Persuasion as a Contest

Stergios Skaperdas* and Samarth Vaidya**

Preliminary draft: October 14, 2005
revised: December 28, 2006

ABSTRACT: From marketing and advertising to political campaigning and court proceedings, contending parties expend resources to persuade an audience of the correctness of their view. We examine how the probability of persuading the audience depends on the resources expended by the parties, so that persuasion can be modelled as a contest. We use a "small-world" Bayesian approach whereby the audience makes inferences solely based on the evidence presented to them. The evidence is produced by the resources expended by the contending parties, but these resources are not taken into account by the audience. We find conditions on evidence production and likelihood functions that yield the well-known additive contest success functions, including the logit function as well as the one that is perfectly discriminatory. We also find conditions that produce a different functional form than the ones previously used. In all cases, there are three main determinants of which side the audience chooses: (i) the truth; (ii) the biases of the audience as distilled in their priors; and (iii) the resources expended by the parties interested in persuading the audience.

*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 Kalle Moene and Karl Warneryd for comments and discussions, as well participants in the conference "advances in the theory of contests and tournaments" at the Social Science Research Council-Berlin (WZB) in October, 2005, and seminar participants at Drexel, 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, 1978) 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 (Baron, 1994, Skaperdas and Grofman, 1995), bureaucratic organization and corporate governance (Mueller and Warneryd, 2001, Konrad, 2004, Castillo and Skaperdas, 2005), and media politics (Vaidya, 2005a,b).

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, Jia, 2005), 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, 2005). 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 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:

$$p_i(r_i, r_j) = \frac{f(r_i)}{f(r_i) + f(r_j)} \tag{1}$$

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 first find necessary conditions that yield (1) as well as more general forms of it. We suppose that the audience is a "small-world" Bayesian, in the sense that it 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. Juries and judges are formally closest to such an assumption, but we think in other instances of persuasion it is a reasonable first approximation, especially when compared to another extreme in which the audience would have full knowledge (or, a probability distribution over) the objectives and strategies of probable or possible contestants far removed from the audience itself. 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 "small 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.

The functional form in (1) and its variations are derived 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., 1986, Che and Gale, 1998). Thus, we find a clean set of circumstances to which the perfectly and imperfectly 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:

$$p_i(r_i, r_j) = \frac{1}{2} + \frac{\alpha}{2}[h(r_i) - h(r_j)] \tag{2}$$

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). To our knowledge, (2) and variations have not been used in contests. In some sense, and for some contexts, the derivation of (2) might be thought of as being more straightforward and rigorous than the derivation of (1) that we obtain in this paper. It would worth exploring the properties and implications of (2) in future research.

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 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]

Our focus is rather different from other work in economics that is related to the mechanics of persuasion. One group of economists have focussed on identifying equilibrium conditions and information environments under which such adversarial generation of evidence could actually contribute to effective decision making. For example, Milgrom and Roberts (1986) show that when verifiable and conclusive evidence can be obtained by the parties without cost and they are constrained to making only non-falsifiable assertions, their incentives to compete with each other could lead to perfect information getting transmitted to a decision-maker even when he is strategically unsophisticated. Froeb and Kobayashi (1996) provide an example of how "naive" and "biased" juries presiding on evidence that is deliberately distorted could nonetheless reach correct verdicts as long as the party on the side of truth has a stronger incentive to supply evidence than its rival. Another group of economists have compared adversarial versus inquisitorial methods of evidence collection and interpreted institutions that rely on evidence presented by rival parties as optimal mechanisms for information generation.[2] Dewatripont and Tirole (1999) show that when collecting information is costly for agents, a process of information acquisition that relies on "advocates" who focus on collecting evidence favorable to only their cause (as they are rewarded only if their cause succeeds) can acquire information more cheaply than a system that relies on one agent to collect information sequentially on competing causes. The focus of their analysis is the moral hazard problem associated with agents having to put costly, unobservable effort and the incentive structure needed to overcome it under the two different methods. Sobel (1985) has also looked at the problem of eliciting costly verifiable information from agents: albeit his decision maker possesses fewer options to structure incentives as

[1] Campaign 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.

[2] See Dewatripoint and Tirole (1999) and Shin (1998).

compared to Dewatripont and Tirole (1999). Our approach and motivation are (as it will become evident if it has not become so from out description above) different and more positive in nature than in this literature. We just seek to find possible justification of functions that describe the economic activity of persuasion better than production functions do.

Before proceeding with the modeling, we briefly discuss the importance of persuasion 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.[3] 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.[4] 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.[5] 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 com-

---

[3] As 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. In the 2000 U.S. presidential race, George W. Bush and his main political rival Al Gore together spent over $300 million in election campaigning (see http://www.opensecrets.org/2000elect/index/AllCands.htm for further information).

[4] In 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). Large law firms in the U.S. employ more than 1500 full time equivalent attorneys each (See 2002 NLJ 250 Annual survey at http://www.law.com/special/professionals/nlj250/2002/nlj250.shtml as compiled by *The National Law Journal* in 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). This is not surprising, given the sheer sizes of verdicts passed by juries; In 2002, the highest jury verdict crossed the $25 billion mark (see NLJ Verdicts 100 at http://www.verdictsearch.com/news/specials/020303verdicts_chart.jsp).

[5] As an example of such a firm, see www.persuasionstrategies.com.

panies 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. Overall, many parts of the service economy and their employees can be reasonably be thought of as engaging in persuasion.

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 several other contexts as well. We follow a Bayesian approach[6] not because we think courts and individuals necessarily follow such an inference process, but because the approach is relatively simple. We do expect actual courts, citizens, and consumers to exhibit systematic psychological biases which could be used by the parties engaged in persuasion and the implications of such tendencies would need to be examined in future work.[7]

2 Evidence Production and Inference

For concreteness we use a court setting to lay out the building blocks of persuasion. The defense (denoted by D) and the prosecution or plaintiff (denoted by P) 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

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

[7]See Lam (2001) for less orthodox ways of looking at the problem.

is innocent (I). We consider two main stages in the process of evidence production and its assessment by the court:

1. Firstly, we posit that collection of evidence is costly; the defense and the prosecution expend resources R^d and R^p to gather evidence favorable to their cause. Hence, even the side arguing for the truth, must put in effort to find any relevant evidence. They then present their evidence to the court. We assume that this is the only additional information that the court obtains during the trial over and above its priors.

2. Based on the evidence presented, the court makes an inference about guilt or innocence. The court uses Bayes' rule to update its prior belief in light of the evidence and the resulting posterior belief represents the probability of guilt or innocence.

3. Given its posterior belief, the court makes a decision on guilt or innocence.

We take the resources expended in stage 1, R^d and R^p, as given. We are mainly concerned with the construction of posterior beliefs (as functions of R^d and R^p) in stage 2, although we also show what occurs in stage 3 when the court follows particular discrete-choice rules.

Beginning with stage 2, we follow the Bayesian approach of inference from evidence as, for example, has been used by Kadane and Schum (1996) in the pioneering analysis of the Sacco and Vanzetti trial evidence. Let the likelihood ratios L^G & L^I represent the force of the evidence (denoted by $e^D$and e^P respectively) presented by the defense and the prosecution towards establishing the guilt or innocence of the defendant. More precisely,

$$L^G = \frac{\Pr ob(e^D, e^P \mid G)}{\Pr ob(e^D, e^P \mid I)} \tag{3}$$

whereas L^I is just the inverse of L^G :

$$L^I = \frac{1}{L^G} = \frac{\Pr ob(e^D, e^P \mid I)}{\Pr ob(e^D, e^P \mid G)} \tag{4}$$

Leave aside for the moment the issue of how L^G and L^I are constructed by the court, given the evidence presented and how evidence production takes place. Let π represent the prior probability of guilt and let π^* represent posterior probability of guilt based on the evidence presented by the two sides i.e. $\pi^* = prob(G \mid e^D, e^P)$. Then according to Bayes' law:

$$\frac{\pi^*}{1-\pi^*} = \frac{\pi}{1-\pi} L^G$$

which straightforwardly implies

$$\pi^* = \frac{\pi L^G}{(1-\pi) + \pi L^G} \tag{5}$$

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.[8]

3 Deterministic Evidence Production and the Generalized Additive Function

As already mentioned, versions of the additive 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

[8] The court's reliance on the likelihood ratio of guilt based on evidence produced is akin to the notion of "preponderance of evidence": a standard used by judges to instruct the jury regarding how to reach a verdict. To quote Judge Robert M. Parker, who instructed the jury in the *Newman v. Johns-Manville* asbestos exposure case, preponderance of evidence implies "whether it's more likely so than not so. In other words, ... such evidence when compared and considered, compared with the evidence against it produces in your minds that what is sought to be true is more likely true than not true, tipping of the scales, one way or the other" (see Selvin and Picus (1987), pg. 19).

inferential settings is still lacking. In the next section, we uncover some evidence production processes and information environments which could support the usage of the logit function in such contexts.

3.1 Evidence production and resources

Let's suppose that the court ranks all evidence that favors the defendant and the plaintiff individually on a $[0, \infty]$ scale with a higher number representing a more powerful evidence. During the trial, the court only observes a pair of evidence: one piece provided by each side. What pair of evidence does the court actually see? Given that finding evidence is costly, this depends on the amount of resources devoted by the plaintiff and the defendant. In particular we postulate the following evidence production functions:

$e^D = F^D(R^D)$ and $e^P = F^P(R^P)$ where both $F^D(\cdot)$ and $F^P(\cdot)$ are monotonically increasing in resources spent by the defendant and the plaintiff respectively. Since we don't expect the party that spends more to be necessarily on the side of the truth, the level of spending by itself is not informative and need not be helpful in guiding towards the truth. Hence without imposing any further restrictions on the above functions, the evidence they produce could be potentially misleading; the side with the truth might not put much effort and present a weaker evidence than the side that's lying.[9] However, since the court must decide on the basis of evidence actually presented, its inferred likelihood ratio of guilt would be determined by the evidence it observes and hence by the amount of resources put in by the two sides. The side that would put in more effort would come up with more powerful evidence relative to its rival and tilt the likelihood ratio in its favor. Hence, we can expect the likelihood of guilt L^G to become a function of R^P and R^D, denoted by $L^G(R^P, R^D)$, that is increasing in R^P and decreasing in R^D. From (2) it follows that the likelihood of innocence $L^I(R^P, R^D) = \frac{1}{L^G(R^P, R^D)}$, is obviously decreasing in R^P and increasing

[9] 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.

in R^D.

With evidence being a deterministic function of the resources also implies that the posterior probability of guilt as a function of resources reduces to that in (5):

$$p^G(R^d, R^p) = \frac{\pi L^G(R^d, R^p)}{(1-\pi) + \pi L^G(R^d, R^p)} = \frac{\pi}{(1-\pi)L^I(R^d, R^p) + \pi} \tag{6}$$

This formulation does have a family resemblance to the logit functions.

Now, suppose the likelihoods of guilt and innocence takes the following form:

$$L^G(R^d, R^p) = \frac{1}{L^I(R^d, R^p)} = \frac{F^P(R^P)}{F^D(R^D)} \tag{7}$$

It is simply the ratio of the quality of the evidence presented by the plaintiff and the defendant as ranked by the court.

Then, 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 function is described by (7), where $F^D(\cdot)$ and $F^P(\cdot)$ 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:*

$$p^G(R^d, R^p) = \frac{\pi F^P(R^P)}{(1-\pi)F^D(R^D) + \pi F^P(R^P)} \tag{8}$$

3.2 Resources, truth, and biases

How might the truth influence evidence production? 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 $R^D = R^P > 0$.[10]

[10] While 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

Since, the production process is structured in favor of the side with the truth, the court would have a reason to take the evidence produced from such a process seriously even if it actually knew the production functions of evidence. For the moment, let's suppose that this is indeed the case. In this instance, the court's inference from the evidence actually presented critically depends on whether it also observes the levels of spending put in by the two parties or not. In fact, if the court were to observe the levels of spending put in by the two sides, then it could decipher the truth perfectly: To see this clearly, consider the following more simplified process of evidence production:

When the defendant is innocent, let e^D and e^P be determined by the following functions:

$e^D = \varphi F(R^D)$
$e^P = (1 - \varphi)F(R^P)$
$\frac{1}{2} < \varphi < 1$

The function $F(.)$ 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 defendant, who is assumed to be innocent, would be able to supply a larger quantity of evidence if $R^D = R^P > 0$. Notice that the defendant also has a larger marginal benefit from putting in additional resources at this point.

When the defendant is guilty, let e^D and e^P be given by:

$e^D = (1 - \varphi)F(R^D)$
$e^P = \varphi F(R^P)$
$\frac{1}{2} < \varphi < 1$

Now, for any pair of evidence, (e^D, e^P) and the corresponding pair of resources (R^D, R^P), the judge could simply check which of the underlying evidence production

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, its 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. We hope to address such issues in the future.

process is consistent with this observation. This way the judge would be able to ascertain the guilt or the innocence of the defendant with perfect accuracy! In such a world, there would be literally no incentive left for the lying side to put in any resource towards presenting the evidence. The side with the truth would then only devote an epsilon level of effort and win the trial. However, the world we are familiar with is nowhere close to this prediction. Typically, both the defendant and the plaintiffs spend a lot of resources towards influencing the verdict. The scope for influencing is substantial in actual trials and other settings and this would be the case if the judge or the decision-maker were to only observe the evidence and not the resources spent. Further, when this is the case, the judge's knowledge of the underlying production functions of evidence would not be very useful in making an inference of guilt.[11] Let's suppose that the judge's likelihood ratio of guilt is determined in the same way as in (5). Then if the defendant were actually guilty, the likelihood ratio of guilt would be given by:

$$L^G(R^D, R^P) = \frac{1}{L^I(R^D, R^P)} = \frac{\varphi F(R^P)}{(1-\varphi)F(R^D)}$$

Hence, we would have:

$$p^G(R^d, R^p) = \frac{\pi\varphi F(R^P)}{(1-\pi)(1-\varphi)F(R^D) + \pi\varphi F(R^P)} \tag{9}$$

It can be clearly seen in this logit form how the probabilities of guilt (and, therefore, the probability of innocence) are 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; 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 that two possibilities and to a higher number of participants.

[11] While the assumption of the evidence production process being tilted in favor of the side with the truth would tend to justify inferring from the evidence presented, the fact that the process is still influenced by resources spent would tend to reduce its credibility. We assume that the former effect would dominate, so that the judge would consider the evidence presented seriously even if it were aware of the underlying production processes.

3.3 Deciding on guilt versus innocence

Ultimately, we are interested in deriving the contest success function as perceived by Defendant and Plaintiff at the beginning of stage 1. The posterior probability of guilt in (6) would be the appropriate contest success function if at stage 3 the court were to make a probabilistic decision based precisely on that posterior belief. In such a case, $p^G(R^d, R^p)$ would equal the $p^p(R^p, R^d)$, the probability of the Plaintiff winning the contest, and $1 - p^G(R^d, R^p)$ would be the probability of the Defendant winning. If, however, the court employs a discrete rule like "guilty beyond a shadow of a doubt" or just "choose guilt if and only if there is a better than even chance of guilt," then (6) does not in general describe the appropriate contest success function from the point of view of the contestants.

The court's decision rule can be described as follows:

$$\text{Choose Guilt if and only if } p^G(R^d, R^p) \geq \gamma \text{ where } \gamma \in (0,1) \tag{10}$$

Obviously γ should be sufficiently close to 1 when the "beyond-a-shadow-of-a-doubt" rule is employed and it equals $\frac{1}{2}$ when the even-chance rule is used.

If the Defendant and Plaintiff know the prior of the court with certainty, the it can be easily established that the contest success function is an asymmetric perfectly-discriminatory one (or, the all-pay auction). If the contestants have a nondegenerate distribution about the priors of the court (that is, the common prior of the contestants over the court's prior), then we can expect an imperfectly discriminating contest success function. In addition to our finding concerning the all-pay auction, the following Proposition describes the contest success function when the contestant's prior is uniform.

Proposition 2: *(a) Suppose the contestants know the court's prior with certainty and the court uses the decision rule (10). Then, the probability of the Plaintiff winning (and of the Defendant losing) is:*

$$
\begin{aligned}
p^p(R^p, R^d) &= 1 \text{ if } \frac{F^P(R^p)}{F^D(R^d)} \geq \frac{(1-\pi)\gamma}{\pi(1-\gamma)} \\
&= 0 \text{ if } \frac{F^P(R^p)}{F^D(R^d)} < \frac{(1-\pi)\gamma}{\pi(1-\gamma)}
\end{aligned}
\tag{11}
$$

(b) Suppose the contestants have a uniform prior about the court's prior and the court uses the decision rule (10). Then, the probability of the Plaintiff winning (and of the Defendant losing) is:

$$
p^p(R^p, R^d) = \frac{(1-\gamma)F^P(R^p)}{(1-\gamma)F^P(R^p) + \gamma F^D(R^d)} \tag{12}
$$

Proof: (a) Given the decision rule in (10) and the form of $p^G(R^d, R^p)$ in (6), the Plaintiff will win if and only if:

$$
\frac{\pi F^P(R^p)}{(1-\pi)F^D(R^d) + \pi F^P(R^p)} \geq \gamma
$$

or if and only if

$$
\frac{F^P(R^p)}{F^D(R^d)} \geq \frac{(1-\pi)\gamma}{\pi(1-\gamma)}
$$

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 (a).

(b) As in part (a), the decision rule in (10) and the form of $p^G(R^d, R^p)$ in (6) imply:

$$
\frac{\pi F^P(R^p)}{(1-\pi)F^D(R^d) + \pi F^P(R^p)} \geq \gamma
$$

or for any given prior π, we have

$$\pi \geq \frac{\gamma F^D(R^d)}{(1-\gamma)F^P(R^p) + \gamma F^D(R^d)}$$

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

$$\begin{aligned} Prob(\pi &> \frac{\gamma F^D(R^d)}{(1-\gamma)F^P(R^p) + \gamma F^D(R^d)}) \\ &= 1 - \frac{\gamma F^D(R^d)}{(1-\gamma)F^P(R^p) + \gamma F^D(R^d)} \\ &= \frac{(1-\gamma)F^P(R^p)}{(1-\gamma)F^P(R^p) + \gamma F^D(R^d)} \end{aligned}$$

which is the form in part (b) of the Proposition's statement. *QED*

The contest success function in (a) differs from typical applications of all-pay auctions (Hillman and Riley, 1989, Baye et. al., 1996, Che and Gale, 1998) only in that it asymmetric and also it does not include an outcome that has a probability of 1/2 when the probability of guilt just equals γ. (This latter difference, of course, could simply be eliminated by having the court flip a coin when that is the case. In such a case, especially when $\gamma = 1/2$, one could then ask why use the decision rule in (10) instead of just flipping an unfair coin with the probabilities in (6).)

In part (b) 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 bring evidence in its favor more easily than your opponent).

Thus, it is clear that both types of contest success functions in Proposition 2 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: Another functional form

In this section we examine an alternative evidence production process. Suppose each of the two interested parties, D and P, can produce one piece of evidence. The Defendant D can either produce evidence in favor of innocence, denoted by I, or offer no evidence, denoted by $\emptyset$. Similarly, the Plaintiff P can either produce evidence of guilt, denoted by G, or offer no evidence, $\emptyset$. However, the production of such evidence is not deterministic. The amount of resources simply enhance the probability of finding a favorable piece of evidence. Let $h^d(R^d)$ denote the probability that D will find evidence in favor of innocence (I). This probability is increasing in R^d, the resources expended on finding that evidence. Similarly, let $h^p(R^p)$ denote the probability that P will find evidence in favor of guilt (G), with that probability also increasing in the resources R^p expended by the plaintiff. These probabilities should of course be related to the true state of the world; that is, the probabilities should depend on whether D is truly guilty or innocent. We shall discuss this issue later. Thus, in terms of evidence there are four possible states of the world that can be faced by the court: $(I, G), (I, \emptyset), (\emptyset, G)$, and $(\emptyset, \emptyset)$ occurring with the following probabilities: $h^d(R^d)h^p(R^p)$, $h^d(R^d)[1 - h^p(R^p)]$, $[1 - h^d(R^d)]h^p(R^p)$, and $[1 - h^d(R^d)][1 - h^p(R^p)]$ respectively. Given evidence (i, g) (where $i \in \{I, \emptyset\}$ and $g \in \{G, \emptyset\}$), let $L^G(i, g)$ 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 I and G are to be potentially useful pieces of evidence to the two parties, then it must be the case that $L^G(i, g)$ observes the following properties:

$$L^G(\emptyset, G) > L^G(\emptyset, \emptyset) = 1 > L^G(I, \emptyset) \tag{13}$$

and

$$L^G(\emptyset, G) > L^G(I, G) > L^G(I, \emptyset), \tag{14}$$

whereas we cannot restrict *a priori* the relationship between $L^G(I, G)$ and $L^G(\emptyset, \emptyset)$.

In line with the arguments in the previous section, we posit that the court does not observe the level of resources and so its likelihood ratio of guilt is purely determined by the evidence pair it observes. Hence the posterior probability of guilt $\pi^*(i, g)$ for each pair of evidence (i, g) will be given by (5) using $L^G(i, g)$. From this it clearly follows that:

$$\pi^*(\emptyset, G) > \pi^*(\emptyset, \emptyset) > \pi^*(I, \emptyset) \tag{15}$$

$$\pi^*(\emptyset, G) > \pi^*(I, G) > \pi^*(I, \emptyset) \tag{16}$$

Given the posterior probability of guilt that will be induced by each realized combination of evidence and given the functions $h^d(\cdot)$ and $h^p(\cdot)$, the *ex ante* probability of guilt as a function of the resources expended by the two sides can be straightforwardly calculated.

Proposition 3: *With discrete evidence production, the probability of guilt is as follows:*

$$p^g(R^d, R^p) = \pi^*(\emptyset, \emptyset) + Bh^p(R^p) - \Delta h^d(R^d) + Ah^d(R^d)h^p(R^p) \tag{17}$$

where $\Delta \equiv \pi^*(\emptyset, \emptyset) - \pi^*(I, \emptyset) > 0, B \equiv \pi^*(\emptyset, G) - \pi^*(\emptyset, \emptyset) > 0$, and $A \equiv \pi^*(I, G) - \pi^*(\emptyset, G) + \pi^*(\emptyset, \emptyset) - \pi^*(I, \emptyset)$.

Proof: It follows straightforwardly from the following calculations:

$$\begin{aligned} p^g(R^d, R^p) &= h^d(R^d)[1 - h^p(R^p)]\pi^*(I, \emptyset) + h^d(R^d)h^p(R^p)\pi^*(I, G) + \\ & \quad [1 - h^d(R^d)]h^p(R^p)\pi^*(\emptyset, G) + [1 - h^d(R^d)][1 - h^p(R^p)]\pi^*(\emptyset, \emptyset) \\ &= \pi^*(\emptyset, \emptyset) + Bh^p(R^p) - \Delta h^d(R^d) + Ah^d(R^d)h^p(R^p) \end{aligned}$$

QED

Given the properties of the posterior probabilities, the probability of guilt can be shown to be increasing in the resources R^pexpended by the Plaintiff and decreasing in the resources R^d expended by the Defendant. Naturally, the probability of innocence, $1 - p^g(R^d, R^p)$, has the reverse properties in terms of the resources expended by each side. It can also be shown that $p^g(R^d, R^p)$ is increasing in π, the prior belief about guilt. The sign of the parameter A is indeterminate without additional restrictions on the posterior probabilities. It is positive, the higher is the prior belief about guilt and the more the likelihood ratios are tilted in favor of the plaintiff (and, therefore, more in favor of guilt).

In addition to the effect that resources have in uncovering evidence, the *truth* – whether the Defendant is guilty or innocent – should have an effect on the ease or difficulty with which each side can uncover evidence in favor of its cause. Thus, 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 D is guilty that when D is innocent. Similarly, for all R we can expect $h^d(R; G) < h^d(R; I)$, or that it would be more difficult for the Defendant to find evidence of innocence when he is guilty than when he is innocent. 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 D is guilty, and $\theta < 1/2$ when D is innocent. The closer θ is to 1 (when D is guilty) or the closer it is to 0 (when D is innocent), the more discerning of the truth can the technology of evidence production can be thought of.

The multiplicative term $Ah^d(R^d)h^p(R^p)$ in (17) can be eliminated only if $A = 0$. Such a condition holds under rather narrow set of circumstances since it is equivalent to $\pi^*(I, G) + \pi^*(\emptyset, \emptyset) = \pi^*(\emptyset, G) + \pi^*(I, \emptyset)$. The condition holds when, for example, $L^G(\emptyset, \emptyset) = L^G(I, G) = 1$, $L^G(\emptyset, G) = \Lambda$,and $L^G(I, \emptyset) = \frac{(2\pi-1)(\Lambda-1)+1}{2\pi(\Lambda-1)+1}$ for some $\Lambda > 1$. Even under these conditions, however, we would not have $B = \Delta$.

The function $p^g(R^d, R^p)$, however, is not the one that the contestants would generally face in stage 1 unless in stage 3 the court were to flip a biased coin in each state (i, g) (with the bias based on the posterior $\pi^*(i, g)$). There are ways to eliminate the multiplicative term, and do it so by taking into account stage 3 when the court uses the decision rule (10). Before discussing that possibility, though, we need to derive the operative contest success functions as of stage 1 using that rule. 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 we need to impose some additional restrictions on the likelihoods and the resultant posteriors in the four states. In particular, we suppose the following:

$$\begin{aligned} \pi^*(I, G) &= \pi^*(\emptyset, \emptyset) = \pi; \\ \pi^*(I, \emptyset) &= \delta\pi \text{ for some } \delta \in (0, 1); \\ \pi^*(\emptyset, G) &= \left\{ \begin{smallmatrix} \Gamma\pi \text{ if } \Gamma \leq 1/\pi \\ 1 \text{ if } \Gamma > 1/\pi \end{smallmatrix} \right. \text{ where } \Gamma > 1 \end{aligned} \tag{18}$$

Obviously the inequalities in (15) and (16) are satisfied since

$$\Gamma\pi = \pi^*(\emptyset, G) > \pi = \pi^*(I, G) = \pi^*(\emptyset, \emptyset) > \delta\pi = \pi^*(I, \emptyset$$

Proposition 4: *(a) Suppose the contestants know the court's prior with certainty and the court uses the decision rule of Guilty if and only if* $\pi^*(i, g) \geq \gamma$ for some $\gamma \in (0, 1)$ and all states (i, g). *In addition, suppose* $\pi^*(I, G) = \pi^*(\emptyset, \emptyset) = \pi$. *Then, the probability of the Plaintiff winning (and of the Defendant losing) is:*

$$
\begin{aligned}
p^p(R^p, R^d) &= 0 \text{ if } \gamma \in (\pi^*(\emptyset, G), 1] \\
&= [1 - h^d(R^d)]h^p(R^p) \text{ if } \gamma \in (\pi, \pi^*(\emptyset, G)] \\
&= 1 - h^d(R^d)[1 - h^p(R^p)] \text{ if } \gamma \in (\pi^*(I, \emptyset), \pi] \\
&= 1 \text{ if } \gamma \in [0, \pi^*(I, \emptyset)]
\end{aligned} \tag{19}
$$

(b) Suppose the contestants have a uniform prior about the court's prior and the court uses the decision rule of Guilty if and only if $\pi^*(i,g) \geq \gamma$ for some $\gamma \in (0,1)$ and all states (i,g)*. In addition, assume (18).[this is not an assumption but a parametrization.] Then, the probability of the Plaintiff winning (and of the Defendant losing) is:*

$$
p^p(R^p, R^d) = 1 - \gamma + \gamma[\frac{\Gamma - 1}{\Gamma} h^p(R^p) - \frac{1-\delta}{\delta} h^d(R^d) + (\frac{1-\delta}{\delta} - \frac{\Gamma - 1}{\Gamma}) h^p(R^p) h^d(R^d)] \tag{20}
$$

Proof: (a) There are four possible states (i,g) 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 $\pi^*(i,g)$ is greater or smaller than γ .

When $\gamma \in (\pi^*(\emptyset, G), 1]$, 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 $\gamma \in (\pi, \pi^*(\emptyset, G)]$, the court decides in favor of the Plaintiff only in state $(\emptyset, G)$. Thus, in this case we have $p^p(R^p, R^d) = [1-h^d(R^d)]h^p(R^p)1 = [1-h^d(R^d)]h^p(R^p)$.

When $\gamma \in (\pi^*(I, \emptyset), \pi]$, the court decides in favor of the Plaintiff in states $(\emptyset, G), (I, G)$, and $(\emptyset, \emptyset)$, and against the Plaintiff in state $(I, \emptyset)$, yielding $p^p(R^p, R^d) = 1 - h^d(R^d)[1 - h^p(R^p)]$.

Finally, when $\gamma \in [0, \pi^*(I, \emptyset)]$, the court decides in favor of the Plaintiff in all states and thus $p^p(R^p, R^d) = 1$.

(b) Given the parametrization of posterior probabilities of guilt in *(18)* and the decision rule, there are four intervals within which γ could fall: $(\pi\Gamma, 1], (\pi, \pi\Gamma], (\delta\pi, \pi]$, and $[0, \delta\pi]$. In those intervals the states in which the Plaintiff wins are the same as those described above in the proof of part (a) of this Proposition. Corresponding to the same intervals are the intervals of the prior π within which the same choices are made by the court:

$$\begin{aligned}
\text{For } \pi &\in [0, \frac{\gamma}{\Gamma}) \text{ the Plaintiff always loses} \\
\text{For } \pi &\in [\frac{\gamma}{\Gamma}, \gamma) \text{ the Plaintiff wins only in state } (\emptyset, G) \\
\text{For } \pi &\in [\gamma, \frac{\gamma}{\delta}) \text{ the Plaintiff wins in states } (\emptyset, G), (I, G), \text{ and } (\emptyset, \emptyset) \\
\text{For } \pi &\in [\frac{\gamma}{\delta}, 1] \text{ the Plaintiff always wins.}
\end{aligned}$$

Given that the contestants' prior about π is uniformly distributed, the Plaintiffs probability of winning can be calculated:

$$\begin{aligned}
p^p(R^p, R^d) &= 0(\frac{\gamma}{\Gamma} - 0) + [1 - h^d(R^d)]h^p(R^p)(\gamma - \frac{\gamma}{\Gamma}) \\
&\quad +1 - h^d(R^d)[1 - h^p(R^p)](\frac{\gamma}{\delta} - \gamma) + 1(1 - \frac{\gamma}{\delta}) \\
&= \gamma\{[1 - h^d(R^d)]h^p(R^p)(\frac{\Gamma - 1}{\Gamma}) + 1 - h^d(R^d)[1 - h^p(R^p)](\frac{1-\delta}{\delta}) - \frac{1}{\delta})\} + 1 \\
&= 1 - \gamma + \gamma[\frac{\Gamma - 1}{\Gamma}h^p(R^p) - \frac{1-\delta}{\delta}h^d(R^d) + (\frac{1-\delta}{\delta} - \frac{\Gamma - 1}{\Gamma})h^p(R^p)h^d(R^d)]
\end{aligned}$$

QED

The functional form in part (a) could be of interest, and be used, in settings where the cut-off parameter γ is close enough to the prior π so that the choices by the contestants have an effect. Perhaps of more applied interest is the form in part (b). In particular, when $\frac{1-\delta}{\delta} = \frac{\Gamma-1}{\Gamma}$, $A = 0$ and the probability of each winning depends on the difference between $h^p(R^p) - h^d(R^d)$. When $\gamma = 1/2$, as we might expect it to be in cases of product or political choice, the probability of P winning takes the following simple form:

$$p^p(R^p, R^d) = \frac{1}{2} + \frac{\alpha}{2}[h^p(R^p) - h^d(R^d)] \text{ where } \alpha \equiv \frac{1-\delta}{\delta} = \frac{\Gamma - 1}{\Gamma} < 1$$

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, given that we have imposed the symmetry condition that yields $A = 0$, no side has an advantage, although they both have a greater return to their resource investment the higher is α.

The functional forms in (17) and in Proposition 4 ((18) and (19)) can be in principle generalized for the case with more alternatives than two, since the "truth," especially in less clear-cut instances than those about guilt or innocence, can take one of a large number of alternatives. Furthermore, with more alternatives present, we could also in principle allow for more than two actors who would presumably try to promote their own favorite alternative. However, it can easily be seen that such generalizations would become increasingly unwieldy as more alternatives and actors were to be added. And, although the functional forms derived in this section appear to have good foundations and we would encourage its usage in future work, it has not been used anywhere that we are aware thus far.

5 Concluding remarks

In this paper, we characterize the process of persuasion: the attempt by contending parties to influence the opinion or judgement of an audience by devoting costly resources into evidence production. 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. One of the functional forms happens to be the relatively well-known logit function which has been recently used to model win-probabilities of contestants in lawsuits, election campaigns, and other similar settings. Our contribution is in providing an explicit inferential foundation for using this function in such contexts. We also uncover a new functional form which could also be suitably applied to such settings. We wish to stress that although the models described in the paper are motivated in the context of inference by the court, we have hardly imposed any court-specific structural assumptions. We believe that our models apply to various settings where eventually a judgement must be made on the basis of priors and evidence released by the competing parties.

When some of the audience has the same model as the contestants.

Suppose from the Contestants' perspective there is a probability σ that the audience is so sophisticated that it has exactly the same model of the world as Contestants as well as exactly the same knowledge as they have. (Equivalently, we assume a large audience with a proportion σ having such knowledge.) Such an audience would be able to completely infer guilt or innocence [would it? I have to check.]

References

[1] Baron, D.P., "Electoral Competition with Informed and Uninformed Voters," American Political Science Review, March 1994, 88(1), 33-47.

[2] Baye, M.R. and Hoppe, H.C., "The Strategic Equivalence of Rent-Seeking, Innovation, and Patent-Race Games," Games and Economic Behavior, 2003, 44(2), 217-226.

[3] Baye, M.R., Kovenock, D. and de Vries, C.G., "The all-pay auction with complete information," Economic Theory, 1996, 8, 291-305.

[4] Bell. D.E., Keeney, R.L. and Little, J.D.C., "A market share theorem," Journal of Marketing Research, 1975, 12, 136-141.

[5] Bernardo, A.E., Talley, E., and Welch, I., "A Theory of Legal Presumptions," Journal of Law, Economics, and Organization, 16(1), 1-49, 2000.

[6] Castillo, R.A. and Skaperdas, S., 2005. "All in the family or public? Law and appropriative costs as determinants of ownership structure," Economics of Governance, 6(2), 93-104.

[7] Che, Y.K and Gale, I., "Caps on political lobbying," American Economic Review, 1998, 88, 643-651.

[8] Cialdini, R.B., Influence: Science and Practice, Fourth Edition, 2001, Needham Heights, MA: Allyn & Bacon.

[9] Clark, D., Riis, C. "Contest success functions: an extension", Economic Theory, 11(1), 201-204, 1998.

[10] Dewatripoint, M., Tirole, J. "Advocates", Journal of Political Economy, 107, 1-39, 1999.

[11] Farmer, A., Pecorino, P. "Legal expenditure as a rent-seeking game", Public Choice, 100(3-4), 271-288, 1999.

[12] Froeb, L., Kobayashi, B. "Naive, biased, yet Bayesian: can juries interpret selectively produced evidence", Journal of Law, Economics and Organization 12, 257-71, 1996.

[13] Fullerton, R.L. and McAfee, R.P., "Auctioning Entry into Tournaments," Journal of Political Economy, 1999, 107(3), 573-605.

[14] Grossman, G., Helpman, E. "Protection for sale", *American Economic Review*, 84, 833-850, 1994.

[15] Hillman, A. and Riley, J.G., "Politically contestable rents and transfers," Economics and Politics, 1989, 1, 17-40.

[16] Hirshleifer, J. and Osborne, E. “Truth, effort and the legal battle”, Public Choice, 108(1-2), 169-195, 2001.

[17] Hirshleifer, J. and J. Riley (1992) The Analytics of Uncertainty and Information (Cambridge University Press, New York).

[18] Jia, H. (2005) A Stochastic Derivation of Contest Success Functions, University of California, Irvine, mimeo.

[19] Kadane, J., Schum, D. A Probabilistic Analysis of the Sacco and Vanzetti Evidence, John Wiley & Sons, Inc., New York, 1996.

[20] Konrad, K.A., "Bidding in hierarchies," European Economic Review, 2004, 48(6),1301-1308.

[21] Konrad, K.A., "Strategy in Contests," manusript, February 2005.

[22] Laband, D., Sophocleus, P. “An estimate of resource expenditures on transfer activity in the United States”, Quarterly Journal of Economics, 107, 959-983, 1992.

[23] Lam, R. “Learning through stories and other essays”, Ph.D. Dissertation, Yale University, 2001.

[24] Legros, P., Newman, F. “Courts, contracts, and interference”, European Economic Review, 46, 734-744, 2002.

[25] Lippmann, W., Public Opinion, 1997(1922), New York: Free Press.

[26] Luce, R.D., 1959, Individual Choice Behavior New York: Wiley.

[27] Milgrom, P., Roberts, R. “Relying on the information of interested parties”, Rand Journal of Economics, 17, 18-32, 1986.

[28] Mueller, Holger and Warneryd, Karl, ”Inside Versus Outside Ownership: A Political Theory of the Firm,” RAND Journal of Economics, Fall 2001, 32(3), 527-541.

[29] Nitzan, S., "Modeling rent-seeking contests," European Journal of Political Economy, 1994, 10(1), 41-60.

[30] Selvin, M. and Picus, L. The Debate over Jury Performance: Observations from a Recent Asbestos Case, Rand Corporation, Santa Monica, 1987.

[31] Shin, H. "Adversarial and inquisitorial procedures in arbitration", Rand Journal of Economics, 29, 378-405, 1998.

[32] Schmalensee, R. (1978) The Economics of Advertising, (North-Holland, Amsterdam-London)

[33] Skaperdas, S. (1996), Contest success functions, Economic Theory 7:283-90.

[34] Skaperdas, S and Grofman, B.G. "Modeling Negative Campaigning," American Political Science Review, 1995, 89(1), 49-61.

[35] Sobel, J. "Disclosure of evidence and resolution of disputes: who should bear the burden of proof?" in Alvin E. Roth, ed., Game Theoretic Models of Bargaining, Cambridge University Press, Cambridge, 1985.

[36] Tullock, G. 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.

[37] Tullock, G. The Economics of Special Privilege and Rent Seeking, Kluwer Academic Publishers, Boston, 1989.

[38] Vaidya, S., (2005a). "Corruption in the Media's Gaze", European Journal of Political Economy, Volume 21, Issue 3, pages 667-687.

[39] Vaidya, S. (2005b), "The Nature of Corruption Deterrence in a Competitive Media Sector", forthcoming, Economics of Goverance.