

Persuasion as a Contest

By

Stergios Skaperdas* and Samarth Vaidya**

February 23, 2009

ABSTRACT: We examine how the probability of persuading an audience depends on resources expended by contending parties as well as on other factors. We use a Bayesian approach whereby the audience makes inferences solely based on the evidence produced by the contestants. We find conditions that yield the well-known additive contest success function, including the logit function. We also find conditions that produce a generalized "difference" functional form. In all cases, there are three main determinants of audience choice: (i) the truth and other objective parameters of the environment; (ii) the biases of the audience; and (iii) the resources expended by the interested parties.

JEL: C70, D20, D70

Keywords: rent-seeking, advertising, litigation, political campaigning, property rights.

*Department of Economics, University of California, Irvine, CA 92697, USA,
email: sskaperd@uci.edu

**School of Accounting, Economics, and Finance, Faculty of Business and Law,
Deakin University, 221 Burwood Hwy, Burwood, VIC 3125, Australia
email: svaidya@deakin.edu.au

We would like to thank Hao Jia, R. Duncan Luce, Kalle Moene, Karl Warneryd, and two referees for comments or discussions, as well participants in the conference "Advances in the theory of contests and tournaments" at the Social Science Research Council-Berlin (WZB), the 2007 AEA meeting session on "transparency", and seminar participants at Deakin University, Drexel University, Monash University, UC Irvine, and USC.

1 Introduction

Much costly economic activity – be it litigation, advertising, lobbying, electoral campaigning or argumentation in policy debates or within organizations – could be thought of as being about persuasion. All such activities involve interested parties devoting resources to influence the opinion of an audience – a court, voters, consumers, fellow employees, the public at large – in their favor. Unlike the case of ordinary economic production, in which inputs are combined cooperatively, the inputs of persuasion are contributed by interested parties usually in an *adversarial* fashion. Ordinary production functions could hardly, then, describe the process by which the inputs of persuasion are translated into outputs. We argue that, instead, persuasion could better be thought of as a contest, a game in which players invest in costly effort to win a prize.

Indeed there are sizable literatures in different areas, in certain cases isolated from one another, that treat persuasion as a contest. Perhaps the most well-known area is that of rent-seeking and lobbying that started with Tullock (1980) (see Nitzan, 1994, for a review). Earlier, in the 1960s and 1970s, there was an extensive research effort on advertising (see, e.g., Schmalensee, 1972) and for decades there has been extensive related research on marketing (Bell et. al., 1975). More recently there has been research using contests in the study of litigation (Farmer and Pecorino 1999, Bernardo et. al., 2000, Hirshleifer and Osborne, 2001), political campaigns (Snyder, 1989, Baron, 1994, Skaperdas and Grofman, 1995), bureaucratic organization and corporate governance (Warneryd, 1998, Mueller and Warneryd, 2001, Konrad, 2004), or media politics (Vaidya, 2005, 2006).

A common ingredient of research in contests is the contest success function, the function that translates efforts into probabilities of winning and losing for the participating contestants. Such probabilistic choice functions have been explored since Luce (1959). In the case of contest success functions, there are three types of derivations: axiomatic (Bell et.al, 1975, Skaperdas, 1996, Clark and Riis, 1998), stochastic (Hirshleifer and Riley, 1992, p.380 and Jia, 2008), and those with microfoundations on innovation-type of races (Fullerton and McAfee, 1999, and Baye and Hoppe, 2003).

(For more detailed comparisons, see the illuminating survey of Konrad, 2007). None of these derivations, however, considers the role of the audience as an active participant in the persuasion process.

Our primary aim in this paper is to derive contest success functions used in the literature from the inferential process of an audience that observes evidence produced by the contestants who seek to persuade the audience of the correctness of their respective views. The main class of functional forms that have been used is the following additive form (in its symmetric version):

$$P_i(R_i, R_j) = \frac{f(R_i)}{f(R_i) + f(R_j)} \quad (1)$$

where R_i and R_j are resources expended by contestants i and j , $f(\cdot)$ is a positive, increasing function, and P_i represents contestant i 's probability of winning. We derive the functional form in (1) and its variations as an outcome of an audience's inference process when evidence production is deterministic and the contestants have a continuous probability distribution over the audience's priors. When the contestants know the audience's priors with certainty, we obtain the limiting case of the perfectly discriminating contest or the all-pay auction, whereby the contestant who puts more resources into the contest wins with certainty (see Hillman and Riley, 1989, Baye et al., 1996, Che and Gale, 1998). Thus, we find a clean set of circumstances to which the imperfectly and perfectly discriminating contests might apply.

Furthermore, we examine a simple stochastic evidence production process which, in its most simplified and symmetric form, yields the following functional form:

$$P_i(R_i, R_j) = \frac{1}{2} + \frac{\alpha}{2}[h(R_i) - h(R_j)] \quad (2)$$

where $h(\cdot)$ is a positive, increasing function (with appropriately defined bounds), $\alpha > 0$, and the other variables and functions are similarly defined to those in (1). While this functional form bears resemblance to the one analyzed by Che and Gale (2000), we are the first to provide an explicit inferential foundation for such a contest function. Also, unlike Che and Gale (2000), in general, (2) allows for a non-linear

relationship between probability of winning and resources devoted by rival parties.

Also, for both classes of functional forms that we derive, we find intuitively appealing ways in which the *truth*, the *biases* and *preconceptions* of the audience, as well as the *resources* expended by the contestants affects the probabilities that the audience will choose one of the alternatives.

While competition among parties to influence an audience bears a resemblance to the literature on special interest politics, in our context these efforts are not simply money exchanged or bribes (as in Grossman and Helpman, 1994) but rather inputs into the production of arguments and evidence. The literature on rent-seeking contests also tends to interpret the efforts as bribes, or at least it is agnostic about what the efforts are. Our derivation of contest functions as “persuasion” functions at least makes us think seriously about distinguishing between outright bribing and information provision and that these two types of influence might have distinct effects.¹

Throughout the analysis, we suppose that the audience makes inferences solely based on the evidence that it sees, and does not even attempt to infer who or how the evidence might have been produced.² Juries and judges are formally closest to

¹Campaign contributions are different from the salaries of lobbyists and the costs of maintaining offices. Campaign contributions can be thought of as a combination of bribes and the buying of access, the latter allowing the lobbyists to make their arguments, whereas the other costs of lobbying could be thought as being costs of persuasion. It would be reasonable then to conjecture that the cost of bribing is only a small fraction of the costs of persuasion that have to be undertaken.

For a lobbying-as-bribe interpretation of contests and a derivation of contest success functions from the objectives of politician, see Epstein and Nitzan (2006).

Apart from our paper, there exists an alternative literature (in particular see Potters and Van Winden, 1992, Lagerlof, 1997, 2007) which also emphasizes resources invested by parties in rent-seeking and lobbying games as going towards costly information provision and then re-examines the welfare implications of such expenditures. However, despite sharing some commonalities with this approach (especially, Lagerlof, 2007), our line of inquiry is different as we are interested in identifying sufficient conditions involving plausible evidence production environments which can support some well known contest success functions - something that this strand of literature does not attempt to do. At a later stage in the paper we spell out more clearly what the similarities and differences are between our analysis and this approach.

²In this respect, our approach is similar to Mullainathan Schwartzstien and Shleifer(2006) who also examine a case where the target of persuasion takes the messages at their face value without concerns of credibility even though the objective of the party sending messages need not be aligned with his own. Froeb and Kobayashi (1996) also assume their decision maker to be naive who takes evidence at face value ignoring strategic suppression of unfavorable evidence although their line of

such an assumption, but we think in other instances of persuasion it is a reasonable approximation, especially when compared to another extreme in which the audience would have full knowledge of (or, a probability distribution over) the objectives and strategies of probable or possible contestants far removed from the audience itself.³ There is a considerable amount of research in psychology and related areas indicating that audiences, in making inferences from messages they receive, do focus on less than the full universe of those who might produce the messages and their motivations (for an overview of research in this area, see Cialdini, 2001). People are affected by what they see – the product or candidate commercial, the trial evidence, the arguments made by their colleagues in their organization – without considering all the imaginable, let alone all the unimaginable but real, possibilities and contingencies. In other words, they can be considered to inhabit a “limited world” where they might not question all the possible strategic reasons that some of the messages that they observe might be generated. There is also considerable practice in the area of public relations, advertising, and the formation of public opinion, at least since the 1920s (see Lippmann, 1997, for one pioneer in the area), with insights similar to those of psychology and related areas.

Our focus is rather different from other work in economics that considers persuasion as transmission of private information. This literature examines persuasion as communication of privately observed information by one or more agents to a decision maker so as to influence his or her judgement on a matter at hand in their favour. Since the objective of the decision maker is not necessarily aligned with those of the agents, the latter may have an incentive to transmit only information that might tilt the decision in their favor. One branch of this literature examines the case where acquiring private information is costless for agents and uncovers conditions or mechanisms that might allow the decision maker to best extract the relevant information from the agents so as to minimize or eliminate the prospect of a wrong decision (see Milgrom and Roberts, 1986). Froeb and Kobayashi (1996) also pursue this line of

inquiry is quite different from ours.

³Recent research in game theory (e.g., Filiz, 2006, Ozbay, 2006) explicitly allows for a subset of the players to be “unaware” of the full strategy space as perceived by some other players.

investigation but by allowing for information acquisition to be costly for the participants. By incorporating costly collection of information by agents, yet another branch of the literature examines properties of optimal contracts or institutions that the decision maker could adopt to cope with the dual issues of providing incentives for information acquisition as well as discouraging the agents to manipulate the communicated information in their favour when their payoffs depend on the decision made (see Dewatripoint and Tirole, 1999 and Shin, 1998). The approach taken in this paper is fundamentally different from all of these approaches as in our model the incentives of the parties attempting to influence the decision maker are outside his or her control. Hence our decision maker could be a judge who must decide on a case by examining presented evidence without having much control over stakes or the resources put in by the parties supplying the same or it could be a voter deciding on the basis of information supplied by competing candidates or a consumer choosing among alternative brands, influenced by advertising expenses of rival companies. In all these contexts the stakes of competing parties are typically high and external to the decision makers. Hence our focus is mainly positive in nature as we just seek to find possible inferential justification of functions that capture the impact of such persuasion better than production functions do.

In this sense, our paper has more in common with the newer psychological approach to modelling persuasion such as Aragonés, Gilboa, Postlewaite and Schmeidler (2005) and Mullainathan, Schwartzstein and Shleifer (2006). While Aragonés, Gilboa, Postlewaite and Schmeidler (2005) formally model inference via analogies using a non-Bayesian learning model, Mullainathan, Schwartzstein and Shleifer (2006) present a Bayesian model of inference which incorporates associative thinking identified in the psychology literature. Our approach is closer in spirit to Mullainathan, Schwartzstein and Shleifer (2006) as they also consider a “limited world” Bayesian audience as a sub case. However, while their focus is to formalize a specific psychological pattern of thinking, our approach is less specific. We simply assume that the competing parties that attempt to influence the audience are aware of what information the audience might find potentially convincing.

Before proceeding with the modeling, we briefly discuss the importance of per-

sualion as an economic activity in modern economies. In modern democracies, the importance of mobilizing voters through persuasive campaigns and lobbying politicians for favors can hardly be over-emphasized.⁴ Besides political influence, persuasion has also gained significance in purely economic activities. The suppliers of persuasion be it consulting agencies, law firms or advertising conglomerates have become indispensable to the working of corporations. The size and revenue earnings of law firms worldwide is testimony to their importance.⁵ What's more, these law firms also have access to litigation consultants who specialize in refining trial strategies to maximize the persuasive impact on juries. Today's advertising agencies are also multi-billion dollar industries providing an array of specialized services such as: creating and maintaining brand loyalties, crisis management (helping companies protect their image when under critical scrutiny), public relation strategies to moderate communication within organizations and public communication strategies to influence legislative actions, among a host of other things. Social science, even the physical sciences, can be considered to involve costly persuasion. Otherwise, all scientifically correct hypotheses would become accepted immediately at the instant they become formulated and proposed. Instead, what we have is a long process of costly thought, writing, reformulation, testing and retesting, argumentation, and so on. In such a process, the "truth" and presented evidence are not the sole determinants of what is accepted as true. Biases and prejudices are relevant as well.⁶ Overall, many parts of the service economy and their employees can be reasonably be

⁴As estimated by Laband and Sophocleus (1992) (see page 966), in 1985, private parties spent approximately \$4.6 billion on lobbyists, \$1 billion on political action committees and \$1.7 billion on individual contributions to political candidates.

⁵In 2002, each of the world's top three law firms earned more than a billion US dollars in gross revenues. (See 2002 Global 100 ranking of law firms based on gross revenues in *The American Lawyer*, November, 2002). Further, major corporations not only use outside law firms but maintain an arsenal of lawyers internally to protect their business practices in the face of litigation (See "Who Defends Corporate America?" in *The National Law Journal*, October 15, 2001. As per the survey, companies like Exxon Mobil Corp. and General Electric maintain over 600 in-house lawyers). In 1985, the total estimated expenses by both the defendants and the plaintiffs in Tort litigation in the U.S. was approximately \$17.3 billion (See Laband and Sophocleus (1992), page 964).

⁶Precisely such a process of scientific discovery – whereby typically old hypotheses take much effort to be replaced by new ones – has been elaborated in the seminal work by Kuhn (1996).

thought of as engaging in persuasion. Ordinary production functions cannot capture such activities as they are fundamentally adversarial.

We motivate our models by considering a trial setting involving a court, a Plaintiff and a Defendant. A trial setting is helpful because it is focused on establishing the truth for a particularly narrow subject and has relatively well-defined rules for evidence production and a structured evaluation process. However our models are not specific to court settings and have applicability in the other contexts we have mentioned.⁷ We follow a Bayesian approach not because we think courts and individuals necessarily closely follow such an inference process, but because the approach provides a consistent structure to the inference process in a simple way.⁸

2 The Basic Setting

For concreteness we use a court setting to lay out the building blocks of persuasion. The prosecution or Plaintiff (denoted by P) and the defense or the Defendant (denoted by D) compete to gather and present evidence so as to influence the verdict of the court in their favor. There are two competing hypotheses: either the Defendant is guilty (G) or he is innocent (I). We consider three stages in the process of evidence production and its assessment by the court:

1. Firstly, we posit that collection of evidence is costly; the Plaintiff and the Defendant expend resources R_p and R_d respectively to gather evidence favorable to their cause. Hence, even the side arguing for the truth, must put in effort to find any relevant evidence. Upon investing resources, each side may obtain a piece of evidence e_i in its favour where $i = p, d$ which it then presents to the court. We postulate evidence production functions that specify the link between resources spent and

⁷Furthermore, our approach is not meant explicitly to model the different ways in which evidence is produced, admitted or treated in different legal systems. For a more detailed axiomatic treatment of such issues, see Daughety and Reinganum (2000).

⁸See Kadane and Schum (1996) for an excellent discussion on Bayesian inference from the evidence presented in the Sacco and Vanzetti trial.

evidence produced and they may be either deterministic or stochastic as will become clear in the following sections.

2. Based on the evidence presented, the court updates its prior belief about guilt or innocence to determine its posterior beliefs. We assume that the court uses Bayes' rule to do so. Throughout the analysis, we assume that the court neither observes the level of resources chosen by either parties nor the underlying evidence production functions. Hence the only additional information it receives over and above its priors is the evidence pair presented at the trial.

3. Given its posterior belief, the court makes a decision on guilt or innocence. The nature of court's decision in turn helps determine the win probability for the Plaintiff and the Defendant.

Throughout the analysis, we take the resources expended in stage 1, R_p and R_d , as given. We are mainly concerned with the construction of posterior beliefs (as functions of R_p and R_d) in stage 2 and in the court's decision based on its posterior belief in stage 3. The key ingredient towards determining the court's posterior probability of guilt is its inferred likelihood ratio of guilt L^g which represents the force of the presented evidence in establishing guilt. By definition, L^g is given by:⁹

$$L^g = \frac{\text{prob}(e_p, e_d | G)}{\text{prob}(e_p, e_d | I)} \quad (3)$$

Kadane and Schum (1996), in the pioneering analysis of the Sacco and Vanzetti trial evidence, use such likelihood ratios to estimate the probative force of the evidence in establishing guilt. These ratios can be subjectively constructed directly without having specific values for the two conditional probabilities in (3).¹⁰

⁹Analogously, by definition, the likelihood ratio of innocence, L^i , is just the inverse of L^g :

$$L^i = \frac{1}{L^g} = \frac{\text{prob}(e_p, e_d | I)}{\text{prob}(e_p, e_d | G)}.$$

¹⁰As stated by Kadane and Schum (1996, p.127) in making a holistic assessment of the probative force of evidence: "All that is required is a person's judgement of [the likelihood ratio]; there is no need for separate or individual values of these two conditional probabilities."

Let π represent the prior probability of guilt and π^* the posterior probability of guilt based on the evidence presented by the two sides (i.e. $\pi^* = \text{prob}(G \mid e_p, e_d)$). Then, using Bayes' rule, we have:

$$\pi^* = \frac{\pi L^g}{(1 - \pi) + \pi L^g} \quad (4)$$

In what follows, we explore two different evidence production processes, postulate their impact on L^g , and accordingly, examine their effect on the posterior probability of guilt π^* and on the final decision by the target of persuasion in stage 3.

3 Deterministic Evidence Production and the Additive Contest Success Function

As already mentioned, different versions of the additive contest success function have recently been used in applications to law and economics, marketing, advertising, and political lobbying to model the win-probabilities of parties devoting costly resources to influence the judgement of the relevant audience. While such functions have been justified axiomatically or in terms of probabilistic choice, a justification for their usage in such inferential settings is still lacking. In this section, we link evidence production and Bayesian inference to identify sufficient conditions for the usage of the additive function.

3.1 Evidence and the determination of posteriors

Let each potential piece of evidence favouring guilt or innocence (that is, e_p and e_d) be on a $(0, \infty]$ scale. Suppose that, given the evidence pair (e_p, e_d) , the court constructs a likelihood ratio of guilt (see (3)) which takes the following *power-law form*:

$$L^g(e_p, e_d) = \lambda \left(\frac{e_p}{e_d} \right)^\mu \quad \text{where } \lambda, \mu > 0 \quad (5)$$

Using the link between probabilistic reasoning and human judgements in general, as discussed in Kadane and Schum (1996), $L^g(e_p, e_d)$ in (5) can be interpreted as the court's perception of the intensity of evidence (as measured by $\frac{e_p}{e_d}$) pointing towards guilt.¹¹ In particular, the power-law form in (5) has been identified in psychophysical experiments, where it is well established that quantitative human perception (such as sensation of relative brightness of light or loudness of sound, as well as judgements concerning intensity of attitudes and opinions) of stimuli follows a power law.¹² Moreover, (5) can be derived from certain probability distributions that the court might have about the generation of the evidence.¹³ The parameter

¹¹On page 163, Kadane and Schum (1996) quote David Hume from his *Treatise of Human Nature* "Thus all probabilistic reasoning is a species of sensation.... When I am convinc'd of any principle, 'tis only an idea, which strikes more strongly upon me. When I give preference to one set of arguments above another, I do nothing but decide from my feeling concerning the superiority of their influence." Referring to Steven's work in psychophysics, they interpret Hume's quote as follows: "in Steven's terms, Hume said that all probabilistic reasoning involves judgments that are prothetic or intensive in nature." They also note on page 165, that "Stevens has gone to considerable lengths to show how methods and issues of interest in sensory psychophysics are applicable to human judgments concerning the intensity of attitudes and opinions about value related and other matters (1966, 1972)." Similarly, our choice of $\frac{e_p}{e_d}$ as a measure of evidence intensity also relies on the observation that "as Stevens noted (1975, 17), a wide variety of our [human] perceptual judgemental capabilities rest on preservation of information about ratios rather than about differences." See Kadane and Schum (1996), page 165.

¹²See Stevens (1966, 1975, p13) and Sinn (1983) pages 127-133 for a survey of these findings.

¹³For example, suppose that the court believes that $prob(e_p, e_d | G)$ and $prob(e_p, e_d | I)$ are described by the following gamma distributions:

$$prob(e_p, e_d | G) = \frac{\left(\frac{e_p}{e_d}\right)^{a_g-1} \exp\left(-\frac{e_p}{be_d}\right)}{b^{a_g} \Gamma(a_g)}$$

where a_g and b are positive parameters and $\Gamma(\cdot)$ is the gamma function; and

$$prob(e_p, e_d | I) = \frac{\left(\frac{e_p}{e_d}\right)^{a_i-1} \exp\left(-\frac{e_p}{be_d}\right)}{b^{a_i} \Gamma(a_i)}$$

where $a_i > 0$.

Then, from (3), we have

$$L^g(e_p, e_d) = \frac{prob(e_p, e_d | G)}{prob(e_p, e_d | I)} = \frac{b^{a_i-a_g} \Gamma(a_i)}{\Gamma(a_g)} \left(\frac{e_p}{e_d}\right)^{a_g-a_i}$$

Note that this takes the same form as (5) for $\lambda = \frac{b^{a_i-a_g} \Gamma(a_i)}{\Gamma(a_g)} > 0$ and $\mu = a_g - a_i > 0$.

λ is an indicator of the bias the court has in weighing (or, taking account of) the evidence of the two sides. Obviously, $\lambda > 1$ would indicate a bias in favor of the Plaintiff's side and $\lambda < 1$ would indicate bias in favor of the Defendant.

Next, we assume that both the Plaintiff and the Defendant are aware as to how the court infers from evidence presented as per (5). Producing evidence is costly as it depends on the amount of resources devoted by the Plaintiff and the Defendant. In particular we consider the following evidence production functions:

$$\begin{aligned} e_p &= F_p(R_p) \\ e_d &= F_d(R_d) \end{aligned}$$

As per the above functions, the values of evidence obtained (by Plaintiff and Defendant) $F_p(\cdot)$ and $F_d(\cdot)$ are deterministically and monotonically increasing in resources. Via these evidence production functions, (5) allows us to establish a direct link between resources and the court's likelihood ratio of guilt. By substituting for $L^g(e_p, e_d)$ into (4), we can express the court's posterior probability of guilt as a function of resources as given by $\pi^*(R_p, R_d) = \frac{\lambda\pi(F_p(R_p))^\mu}{(1-\pi)(F_d(R_d))^\mu + \lambda\pi(F_p(R_p))^\mu}$.

With this assumption, as the above expression reveals, the probability of guilt as a function of resources turns out to be a generalized additive function, a result summarized below.

Proposition 1: *Assume the likelihood ratio is described by (5), where $F_p(\cdot)$ and $F_d(\cdot)$ are deterministic evidence production functions for the Defendant and the Plaintiff. Then, the court's posterior probability of guilt is described by the following function:*

$$\pi^*(R_p, R_d) = \frac{\lambda\pi(F_p(R_p))^\mu}{(1-\pi)(F_d(R_d))^\mu + \lambda\pi(F_p(R_p))^\mu} \quad (6)$$

The prior π and the multiplicative parameter λ in the likelihood function (5) play a similar role, in the sense that having a higher prior of guilt or weighing evidence in favor of the Plaintiff (and against the Defendant) more heavily increases the court's posterior of guilt. The side that puts relatively more resources tilts the verdict in

its favor by obtaining a relatively higher valued evidence.¹⁴ The parameter μ is a measure of the sensitivity of the court to the evidence and, through that, to the resources expended (see Hirshleifer, 1989, and Jia, 2008, for interpretations of this parameter). For brevity, from now on we set both parameters λ and μ equal to 1, and thus we will be considering the following version of (6):

$$\pi^*(R_p, R_d) = \frac{\pi F_p(R_p)}{(1 - \pi)F_d(R_d) + \pi F_p(R_p)} \quad (7)$$

From Proposition 1 it is clear that the specific form of $\pi^*(R_p, R_d)$ depends on the evidence production functions $F_p(R_p)$ and $F_d(R_d)$. In their current form, the evidence production functions explicitly capture only the importance of resources on evidence produced. However, apart from resources, it is natural to think that the truth itself would influence evidence production too. While there is no unique way of characterizing this, we consider the following restrictions intuitive:

When truth is on the side of the Defendant, we expect $F_d(R_d) > F_p(R_p)$ when $R_d = R_p > 0$. Similarly, when the truth is with the Plaintiff, we expect that $F_p(R_p) > F_d(R_d)$ when $R_d = R_p > 0$.¹⁵ In particular, consider the following evidence production functions:

$$\begin{aligned} e_p &= F_p(R_p) = \varphi f(R_p) \\ e_d &= F_d(R_d) = (1 - \varphi)f(R_d) \end{aligned}$$

¹⁴Indeed, given such a production process, the court's inference on the basis of face value of evidence allows it to be misled especially since we don't necessarily expect the party that spends more to be on the side of truth. Legros and Newman (2002) also allow for the possibility that by devoting greater resources, a party can mislead the judge. In their model the party which invests more resources into trial hires a better lawyer who can then produce an "interference" or jam the argument of the opponent's lawyer from reaching the judge.

¹⁵While this is clearly desirable, it might not always be true. Generally, the evidence production process would also be sensitive to the nature of dispute at hand. At times, it might be much harder for the plaintiff to provide evidence in favor of its cause even when he is arguing for the truth and vice-versa. For example, in tort litigation regarding asbestos exposure, it is difficult for the plaintiffs to prove that their illness is caused by asbestos exposure and how serious is the potential consequences of their current symptoms.

where, when the Defendant is innocent we have $\varphi \in (0, \frac{1}{2})$ and, similarly, when the Defendant is guilty, $\varphi \in (\frac{1}{2}, 1)$. We assume that the function $f(\cdot)$ is monotonically increasing in its argument and captures the importance of resources devoted on evidence production. The parameter φ captures the fact that truth does matter in the production of evidence. Given the restriction on the range of φ , it follows that the party arguing for the truth, would be able to supply a higher valued evidence if $R_d = R_p > 0$.

Accordingly, substituting these evidence production functions into (7), we obtain:¹⁶

$$\pi^*(R_p, R_d) = \frac{\pi\varphi f(R_p)}{(1-\pi)(1-\varphi)f(R_d) + \pi\varphi f(R_p)} \quad (8)$$

It can be clearly seen in this additive form how the probability of guilt (and, therefore, the probability of innocence) is affected by the three main factors: (i) The *truth*, as represented by the parameter φ ; (ii) the *preconceptions or bias* of the court, as represent by the prior of guilt π (as well as the bias parameter λ that was set to 1 for brevity); and (iii) the *efforts* of the two sides (R_p and R_d). One advantage of this functional form is that it can easily generalize to more than two possibilities and to a higher number of participants.

3.2 Deciding on guilt versus innocence

So far we have seen how evidence production functions and the court's inference from the presented evidence can determine its posterior probability of guilt. However, ultimately, we are interested in deriving the contest success functions that capture the win probabilities as perceived by the Defendant and Plaintiff at the beginning of stage 1. To do this, we must consider alternative decision rules that the court might use to arrive at its verdict and below we present two possibilities and their implications for Plaintiff and Defendant's win probabilities.

¹⁶Versions of this functional form have been employed, among others, by Farmer and Pecorino (1999), Hirshleifer and Osborne (2001), Vaidya (2005), and Robson and Skaperdas (2008).

Suppose that the court makes a probabilistic decision on guilt (akin to tossing an unfair coin) as follows:

Choose G with probability $\pi^*(R_p, R_d)$ and I with probability $1 - \pi^*(R_p, R_d)$ (Rule 1)

With the above decision rule, it is immediately obvious that the contest success function representing the win probability of the Plaintiff $P_P(R_p, R_d)$ would be the posterior probability of guilt in (7) as stated below.

Proposition 2: *Suppose the contestants know the court's prior with certainty and the court uses Rule 1. Then the probability of the Plaintiff winning (and of the Defendant losing) is:*

$$P_P(R_p, R_d) = \pi^*(R_p, R_d) = \frac{\pi F_p(R_p)}{(1 - \pi)F_d(R_d) + \pi F_p(R_p)}$$

If, however, the court employs a discrete rule like “guilty beyond reasonable doubt” or just “choose guilt if and only if there is a better than even chance of guilt,” then (7) does not in general describe the appropriate contest success function from the point of view of the contestants. More precisely, suppose the court's decision rule is given by:

Choose G if and only if $\pi^*(R_p, R_d) > \gamma$ where $\gamma \in (0, 1)$ (Rule 2)

Obviously γ should be sufficiently close to 1 when the “beyond reasonable doubt” rule is employed and it equals $\frac{1}{2}$ when the even-chance rule is used. In the appendix we derive both decision rules as the outcome of a court's maximization problem.

With this decision rule, if the Defendant and Plaintiff know the prior of the court with certainty, then it can be easily established that the contest success function is an asymmetric perfectly-discriminatory one (or, the all-pay auction). If the contestants do not know the court's prior with certainty but rather have a common nondegenerate distribution about the court's prior (that is, the common prior of the contestants over

the court's prior), then we can expect an imperfectly discriminating contest success function. In what follows we assume a uniform prior distribution when this case applies. In particular, consider the following proposition which describes the contest success functions applicable when court's decision rule is discrete:

Proposition 3: (i) *Suppose the contestants know the court's prior with certainty and the court uses Rule 2. Then, the probability of the Plaintiff winning (and of the Defendant losing) is:*

$$P_P(R_p, R_d) = \begin{cases} 1 & \text{if } \frac{F_p(R_p)}{F_d(R_d)} > \frac{(1-\pi)\gamma}{\pi(1-\gamma)} \\ 0 & \text{if } \frac{F_p(R_p)}{F_d(R_d)} \leq \frac{(1-\pi)\gamma}{\pi(1-\gamma)} \end{cases}$$

(ii) *Suppose the contestants have a uniform prior about the court's prior and the court uses Rule 2. Then, the probability of the Plaintiff winning (and of the Defendant losing) is:*

$$P_P(R_p, R_d) = \frac{(1-\gamma)F_p(R_p)}{(1-\gamma)F_p(R_p) + \gamma F_d(R_d)}$$

Proof: Please see Appendix.

The contest success function in (i) differs from typical applications of all-pay auctions (Hillman and Riley, 1989, Baye et. al., 1996, Che and Gale, 1998) only in that it is asymmetric and also it does not include an outcome that has a probability of 1/2 when the probability of guilt just equals γ .¹⁷

In part (ii) we have a regular additive contest success function which is also asymmetric but without the priors playing any role (because they have cancelled out due to the uniformity of the contestant's prior over the court's prior) and with the decision parameter γ playing an important role whenever it differs from 1/2. When that parameter is close to 1, of course, the Plaintiff's position is hard and can be overcome only with the investment of vastly greater resources than the Defendant or with a much more favorable evidence production function (it helps to have the truth with you and be able to bring evidence in its favor more easily than your opponent).

¹⁷This latter difference, of course, could simply be eliminated by having the court flip a fair coin when that is the case.

Thus, it is clear that both types of contest success functions in Proposition 3 allow for the resources, biases, and the truth to affect outcome, albeit in different ways and taking into account the relative difficulties of proving one case or the other as indicated by the “handicap” parameter γ .

4 Stochastic Evidence production and the Difference Form

In this section we examine another evidence production process where the link between resources devoted and evidence obtained is probabilistic. As before, our ultimate goal is to show how such a process can lead to functional forms that link resources with the win probabilities of the contestants wishing to influence the decision of the court or the relevant audience. We begin by characterizing such an evidence production process and examine its impact on the court’s inference. Suppose each of the two parties can produce one piece of evidence. The Plaintiff can either produce evidence of guilt, denoted by e_p , or offer no evidence, e_\emptyset . Similarly the Defendant can either produce evidence in favor of innocence, denoted by e_d , or offer no evidence, denoted by e_\emptyset . As with the previous section, the court’s inference of guilt is based purely on observed evidence and without being aware of the evidence production process. Accordingly, given evidence (e_k, e_l) (where $k \in \{p, \emptyset\}$ and $l \in \{d, \emptyset\}$), let $L^g(e_k, e_l)$ denote the court’s likelihood of guilt purely on the basis of the evidence pair it observes.

Note for any given prior beliefs π , if e_p and e_d are to be potentially useful pieces of evidence to the two parties, then it must be the case that $L^g(e_k, e_l)$ observes the following properties:

$$L^g(e_p, e_\emptyset) > L^g(e_\emptyset, e_\emptyset) = 1 > L^g(e_\emptyset, e_d) \tag{9}$$

and

$$L^g(e_p, e_\emptyset) > L^g(e_p, e_d) > L^g(e_\emptyset, e_d) \tag{10}$$

For now, we impose no other restrictions on these likelihood ratios, including on the relationship between $L^g(e_p, e_d)$ and $L^g(e_\emptyset, e_\emptyset)$.

The posterior probability of guilt $\pi^*(e_k, e_l)$ for each pair of evidence (e_k, e_l) will be given by (4) using $L^g(e_k, e_l)$. From this it clearly follows that:

$$\pi^*(e_p, e_\emptyset) > \pi^*(e_\emptyset, e_\emptyset) = \pi > \pi^*(e_\emptyset, e_d) \quad (11)$$

$$\pi^*(e_p, e_\emptyset) > \pi^*(e_p, e_d) > \pi^*(e_\emptyset, e_d) \quad (12)$$

Both the Plaintiff and the Defendant are supposed to know the court's posteriors for any given combination of evidence. Unlike the previous section where evidence was a deterministic function of resources expended, here the resources devoted by either party determine the probability of finding favorable piece of evidence. Let $h_d(R_d)$ denote the probability that the Defendant will find evidence in favor of innocence (e_d). This probability is increasing in R_d , the resources expended on finding that evidence. Similarly, let $h_p(R_p)$ denote the probability that the Plaintiff will find evidence in favor of guilt (G), with that probability also increasing in the resources R_p .¹⁸ Thus, in terms of evidence there are four possible states of the world that can be faced by the court: (e_p, e_d) , (e_\emptyset, e_d) , (e_p, e_\emptyset) , and $(e_\emptyset, e_\emptyset)$ occurring with the following probabilities: $h_p(R_p)h_d(R_d)$, $[1 - h_p(R_p)]h_d(R_d)$, $h_p(R_p)[1 - h_d(R_d)]$, and $[1 - h_p(R_p)][1 - h_d(R_d)]$, respectively.¹⁹

¹⁸These probabilities should of course be related to the true state of the world; that is, the probabilities should depend on whether the defendant is truly guilty or innocent. We shall discuss this issue later.

¹⁹There are some interesting similarities between this setting and the work of Lagerlof (2007) which interprets efforts of competitors in a rent-seeking game as attempts to supply information to the distributor of the prize. In particular, in his model, the distributor does not have information about a firm's type, but each firm privately knows his type: efficient or inefficient. Firms in turn can spend resources to produce verifiable evidence of their type which they may choose to reveal to the distributor. As with our stochastic evidence production model, the link between resources and evidence is stochastic. The more resources a firm devotes, the greater is the probability of finding evidence. Further, as with our model, the distributor of the prize is not aware either of the level of resources devoted by any firm nor the underlying probability functions that govern the discovery of evidence. However, in Lagerlof's model, evidence is hard information. Efficient firms can either

Given the posterior probability of guilt that will be induced by each realized combination of evidence and given the functions $h_p(\cdot)$ and $h_d(\cdot)$, the expected posterior probability of guilt as a function of the resources expended by the two sides can be straightforwardly calculated:

$$\begin{aligned}
E[\pi^*] &= \pi^*(e_\emptyset, e_\emptyset) + Bh_p(R_p) - \Delta h_d(R_d) + Ah_p(R_p)h_d(R_d) & (13) \\
\text{where } B &\equiv \pi^*(e_p, e_\emptyset) - \pi^*(e_\emptyset, e_\emptyset) > 0, \Delta \equiv \pi^*(e_\emptyset, e_\emptyset) - \pi^*(e_\emptyset, e_d) > 0 \\
\text{and } A &\equiv \pi^*(e_p, e_d) - \pi^*(e_p, e_\emptyset) + \pi^*(e_\emptyset, e_\emptyset) - \pi^*(e_\emptyset, e_d).
\end{aligned}$$

We proceed as in the previous section by considering the two types of decision rules that the court might use to arrive at its verdict. First, consider the case whereby the court makes a probabilistic decision on guilt (akin to tossing an unfair coin) as follows:

Choose G with probability $\pi^*(e_k, e_l)$ and I with probability $1 - \pi^*(e_k, e_l)$ (Rule 1')

It is then straightforward to show that the contest success function from the viewpoint of the Plaintiff and the Defendant are directly obtained from (13) and it has similar qualitative properties to the additive forms of the previous sections.

Proposition 4: *Suppose the contestants know the court's prior with certainty*

find verifiable evidence of their type or nothing at all. Similarly, an inefficient firm can either find verifiable evidence of their inefficiency or nothing at all. Hence in his model, only the efficient have an incentive to lobby. However, we feel that such an information environment would not necessarily be suitable to settings such as courts where even though the truth may be fundamentally with only one side, both sides have scope to produce convincing bits of evidence as in the case of our model. Hence, instead of the hard information definition, we define evidence as simply information that appears to the court as convincing and there is scope for both sides to present favorable evidence including the side not arguing for the truth and the court's inference is completely separate from the production functions of evidence. This feature of our model also sets it apart from Lagerlof (1997) which also incorporates a hard evidence stochastic information model in the context of lobbying as information provision. Further, unlike Lagerlof (1997) our decision maker does not care about the welfare of either contesting parties but simply about the quality of his judgement per se.

and the court uses Rule 1'. Then the probability of the Plaintiff winning (and of the Defendant losing) is given by:²⁰

$$P_P(R_p, R_d) = E[\pi^*] \text{ described in (13)}$$

$$\text{with } \frac{\partial P_P}{\partial R_p} > 0, \frac{\partial P_P}{\partial R_d} < 0 \text{ and } \frac{\partial P_P}{\partial \pi} > 0.$$

Proposition 4 suggests a difference-form contest success function where the win-probability of the Plaintiff (same as expected posterior probability of guilt) is increasing in the resources R_p expended by the Plaintiff and decreasing in the resources R_d expended by the Defendant. Naturally, the win probability of the Defendant $1 - P_P(R_p, R_d)$, has the reverse properties in terms of the resources expended by each side. The parameters B and Δ are similar and have natural interpretations. $B(\equiv \pi^*(e_p, e_\emptyset) - \pi^*(e_\emptyset, e_\emptyset))$ represents the marginal increase in the court's posterior as a result of the Plaintiff producing evidence of guilt in the absence of evidence of innocence, whereas Δ represents the marginal decrease in the court's posterior as a result of the Defendant producing evidence of innocence in the absence of evidence of guilt. Parameter A can take positive, zero, or negative values. If A is positive, then the Plaintiff has an advantage in the sense that the greater the product $h_p(R_p)h_d(R_d)$ is, the greater is the Plaintiff's probability of winning. The effect on the Defendant would be exactly the opposite. Thus, a non-zero A confers an asymmetric advantage to one of the contestants that is proportional to the $h_p(R_p)h_d(R_d)$ term. To better the intuition behind the value of A consider the symmetry condition $\pi^*(e_p, e_d) = \pi^*(e_\emptyset, e_\emptyset)(= \pi)$ (see (11) and (12)). Then, $A = \Delta - B$, and thus the party that can induce the higher aforementioned marginal increase in the court's posterior has the significant advantage in the contest. That is, a non-zero A , and the advantage it confers to one party, stems from the tilt that the court might have in evaluating favorable evidence to that party versus its opponent.

We now allow for the court to use the alternative decision rule, similar to rule 2,

²⁰The comparative statics can be easily verified using the restrictions in (11) and (12) and the properties of the evidence production functions $h_p(\cdot)$ and $h_d(\cdot)$.

and derive the operative contest success functions as perceived by the two parties:

Choose G if and only if $\pi^*(e_k, e_l) > \gamma$ for every (e_k, e_l) where $\gamma \in (0, 1)$ (Rule 2')

As with the case of deterministic evidence production in the previous section, we consider two cases: (i) when the contestants know the court's prior with certainty and (ii) when the contestants have a uniform prior over the court's prior. However, in the second case it is convenient to parameterize the resultant posteriors in the four states as follows:²¹

$$\begin{aligned}\pi^*(e_p, e_d) &= \pi^*(e_\emptyset, e_\emptyset) = \pi; \\ \pi^*(e_\emptyset, e_d) &= \delta\pi \text{ for some } \delta \in (0, 1); \\ \pi^*(e_p, e_\emptyset) &= \Gamma\pi \text{ for some } \Gamma > 1\end{aligned}\tag{14}$$

Proposition 5: (i) *Suppose the contestants know the court's prior with certainty and the court uses Rule 2'. Then, the probability of the Plaintiff winning (and of the Defendant losing) is:*

$$P_P(R_p, R_d) = \begin{cases} 0 & \text{if } \gamma \in [\Gamma\pi, 1) \\ h_p(R_p)[1 - h_d(R_d)] & \text{if } \gamma \in [\pi, \Gamma\pi) \\ 1 - [1 - h_p(R_p)]h_d(R_d) & \text{if } \gamma \in [\delta\pi, \pi) \\ 1 & \text{if } \gamma \in (0, \delta\pi) \end{cases}\tag{15}$$

(ii) *Suppose the contestants have a uniform prior about the court's prior and the court uses Rule 2'. Then, the probability of the Plaintiff winning (and of the*

²¹Obviously the inequalities in (11) and (12) are satisfied since

$$\Gamma\pi = \pi^*(e_p, e_\emptyset) > \pi = \pi^*(e_p, e_d) = \pi^*(e_\emptyset, e_\emptyset) > \delta\pi = \pi^*(e_\emptyset, e_d)$$

Defendant losing) is:

$$P_P(R_p, R_d) = \begin{cases} 1 - \gamma + \gamma \left[\frac{\Gamma-1}{\Gamma} h_p(R_p) - \frac{1-\delta}{\delta} h_d(R_d) + \left(\frac{1-\delta}{\delta} - \frac{\Gamma-1}{\Gamma} \right) h_p(R_p) h_d(R_d) \right] & \text{if } \delta > \gamma \\ 1 - \gamma + \gamma \left(\frac{\Gamma-1}{\Gamma} \right) h_p(R_p) - (1 - \gamma) h_d(R_d) & \\ + \left((1 - \gamma) - \gamma \left(\frac{\Gamma-1}{\Gamma} \right) \right) h_p(R_p) h_d(R_d) & \text{if } \delta \leq \gamma \end{cases} \quad (16)$$

Proof: Please see Appendix.

The form in part (i) is a mixture of the all-pay auction and a probabilistic contest success function. In settings in which the cut-off parameter γ is sufficiently close to the prior π , the probabilistic part would apply and the choices by the contestants have an effect. When γ is close enough to 1 (and higher than $\Gamma\pi$), as it could be in the “beyond all reasonable doubt” case, there is nothing the Plaintiff can do to increase the expected probability of winning beyond zero.

Perhaps of more applied interest is the form in part (ii), especially when $\delta > \gamma$. In particular, when $\frac{1-\delta}{\delta} = \frac{\Gamma-1}{\Gamma}$, the probability of winning for each party depends only on the difference between $h_p(R_p)$ and $h_d(R_d)$ (i.e. on $h_p(R_p) - h_d(R_d)$) barring a constant term. Further, when $\gamma = 1/2$, as we might expect it to be in cases of product or political choice, the probability of the plaintiff winning takes the following simple form:

$$P_P(R_p, R_d) = \frac{1}{2} + \frac{\alpha}{2} [h_p(R_p) - h_d(R_d)] \text{ where } \alpha \equiv \frac{1-\delta}{\delta} = \frac{\Gamma-1}{\Gamma} < 1 \quad (17)$$

The function in (17) resembles the piece-wise linear difference form contest explored by Che and Gale (2000). However, there are important differences. (17) has inferential foundations and it can be non-linear in R_p and R_d with h_p and h_d interpreted as probabilities and bounded naturally between 0 and 1. Also notice that in (17), the parameter α depends on the parameters Γ and δ which represent the force of the evidence – against the Defendant when Γ is high and in favor of her when δ is low – by the court. The higher is Γ and the lower is δ , the higher is α

and the easier it is to convince the court in a contestant’s favor, conditional on one contestant producing evidence and the opponent not producing any evidence. But in expectation, under symmetry, no side has an advantage although they both have a greater return to their resource investment the higher is α .

As with the previous section, even with stochastic evidence production, we would also expect the truth to play a role in uncovering evidence apart from resources and the functional forms in Propositions 4 and 5 can accommodate this feature easily. To see this, suppose for all R we can expect to have $h_p(R; G) > h_p(R; I)$, where G, I represent the truth. That is, we can expect it to be easier for the Plaintiff to find evidence of guilt when Defendant is guilty than when the Defendant is innocent. Similarly, for all R we can expect $h_d(R; G) < h_d(R; I)$. How much easier or more difficult it is to find favorable evidence will of course depend on the particular circumstances, on the particular “technology” of evidence production. Following the example of Hirshleifer and Osborne (2001) and others, one way of parameterizing the two functions is to have $h_p(R_p) = \theta h(R_p)$ and $h_d(R_d) = (1 - \theta)h(R_d)$, for some increasing function $h(\cdot)$, $\theta > 1/2$ when the Defendant is guilty, and $\theta < 1/2$ when the Defendant is innocent. The closer θ is to 1 (when the Defendant is guilty) or the closer it is to 0, the more discerning of the truth can the technology of evidence production be thought of. The functional forms in propositions 4 and 5 can easily capture such a truth-sensitive evidence production process, by simply substituting for $h_p(\cdot)$ and $h_d(\cdot)$ using the above functions.

5 Concluding remarks

Contest success functions have been used in a variety of economic applications that involve attempts by contending parties to influence the opinion or judgement of an audience by devoting costly resources into evidence production. In this paper, we provide an explicit inferential justification for using such functions and by doing so provide a better foundation for a substantial body of existing research. In particular, by envisaging different types of evidence production functions, we derive two functional forms that provide an explicit link between resources devoted and their

impact on the court's inference about the truth and its verdict. One of the functional forms happens to be the relatively well-known additive function which has been recently used to model the winning probabilities of contestants in rent-seeking, litigation, election campaigns, advertising, and other settings. We also uncover a less well-known difference functional form (similar to Che and Gale (2000)) which could also be suitably applied to such settings. By linking the contest success functions to a Bayesian inference process, we are able to clearly distinguish the role played by priors, psychological biases, the truth and the resources in influencing the relevant audience's judgement by both re-interpreting known parameters and introducing new parameters in these functional forms. Notice also that while the functional forms demarcate the role of priors from the impact of resources in any given period, in a dynamic setting they would also indirectly allow for resources put into evidence production in the current period to influence future priors assuming the current period posterior would become next period's prior. Thus, these functions would be useful in a variety of settings, examining broadly the link between costly adversarial production of evidence or arguments and an audience's judgment given its initial set of priors or biases. Hence, to provide but a few examples, these functions can be suitably employed to better explore the links between campaign contributions and electoral outcomes, advertising expenditures and creation of "brand loyalty" among many others.

More generally, the interpretation of persuasion as a contest forces one to think about its inputs very differently from those of traditional production functions. This is because, a main substantive difference between contest success functions and production functions is that the inputs in the former are combined adversarially whereas in the latter they are combined cooperatively. Then, in contrast to ordinary production large increases in the inputs of persuasion as employed by all contestants could well yield no change in the final outcome (the probability of winning for each adversary). In the case of litigation, for example, a plaintiff and a defendant can both dramatically increase their legal expenditures without changing their probabilities of winning in court (or, the probability of the court discovering the truth). Similarly, lobbying organizations advocating opposing policies can have the same probabilities

of prevailing with widely different levels of expenditures. Given that much economic activity in modern economies, especially in the service sector, involves such adversarial combination of inputs, we must be careful to not evaluate these activities in the same cooperative way as contributions of traditional inputs are valued in light of the standard production theory. Exploring further, then, the theoretical and empirical implications of the adversarial combination of inputs that is found in such contests of persuasion would be a natural next step.

APPENDIX

Proof of Proposition 3, (i): Given Rule 2 and the form of $\pi^*(R_p, R_d)$ in (7), the Plaintiff will win if and only if:

$$\frac{\pi F_p(R_p)}{(1 - \pi)F_d(R_d) + \pi F_p(R_p)} > \gamma$$

or, if and only if

$$\frac{F_p(R_p)}{F_d(R_d)} > \frac{(1 - \pi)\gamma}{\pi(1 - \gamma)}$$

Since both $F_p(\cdot)$ and $F_d(\cdot)$ are deterministic functions of their respective arguments and π is known with certainty, we obtain the probability of winning in (i).

Proof of Proposition 3, (ii): As in part (i), Rule 2 and the form of $\pi^*(R_p, R_d)$ in (7) imply:

$$\frac{\pi F_p(R_p)}{(1 - \pi)F_d(R_d) + \pi F_p(R_p)} > \gamma$$

or for any given prior π , we have

$$\pi > \frac{\gamma F_d(R_d)}{(1 - \gamma)F_p(R_p) + \gamma F_d(R_d)}$$

Given our assumption that π follows uniform distribution over $[0, 1]$, the probability of a guilty verdict is as follows:

$$\begin{aligned}
& Prob \left(\pi > \frac{\gamma F_d(R_d)}{(1-\gamma)F_p(R_p) + \gamma F_d(R_d)} \right) \\
&= 1 - \frac{\gamma F_d(R_d)}{(1-\gamma)F_p(R_p) + \gamma F_d(R_d)} \\
&= \frac{(1-\gamma)F_p(R_p)}{(1-\gamma)F_p(R_p) + \gamma F_d(R_d)}
\end{aligned}$$

which is the form in part (ii) of the proposition's statement. *QED*

Proof of Proposition 5, (i): There are four possible states (e_k, e_l) in calculating the (ex ante) probability. Given the decision rule, in each state the court will decide with certainty whether to rule in favor of the Plaintiff or not depending on whether $\pi^*(e_k, e_l)$ is greater or smaller than γ .

When $\gamma \in [\Gamma\pi, 1)$, there is no state in which the court can rule in favor of the Plaintiff and therefore $P_P(R_p, R_d) = 0$ in this case.

When $\gamma \in [\pi, \Gamma\pi)$, the court decides in favor of the Plaintiff only in state (e_p, e_\emptyset) . Thus, in this case we have $P_P(R_p, R_d) = h_p(R_p)[1 - h_d(R_d)]1 = h_p(R_p)[1 - h_d(R_d)]$.

When $\gamma \in [\delta\pi, \pi)$, the court decides in favor of the Plaintiff in states (e_p, e_\emptyset) , (e_p, e_d) , and $(e_\emptyset, e_\emptyset)$, and against the Plaintiff in state (e_\emptyset, e_d) , yielding $P_P(R_p, R_d) = 1 - [1 - h_p(R_p)]h_d(R_d)$.

Finally, when $\gamma \in (0, \delta\pi)$, the court decides in favor of the Plaintiff in all states and thus $P_P(R_p, R_d) = 1$.

Proof of Proposition 5, (ii): The Plaintiff is aware that for any given pair of resource investment (R_p, R_d) , one of the four alternative states (e_k, e_l) will exist. In each state, depending on the realized value of π the Plaintiff will have a chance of getting the verdict in his or her favor if and only if π is such that $\pi^*(e_k, e_l) > \gamma$.

We first suppose that $\delta > \gamma$ so that the Plaintiff has a positive probability of winning in each state. To identify the Plaintiff's probability of winning in each alternative state, let's examine each of them separately as follows.

State (e_p, e_\emptyset) :

This state occurs with an ex-ante probability of $h_p(R_p)[1-h_d(R_d)]$ and $\pi^*(e_p, e_\emptyset) = \Gamma\pi$. Hence in this state, the Plaintiff wins if $\Gamma\pi > \gamma$ or $\pi > \frac{\gamma}{\Gamma}$. Given uniform distribution of π , this occurs with probability $(1 - \frac{\gamma}{\Gamma})$.

State (e_p, e_d) :

This state occurs with an ex-ante probability of $h_p(R_p)h_d(R_d)$ and $\pi^*(e_p, e_d) = \pi$. Hence in this state, the Plaintiff wins if $\pi > \gamma$. Given uniform distribution of π , this occurs with probability $(1 - \gamma)$.

State $(e_\emptyset, e_\emptyset)$:

This state occurs with an ex-ante probability of $(1 - h_p(R_p))(1 - h_d(R_d))$ and $\pi^*(e_\emptyset, e_\emptyset) = \pi$. Hence in this state, the Plaintiff wins if $\pi > \gamma$. Given uniform distribution of π , this occurs with probability $(1 - \gamma)$.

State (e_\emptyset, e_d) :

This state occurs with an ex-ante probability of $(1-h_p(R_p))h_d(R_d)$ and $\pi^*(e_\emptyset, e_d) = \delta\pi$. Hence in this state, the Plaintiff wins if $\delta\pi > \gamma$ or $\pi > \frac{\gamma}{\delta}$. Given uniform distribution of π , this occurs with probability $(1 - \frac{\gamma}{\delta})$.

Using the above information, it is easy to appreciate that the Plaintiff's ex-ante probability of winning would be:

$$P_P(R_p, R_d) = h_p(R_p)[1-h_d(R_d)] (1 - \frac{\gamma}{\Gamma}) + h_p(R_p)h_d(R_d) (1 - \gamma) + (1-h_p(R_p))(1-h_d(R_d)) (1 - \gamma) + (1 - h_p(R_p))h_d(R_d) (1 - \frac{\gamma}{\delta})$$

By suitable re-arrangement, the above expression reduces to: $1 - \gamma + \gamma[\frac{\Gamma-1}{\Gamma}h_p(R_p) - \frac{1-\delta}{\delta}h_d(R_d) + (\frac{1-\delta}{\delta} - \frac{\Gamma-1}{\Gamma})h_p(R_p)h_d(R_d)]$.

When $\delta \leq \gamma$, $\frac{\gamma}{\delta} \geq 1$ and the δ term vanishes from the above expression. Then, the expression for the contest success function becomes:

$$(1 - \gamma) + \gamma \left(\frac{\Gamma-1}{\Gamma}\right) h_p(R_p) - (1 - \gamma)h_d(R_d) + ((1 - \gamma) - \gamma \left(\frac{\Gamma-1}{\Gamma}\right))h_p(R_p)h_d(R_d). QED$$

A derivation of decision rules 1 and 2:

We begin by showing how the decision rule 2 could arise out of a court's concern towards minimizing the expected cost of making an incorrect verdict. Following each of the two alternative verdicts, G and I , the court is aware that there can be two mutually exclusive states. In one state, the verdict would be correct in

which case we assume that the court suffers no cost so its utility is set at 0. In the other state, the verdict would be incorrect in which case the court suffers a loss in utility. We assume that this loss in utility is L from an incorrect G verdict while it is αL from an incorrect I verdict where $\alpha, L > 0$. Accordingly, following the court's determination of π^* , its expected utility from either verdict is as follows:

$$\begin{aligned} G & : \pi^*(0) + (1 - \pi^*)(-L) = -(1 - \pi^*)L \\ I & : \pi^*(-\alpha L) + (1 - \pi^*)(0) = -\pi^*\alpha L \end{aligned}$$

Hence to maximize its expected utility from the verdict, the court rules G if and only if $-(1 - \pi^*)L > -\pi^*\alpha L$ i.e. if $\pi^* > \gamma \equiv \frac{1}{1+\alpha}$ which is decision rule 2 where the level of γ depends on the cost parameter α which measures the cost to the court of acquitting the guilty relative to punishing the innocent.

Rule 1 can in turn be interpreted as P and D 's perception of how the court decides on its verdict following its determination of π^* when although they know that court uses rule 2, neither P nor D know the true value of γ but rather have a common prior about it being distributed uniformly over $[0, 1]$. Given this prior, once π^* is determined, it follows immediately from rule 2 that the probability of the court making a guilty verdict from P and D 's perspective is $\text{Prob}[\gamma < \pi^* | \pi^*] = \pi^*$ which yields rule 1.

REFERENCES

- Aragones, Enriqueta, Gilboa, Itzhak, Postlewaite, Andrew, and Schmeidler, David. "Fact-Free Learning", *American Economic Review*, 2005, 95, pp. 1355-68.
- Baron, David P. "Electoral Competition with Informed and Uninformed Voters," *American Political Science Review*, 1994, 88(1), pp. 33-47.
- Baye, Michael R. and Hoppe, Heidrun C. "The Strategic Equivalence of Rent-Seeking, Innovation, and Patent-Race Games," *Games and Economic Behavior*, 2003, 44(2), pp. 217-26.

Baye, Michael R., Kovenock, Dan and de Vries, Casper G. "The all-pay auction with complete information," *Economic Theory*, 1996, 8, pp. 291-305.

Bell, David E., Keeney, Ralph L. and Little, John D. C. "A market share theorem," *Journal of Marketing Research*, 1975, 12(2), pp. 136-41.

Bernardo, Antonio E., Talley, Eric, and Welch, Ivo. "A Theory of Legal Presumptions," *Journal of Law, Economics, and Organization*, 2000, 16(1), pp. 1-49.

Che, Yeon Koo and Gale, Ian L. "Caps on political lobbying," *American Economic Review*, 1998, 88, pp. 643-51.

Che, Yeon Koo and Gale, Ian L. "Difference-Form Contests and the Robustness of All-Pay Auctions", *Games and Economic Behavior*, 2000, 30, pp. 22-43.

Cialdini, Robert B. *Influence: Science and Practice*, Fourth Edition, Needham Heights, MA: Allyn & Bacon, 2001.

Clark, Derek and Riis, Christian. "Contest success functions: an extension," *Economic Theory*, 1998, 11(1), pp. 201-4.

Daughety, Andrew F. and Reinganum, Jennifer F. "On the Economics of Trials: Adversarial Process, Evidence, and Equilibrium Bias," *Journal of Law, Economics, and Organization*, 2000, 16(2), 365-394.

Dewatripoint, Mathias and Tirole, Jean. "Advocates," *Journal of Political Economy*, 1999, 107, pp. 1-39.

Epstein, Gil S. and Nitzan, Shmuel. "The Politics of Randomness," *Social Choice and Welfare*, 2006, 27, 423-433.

Farmer, Amy and Pecorino, Paul. "Legal expenditure as a rent-seeking game," *Public Choice*, 1999, 100(3-4), pp. 271-88.

Filiz, Emel. "Incorporating Unawareness into Contract Theory," December 2006.

Froeb, Luke M. and Kobayashi, Bruce H. "Naive, biased, yet Bayesian: can juries interpret selectively produced evidence", *Journal of Law, Economics and Organization*, 1996, 12, pp. 257-71.

Fullerton, Ron L. and McAfee, R. Preston. "Auctioning Entry into Tournaments," *Journal of Political Economy*, 1999, 107(3), pp. 573-605.

Grossman, Gene and Helpman, Elhanan. "Protection for sale", *American Economic Review*, 1994, 84, pp. 833-50.

- Hillman, Arye and Riley, John G. "Politically contestable rents and transfers," *Economics and Politics*, 1989, 1, pp. 17-40.
- Hirshleifer, Jack. "Conflict and Rent-Seeking Success Functions: Ratio vs. Difference Models of Relative Success," *Public Choice*, 1989, 63(2), pp. 101-12.
- Hirshleifer, Jack and Osborne, Evan. "Truth, effort and the legal battle", *Public Choice*, 2001, 108(1-2), pp. 169-95.
- Hirshleifer, Jack and Riley, John G. *The Analytics of Uncertainty and Information*, Cambridge University Press, New York, 1992.
- Jia, Hao. "A Stochastic Derivation of the Ratio Form of Contest Success Functions," *Public Choice*, 2008, 135(3-4), 125-130.
- Kadane, Joseph B. and Schum, David A. *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*, Wiley, New York, 1996.
- Konrad, Kai A. "Bidding in hierarchies," *European Economic Review*, 2004, 48(6), pp. 1301-8.
- Konrad, Kai A. *Strategy in Contests - An Introduction*, monograph manuscript, January 2007.
- Kuhn, Thomas. *The Structure of Scientific Revolutions*, 3rd Edition, Chicago: University of Chicago Press, 1996.
- Laband, David N. and Sophocleus, John P. "An estimate of resource expenditures on transfer activity in the United States", *Quarterly Journal of Economics*, 1992, 107, pp. 959-83.
- Lagerlof, Johan N. "Lobbying, information, and private and social welfare", *European Journal of Political Economy*, 1997, 13, pp. 615-37.
- Lagerlof, Johan N. "A theory of rent seeking with informational foundations", *Economics of Governance*, 2007, 8, pp. 197-218.
- Legros, Patrick and Newman, Andrew F. "Courts, contracts, and interference", *European Economic Review*, 2002, 46, pp. 734-44.
- Lippmann, Walter. *Public Opinion*, New York: Free Press, 1997(1922).
- Luce, R. Duncan. *Individual Choice Behavior*, Wiley, New York, 1959.
- Milgrom, Paul and Roberts, John. "Relying on the information of interested parties", *RAND Journal of Economics*, 1986, 17, pp. 18-32.

Mueller, Holger and Warneryd, Karl. "Inside Versus Outside Ownership: A Political Theory of the Firm," *RAND Journal of Economics*, 2001, 32(3), pp. 527-41.

Mullainathan Sendhil, Schwartzstein, Joshua and Shleifer Andrei. "Coarse Thinking and Persuasion", manuscript, 2006.

Nitzan, Shmuel. "Modeling rent-seeking contests," *European Journal of Political Economy*, 1994, 10(1), pp. 41-60.

Ozbay, Orkut Y. "Unawareness and Strategic Announcements in Games with Uncertainty," November 2006.

Potters, Jan and Van Winden, Frans. "Lobbying and asymmetric information", *Public Choice*, 1992, 74, pp. 269-92.

Robson, Alexander and Skaperdas, Stergios. "Costly Enforcement of Property Rights and the Coase Theorem," *Economic Theory*, July 2008, 36(1), 109-128

Schmalensee, Richard. *The Economics of Advertising*, North-Holland, Amsterdam-London, 1972.

Shin, Hyun S. "Adversarial and inquisitorial procedures in arbitration", *RAND Journal of Economics*, 1998, 29, pp. 378-405.

Sinn, Hans-Werner. *Economic Decisions under Uncertainty*, North Holland, Amsterdam, 1983.

Skaperdas, Stergios. "Contest success functions", *Economic Theory*, 1996, 7, pp. 283-90.

Skaperdas, Stergios and Grofman, Bernard G. "Modeling Negative Campaigning," *American Political Science Review*, 1995, 89(1), pp. 49-61.

Snyder, James M., "Election Goals and the Allocation of Campaign Resources," *Econometrica*, May 1989, 57(3), 637-660.

Stevens, Stanley S. "A Metric for Social Consensus", *Science*, 1966, 151, pp. 530-41.

Stevens, Stanley S. *Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects*, Wiley, New York, 1975.

Tullock, Gordon. "Efficient rent seeking", in: J.M. Buchanan, R.D. Tollison, and G. Tullock, eds., *Toward a Theory of the Rent Seeking Society*, Texas A&M

University Press, College Station, TX, 3-15, 1980.

Vaidya, Samarth. "Corruption in the Media's Gaze", *European Journal of Political Economy*, 2005, 21(3), pp. 667-87.

Vaidya, Samarth. "The Nature of Corruption Deterrence in a Competitive Media Sector", *Economics of Governance*, 2006, 7(3), pp. 229-43.

Warneryd, Karl. "Distributional Conflict and Jurisdictional Organization," *Journal of Public Economics*, 1998, 69(3), pp. 435-450.