_p \supseteq C_x$. Hence, $R_x^p \not\subseteq \emptyset$, and therefore $R_x^p \supseteq R_x^p$, implying that $p_x$ dominates $p$. Hence, by Theorem 2, we obtain $p \notin [p_x^u, p_x^l]$. Likewise, pick an arbitrary $p > p_x$, which again always exists (see Footnote 15). In this case, $D_p \not\subseteq D_{p_x}$, and therefore $D_p \subseteq C_x$. Hence, $R_x^{p_x^u} \supseteq R_x^p$, again implying, by Theorem 2, that $p \notin [p_x^u, p_x^l]$. Therefore, the only threshold in $[p_x^u, p_x^l]$ is $p_x$ itself, thus completing this part of the proof.

**Sufficiency ($\Leftarrow$).** Let $p_x^u = p_x^l$, implying by definition that $\{ p \in [0, 1] : C_x \supseteq D_p \} \cap \{ p \in [0, 1] : C_x \subseteq D_p \} \neq \emptyset$. Hence, there exists some $p_x \in [0, 1]$ such that $C_x = D_{p_x}$. \qed
Proof of Proposition 2. Now take $M := C_x$, which is obviously convex, as it is the intersection of a simplex and a half-space. Moreover, set $q := -W_x$, where similarly to the previous proof, $W_x(P) := 1$ if $P \subseteq G$ and $W_x(P) := 0$ otherwise. Furthermore, let $p^\ell_x := -\delta^*(W_x|A_x)$. Hence,

$$C_x \subseteq D_p \iff M \subseteq \{ \pi \in \Delta(\Omega, \mathcal{R}) : -\pi(G) \leq -p \}$$

$$\iff -p \geq -p^\ell_x$$

$$\iff p \leq p^\ell_x,$$

(A.2)

with the first and the third equivalence being obvious, and the second equivalence following from (A.1). Hence, it suffices to prove that $p^\ell_x \in [0,1]$. We proceed by contradiction, first supposing that $p^\ell_x < 0$. Then, by (A.2), it follows that there is no $p \in [0,1]$ such that $C_x \subseteq D_p$, which obviously contradicts the fact that $D_0 = \Delta(\Omega, \mathcal{R})$. Finally, suppose that $p^\ell_x > 1$. Then, again by (A.2), it follows that $C_x \subseteq D_p$ for all $p \in [0,1]$, and in particular it follows that $C_x \subseteq D_1$. Now recall that $x$ is nontrivial, and therefore there exists some $P \in \mathcal{P}$ with $V_x(P) > 0$. Moreover, take an arbitrary $P' \in \mathcal{P}$ with $P' \subseteq I$, and recall that $V_x(P') < 0$, by (A0). Then, define $\pi_x \in \Delta(\Omega, \mathcal{R})$ so that $\pi(P) = \varepsilon$ and $\pi(P') = 1 - \varepsilon$, and notice that $\mathbb{E}_{\pi_x}V_x = \varepsilon V_x(P) + (1 - \varepsilon)V_x(P')$ is continuous in $\varepsilon$. Hence, there exists some $\varepsilon \in (0,1)$ such that $\mathbb{E}_{\pi_x}V_x > 0$, thus implying that $\pi_x \in C_x \setminus D_1$, which completes the proof.

Proof of Proposition 3. For the most part, we follow similar steps as in the proof of Proposition 2. Take $M := -C_x$, and set $q := W_x$ and $p^u_x := \delta^*(W_x|A_x)$. Hence,

$$C_x \supseteq D_p \iff C_x \supseteq \{ \pi \in \Delta(\Omega, \mathcal{R}) : \pi(G) > p \}$$

$$\iff M \subseteq \{ \pi \in \Delta(\Omega, \mathcal{R}) : \pi(G) \leq p \}$$

$$\iff p \geq p^u_x,$$

(A.3)

with the first equivalence following from the fact that $C_x$ is closed and $D_p = \text{clos}(\{ \pi \in \Delta(\Omega, \mathcal{R}) : \pi(G) > p \})$, the second equivalence following from $M = -C_x$, and the third equivalence following from (A.1). Hence, it suffices to prove that $p^u_x \in [0,1]$ if and only if $V_x(P) \geq 0$ for all $P \subseteq \mathcal{P}$ with $P \subseteq G$.

Step 1. First we prove that $p^u_x \geq 0$ irrespective of whether $V_x(P) \geq 0$ for all $P \subseteq G$ or not. Suppose instead that $p^u_x < 0$. Then, by (A.3), we obtain $C_x \supseteq D_0 = \Delta(\Omega, \mathcal{R})$, which contradicts (A0).

Step 2. Second, we prove that $p^u_x > 1$ if and only if $V_x(P) < 0$ for some $P \subseteq G$, which is obviously equivalent to what we want to show. Indeed, let $p^u_x > 1$, which (by (A.3)) is equivalent to $C_x \not\supseteq D_1$. The latter is true if and only if there exists some $\pi \in \Delta(\Omega, \mathcal{R})$ simultaneously satisfying $\pi(G) = 1$ and $\mathbb{E}_{\pi}V_x < 0$. Finally notice that this is the case if and only if $V_x(P) < 0$ for some $P \subseteq G$.

References


