



סדרת מאמרים לדיון Working paper series

No. 1 – 2011

Subgame Perfect Plea Bargaining in Biform Judicial Contests

Mordechai E. Schwarz

Subgame Perfect Plea Bargaining in Biform Judicial Contests

Mordechai E. Schwarz

*The Open University of Israel and the Open University Research Institute for Policy,
Political Economy and Society (OURI-PPES)*

February 14, 2011

Abstract

The judicial process is modeled as a biform contest with risk-averse contestants. I provide a sufficient condition for a non-empty core and show that the effect of the severity of the charges on the core is ambiguous. I also show that the practice of plea bargaining actually applies Moulin's (1984) mechanism which non-cooperatively implements the Kalai-Smorodinsky (1975) bargain solution in subgame perfect equilibrium. The charge reduction rate is inversely related to the defendant's and positively related to the prosecutor's "fear of ruin" index, but increases with the defendant's stake, implying that the defendant's gain (and social cost) from the guilty plea deal increases with the severity of the crime.

Keywords: plea bargaining, subgame-perfect equilibrium, Kalai-Smorodinsky solution, implementation.

Jel Code: K41, C78, D86.

1. Introduction

Plea bargaining involves negotiations between a defendant and a prosecutor regarding the conditions under which the defendant enters a guilty plea. It is estimated that over 95% of the convictions in criminal cases in the United States result from a negotiated guilty plea,¹ but the rate varies with the seriousness of the crime.²

From a positivist point of view, as Carney and Fuller (1969) pointed out, "[I]t would seem reasonable to expect that the weaker the case against a defendant, the more entitled is he to a trial by jury, and, correspondingly, the more questionable is the propriety of the negotiated plea. In practice, however, it seems that the weaker the case (up to a point), the more likely is the prosecutor to press for a guilty plea, inasmuch as this was the most often mentioned reason given by prosecutors for trying a negotiated plea." Hence, the positivist puzzle of plea bargaining is what drives defendants to plead guilty when evidence is weak, and prosecutors to ask for lenient sentences in return, instead of declining the case.

Apparently, charge reduction may explain the defendants' tendency to plead guilty,³ so the question that remains is what drives a prosecutor to reduce charges in return for a guilty plea. If evidence is insufficient to prove the original charges but strong enough to prove the less serious charges, one would expect the prosecutor to indict for the reduced charges from the start, without making "concessions" to the defendant, which may be perceived by the public as weakness of the law enforcement system. On the other hand, if the evidence is too weak even for the lesser charge, what drives the defendant to plead guilty, even for reduced charges?

Moreover, one would expect that once a prosecutor agrees to reduce charges against a defendant in return for a guilty plea, he would also insist that this defendant receive the maximal sentence for this less serious offence if he pleads guilty. According to the Federal Sentencing Guidelines (FSG), a guilty plea is not recognized

¹ Grossman and Katz (1983), Franzoni (1999), Wright (2005), Ulmer et al. (2009).

² Plea bargains constituted 98% of convictions for embezzlement, immigration and gambling, about 80% for kidnapping and assault but only 67% of those for murder (Reiganum, 2000).

³ Carney and Fuller (1969) reported that 67.8% of their sample entered pleas of guilty and 93.2% of these guilty pleas included charge reduction. Wright and Engen (2006, 2007) found that charge reduction occurs in roughly half of all felony cases that resulted in conviction, and has a large effect on sentencing. See also Wright and Miller (2002).

as relevant for sentencing (Ulmer et al., 2009). Nevertheless, Wright (2005), King et al. (2005) and Ulmer et al. (2009) found significant differences between sentences imposed on trial convicted defendants and those defendants who pleaded guilty for the same offense, even after controlling for legitimate factors, implying a trial penalty for insisting on the constitutional right to be judged. Namely, defendants who plead guilty enjoy charge reduction and additional sentence discounts.

Classical economic studies of law and economics (c.f. Becker, 1968) treated the judicial procedure as a lottery, assuming that conviction probability is exogenously determined by evidence. More recent studies shifted to the contest model,⁴ but the lottery model remained popular in plea bargaining analyses.⁵ The few exceptions either lack an explicit analysis of the judicial process as a contest, assume that litigants are risk-neutral and the prosecutor is sincerely motivated.⁶ These studies investigated, for example the effect of expected trial outcome on litigants' expenditures (c.f. Katz, 1988), or vice versa (c.f. Skaperdas and Gan, 1995), but I am unaware of any study which explicitly analyzes how all these variables are endogenously interrelated. To this end, a *biform contest* model is required.

In this article, I analyze the judicial procedure as a biform contest, namely a two-stage game. The first stage is a rent-seeking contest which results in a second stage bargaining game. I provide a sufficient condition for a non-empty core, show that the effect of the severity of the crime on agreement availability is ambiguous and then show that the common plea bargaining practice of charge reduction actually applies Moulin's (1984) mechanism which non-cooperatively implements the Kalai-Smorodinsky (1975) bargain solution in subgame-perfect equilibrium. The charge reduction rate is inversely related to the defendant's and positively related to the prosecutor's "fear of ruin" index. Charge reduction increases with the defendant's stake, implying that the defendant's gain (and social cost) from a guilty plea deal increases with the severity of the crime. It is also shown that risk-aversion is a necessary condition for guilty-plea deals.

⁴ c.f. Katz (1988), Hirshleifer and Osborne (2001). For a survey of earlier literature, see Cooter and Rubinfeld (1989).

⁵ c.f. Newman (1966), Adelstein (1978), Grossman and Katz (1983), Bebchuk (1984), Lott (1987), Reinganum (1988), Kobayashi (1992), Reinganum (2000), Baker and Mazzetti (2001), Kim (2010).

⁶ c.f. Landes (1971), Lott (1987), Franzoni (1999), Yilankaya (2002), Hay (1995).

The article proceeds as follows. Section 2 contains a brief review of related literature. Section 3 presents the benchmark one-shot contest model equilibrium. Section 4 presents the extended biform game model of judicial process with risk-averse rent-seeking contestants. Section 5 discusses the implications of risk-seeking and section 6 summarizes the article. All proofs are relegated to the appendix.

2. Related Literature

As mentioned above, classical economic studies of law and economics assumed that conviction probability is exogenously given, according to evidence strength, and although more recent studies shifted to the contest model, plea bargaining analyses continued to assume the lottery model. Also, economic analyses of plea bargaining usually assume risk-neutral litigants and a sincere prosecutor, or lack an explicit analysis of the judicial process as a contest.

Nevertheless, the controversy among philosophers and jurists regarding the moral and judicial aspects of plea bargaining indicates that scholars have the contest model in mind.⁷ Proponents of negotiated justice claim that the high costs associated with the production of evidence justifies plea bargaining which "saves the taxpayer money," especially when the defendant is not forced to plead guilty. The plea bargain, they argue, replicates the trial outside the courtroom and enables the parties to estimate their chances. Thus, the plea agreement reflects the expected verdict more efficiently.⁸ Clearly, this argumentation makes sense only if "evidence" is assumed to be a function of "efforts" exerted by the litigants, as assumed by the contest model.

The sincere prosecutor assumption is plausible within the lottery model framework because the exogenously fixed conviction probability neutralizes the prosecutor's incentives and motivation effect. The contest model assumes that "evidence," and thus winning probabilities, are determined endogenously according to

⁷ See, for example, Kipnis (1976), Huff, Rattner and Sagarin (1986), Scott and Stunz (1992), Schulhofer (1992), Gazal and Bar-Gill (2004) and Wright (2005).

⁸ Ulmer et al. (2009) report that 34% of the federal judges they interviewed agreed or strongly agreed that an "efficient case is an end in itself." No data were supplied about the prosecutors' attitude, but it is hard to believe that prosecutors who bear most of the burden of evidence production would be less generous to defendants for saving them time and effort.

efforts and resources invested by the contestants, implying that the sincere prosecutor assumption is implausible. Kipnis (1976) noted that prosecutors operate under public or administrative pressure to "produce" convictions, and are judged professionally by their conviction productivity. Miceli (1990) noted that the adversarial system itself naturally places the prosecutor and the defense attorney on opposite sides of an all-or-nothing contest. Consequently, each side has an incentive to present its most favorable case to the court, but the heavy caseload facing most prosecutors necessitates that the majority of cases be settled out of court through plea bargains. Therefore, to encourage such plea bargains, and to insure that the prosecutor is in a strong bargaining position, it is important that he establish a record of consistent success at trial, especially when the prosecutor's salary, prestige and promotion or reelection are presumably linked to his conviction record as a measure of his success. Reinganum (2000) noted that an elected prosecutor may maximize his chances of reelection by being "tough on crime" on one hand, and "saving the taxpayers' money" on the other. Plea bargaining serves both these apparently contradictory goals. For example, Wright (2005) found that as plea bargains became more common, the conviction rate increased while acquittals and dismissals declined sharply.⁹

These studies and the above mentioned empirical findings of charge reduction and sentencing discounts in return for "saving taxpayers' money" or "trial penalty" for wasting public resources by insisting on the constitutional right to a trial, reflect a principal-agent problem between the public (the principal) and the prosecutor (the agent). Principal-agent situations are characterized by an informational advantage of the agent over the principal. The prosecutor's informational advantage explains the almost automatic approval of plea arrangements by courts, as plea negotiations between the prosecution and the defense include fact bargaining and charge reduction, implying that the court and the public are not exposed to "bad facts." Guilty pleas are mutually beneficial to the prosecutor and the defendant, but not necessarily to society.

⁹ This finding also indicates that plea arrangements did not reflect expected verdicts but replaced inquiries and investigations as the main source of evidence. Wright (2005), Ulmer et al. (2009), Bowers (2007) and other researchers documented some commonly used techniques applied by prosecutors in order to extract guilty pleas from defendants, which cannot be reconciled with sincere truth seeking. These findings also undermine the "voluntary" and "knowing" basis of the defendant's guilty plea. Similarly, assuming sincere judges and jurors ignores their incentive to save their own time and efforts, and to encourage litigants to end the process as quick as possible. The strategic bias of judges and jurors has been studied extensively in the literature. See, for example, Austin-Smith and Banks (1996), Feddersen and Pesendorfer (1998), Dekel and Piccione (2000) and Coughlan (2000).

The prosecution's ability to obtain judicial approval for socially detrimental plea agreements stems from the public's unawareness of these "bad facts." Hence, the deadweight loss associated with the future consequences of lenient sentences and short terms of imprisonment for ruthless criminals is underestimated by the public. Future victims have no faces, and their cries are presently silent.

3. The Model

3.1. Basic Settings

Consider two contestants, 1 (a defendant) and 2 (a prosecutor), competing in a judicial rent-seeking type contest. Denote the contestants' expenditures vector by $\mathbf{x} = \{x_1, x_2\}$ and assume that the winning probability of agent i is given by a Tullock (1980) type Contest Success Function,

$$(1) \quad p_i(\mathbf{x}) = \frac{dh_i(x_i)}{h_i(x_i) + dh_j(x_j)}, \quad i, j = 1, 2$$

Where $h_i(x_i)$ is agent i 's production function, which measures the effectiveness of agent i 's effort in producing evidence, and d is a positive parameter which measures agent i 's advantage (or disadvantage) over agent j . In a judicial context, d reflects the standard of proof. Namely, if $d_1 > 1$, the prosecutor bears a heavier standard of proof. For the sake of simplicity, assume $h_i(x_i) = x_i$ and $d = 1$. This simplifying assumption does not affect the qualitative results of the model.¹⁰

The judicial contest has two contingent results: acquittal (I) or conviction (II). In case of conviction, the defendant loses v_1 and the prosecutor gains v_2 . Denote the

¹⁰ For axiomatization of Contest Success Functions, see Skaperdas (1996), Clark and Riis (1998), Epstein and Nitzan (2007) and Corhón and Dahm (2010).

initial wealth of the contestant i by A_i and his post contest net wealth by w_i^k , $i = 1, 2$, $k = I, II$. It follows that,

$$(2) \quad \begin{aligned} w_1^I &= A_1 - x_1, & w_1^{II} &= A_1 - x_1 - v_1 \\ w_2^I &= A_2 - x_2, & w_2^{II} &= A_2 - x_2 + v_2 \end{aligned}$$

3.2. Attitude toward Risk

Assuming risk aversion significantly complicates the analysis.¹¹ Therefore, the rent-seeking literature generally assumes risk-neutrality.¹² Nevertheless, in a bargaining, and particularly in a plea bargaining context, this assumption is not only extremely unrealistic, but misses an important behavioral factor (see Svejnar, 1986).

The most prevalent risk-aversion indices are Pratt's (1964) absolute risk-aversion index, defined as $R_i = -u_i''/u_i'$, Arrow's (1971) relative risk-aversion index, defined as $AR_i = R_i w_i$, and Aumann and Kurz's (1977) "fear of ruin" index, defined as $F_i \equiv u_i/u_i'$.¹³ According to common definitions that prevail in the literature, an agent i is defined as risk-neutral if $R_i = 0$, risk-averse if $R_i > 0$ and risk-seeker if $R_i < 0$. F_i is a measure of agent i 's "fear of ruin" or "fear of disagreement," in Svejnar's (1986) term, which is more relevant in our context. Suppose that an agent faces a gamble in which he risks his entire wealth, w_i in return for an additional small gain of v_i . As v_i is small relative to w_i , the agent's loss probability, q_i , must be very small in order to make him gamble his entire wealth, and even smaller, as the agent's "fear of ruin" is larger. Thus, the "fear of ruin" is inversely related to "boldness." As v_i is smaller, q_i approaches zero. The "boldness" index is therefore, $\lim_{v_i \rightarrow 0} q_i/v_i$. To see that notice that the agent indifference condition is

¹¹ Hillman and Katz (1984), Skaperdas and Gan (1995), Konrad and Schlesinger (1997) and Cornes and Hartley (2003).

¹² See Epstein and Nitzan (2007).

¹³ On the interrelations between these indices of risk-aversion, see Foncel and Treich (2005).

$$(3) \quad u_i(w_i) = (1 - q_i)u_i(w_i + v_i) + q_i u_i(0)$$

Rearranging (3), assuming $u_i(0) = 0$ yields:

$$(4) \quad \frac{q_i}{v_i} = \frac{[u_i(w_i + v_i) - u_i(w_i)]/v_i}{u_i(w_i + v_i)}.$$

Taking the limit of (4) yields $\lim_{v_i \rightarrow 0} q_i/v_i = u_i'/u_i = F_i$. Notice that even if $R_i = 0$, namely when an agent is risk-neutral according to common definitions, he still may have positive "fear of disagreement."

3.3. The One-Shot Judicial Contest Competitive Equilibrium

As a benchmark, let us begin by analyzing the judicial contest as a one-shot common rent-seeking contest. Assume that agents' preferences are represented by von-Neumann-Morgenstern utility functions $u_i(w_i)$, $u_i' > 0$, $u_i'' < 0$. The agents' target functions are

$$(5) \quad Eu_i(w_i) = p_1(\mathbf{x})u_i(w_i^I) + p_2(\mathbf{x})u_i(w_i^H), \quad i = 1, 2$$

First order conditions for interior solution is

$$(6) \quad p_i'^*(\mathbf{x}^*)\Delta u_i - Eu_i' = 0 \quad \forall i$$

Where asterisks denote equilibrium values, $p'_i = \frac{\partial p_i}{\partial x_i}$, $\Delta u_i = u_i(w_i^I) - u_i(w_i^{II})$ and $Eu'_i = p_i^*(\mathbf{x}^*)u'_i(w_i^I) + (1 - p_i^*(\mathbf{x}^*))u'_i(w_i^{II})$. With risk-neutral contestants, (6) is reduced to $p'_i(\mathbf{x}^*)v_i = 1 \quad \forall i$, implying

$$(7) \quad \hat{\mathbf{x}} = \left\{ \frac{v_1^2 v_2}{(v_1 + v_2)^2}, \frac{v_1 v_2^2}{(v_1 + v_2)^2} \right\}, \quad \hat{\mathbf{p}} = \left\{ \frac{v_1}{v_1 + v_2}, \frac{v_2}{v_1 + v_2} \right\}$$

Where $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ denote equilibrium vectors under a risk-neutrality assumption.

4. Biform Games and Biform Contests

Brandenburger and Stuart (2007) described a *biform n -player game* (or *hybrid game*) as a two-stage game. The first stage is non-cooperative and is designed to describe the strategic moves of the players. However, the consequences of these moves are not payoffs. Instead, each profile of strategic choices at the first stage leads to a second stage cooperative game. This gives the competitive environment created by the choices that the players made in the first stage. Formally, a biform game is a collection $(S, V, \mathbf{a})$, where $S = S^1 \times S^2 \times \dots \times S^n$ is a profile of strategies, $V : A \times 2^n \rightarrow \mathbb{R}$ is a cooperative game coalitional function and $\mathbf{a} = \{\alpha_i\}_{i=1}^n$ is a vector of "confidence indices" for each player ($\alpha_i \in [0, 1]$). Roughly speaking, the confidence indices evaluate every contingent outcome in the core according to player i 's preferences.¹⁴

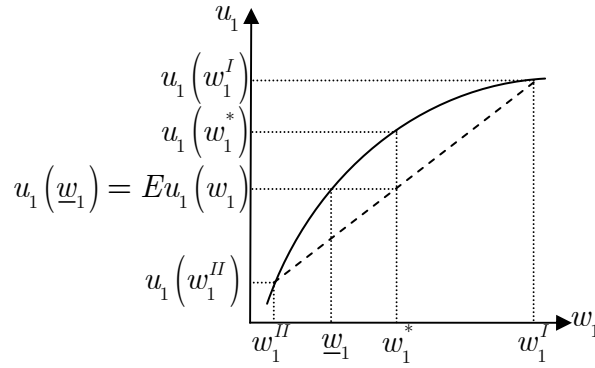
Similarly, a *biform n -player contest* is a two-stage game. The first stage is a rent-seeking contest designed to determine the bargaining power of the contestants. The

¹⁴ For a more accurate definition and explanation of confidence index, see Brandenburger and Stuart (2007) or Stuart (2005).

contest's outcome is not prize allocation, but a vector of equilibrium expenditures and winning probabilities that characterizes each contestant's bargaining power. Formally, a biform contest is a collection $(\mathbf{x}, p, \mathbf{u}, b)$ where $\mathbf{x} = \{x_i\}_{i=1}^n$ is the contestants rent-seeking expenditures vector, p is the contest success function, $\mathbf{u}$ is a vector of contestants' utilities and b is a bargaining problem.¹⁵

The benchmark analysis in the previous section modeled the judicial contest as a one shot zero-sum non-cooperative game, in which the expected equilibrium net wealth of the defendant and the prosecutor are $w_1^* = A_1 - p_2^* v_1 - x_1^*$ and $w_2^* = A_2 + p_2^* v_2 - x_2^*$, respectively. Assuming risk-aversion and applying Jensen's inequality implies that $u_i(w_i^*) \geq Eu_i(w_i)$. Hence, there is a $\underline{w}_i < w_i^*$, satisfying $u_i(\underline{w}_i) = Eu_i(w_i)$, (see Figure 1). $\underline{w}_i$ is the *certainty equivalent* wealth of agent i .

Figure 1



A plea bargain aims at avoiding contest confrontation through a compromise that assigns the defendant a charge reduction of $f(v_1)$ in return for pleading guilty. Skaperdas and Gan (1995) analyzed a compromise in a contest problem, emphasizing that "An agreed division of the prize, though, does not imply the absence of effort on

¹⁵

For early analyses of biform contest situations, see McDonald and Solow (1981), Svejnar (1986), Alexander (1992) and Skaperdas and Gan (1995). More recent studies include, for example, Eesteban and Sákovics (2002), Bayindir-Upmann and Gerber (2003) and Skaperdas (2006).

the contestants' part. Each contestant would like to establish a good bargaining position and the best way to do that is to have a credible fallback position in case negotiations fail." Skaperdas and Gan's study is thus an early analysis of a biform contest, where in the first stage contestants exert efforts to establish bargaining power and a credible negotiation threat in case of fallback, but also to be well equipped for the conflict, if necessary. If an agreement is achieved, the defendant's and the prosecutor's net wealth are $w_1 = A_1 - p_2 f(v_1) - x_1$ and $w_2 = A_2 + p_2 f(v_2) - x_2$, respectively. A compromise agreement is achievable if it is mutually beneficial, namely, if it ensures each contestant $u_i(w_i) \geq u_i(\underline{w}_i)$.¹⁶

4.1. The Core of a Plea Bargaining Biform Contest

For simplicity, assume $f(v_i) = \beta v_i$. A subgame perfect equilibrium is calculated by backward induction. Hence, assume that $\beta \in [0, 1]$ is the second stage bargaining solution and is common knowledge. Thus, the defendant's target function is

$$(8) \quad \pi_1(\mathbf{x}) = u_1(A_1 - p_2 \beta v_1 - x_1),$$

and the prosecutor's target function is

$$(9) \quad \pi_2(\mathbf{x}) = u_2(A_2 + p_2 \beta v_2 - x_2),$$

¹⁶ Skaperdas and Gan (1995) assumed that the sharing rule, $f(p_2)$, refers to the winning probability of agent 2. Here I assume that the sharing rule refers to v_i . This formulation has a straightforward interpretation. $p_2 v_1$ is the penalty discount for the defendant which reflects its expected value, and $f(v_i)$ is the rate of charge reduction.

where p_2 is the prosecutor's winning probability. Clearly, the defendant seeks to minimize p_2 and the prosecutor seeks to maximize it. Solving the two optimization problems simultaneously yields:

$$(10) \quad \tilde{\mathbf{x}} = \beta \hat{\mathbf{x}} \Rightarrow \tilde{\mathbf{p}} = \hat{\mathbf{p}}$$

Where $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ are defined above in (7). Namely, as pointed out by Skaperdas and Gan (1995), bargaining position optimization implies that contestants' expenditures are proportional to risk-neutral agents' expenditures, where β is the proportionality coefficient. Denote the expected payoff of the defendant and the prosecutor by $c_1 = p_2 v_1 + x_1$ and $c_2 = p_2 v_2 - x_2$, respectively, and it is easily obtained from (10):

$$(11) \quad \tilde{c}_1 = \beta \hat{c}_1, \quad \tilde{c}_2 = \beta \hat{c}_2.$$

The derivation of (11) was based on the assumption that in the first stage, the sharing rule, β , is exogenous and common knowledge. In real world plea bargaining negotiations, however, β lies in the center of the dispute between the parties. Therefore, we first have to define the upper and lower limits of β . In other words, we now turn to explore the core of the plea bargain game.

A sharing rule β (namely, a deal), is *privately rational* if it ensures each agent at least his certainty equivalent utility. A deal is *rational* if it is privately rational for *all* agents. The *core*, or the *bargaining set* (denoted by B), is the set of all rational deals.

Formally, define an agents' *gain function* by

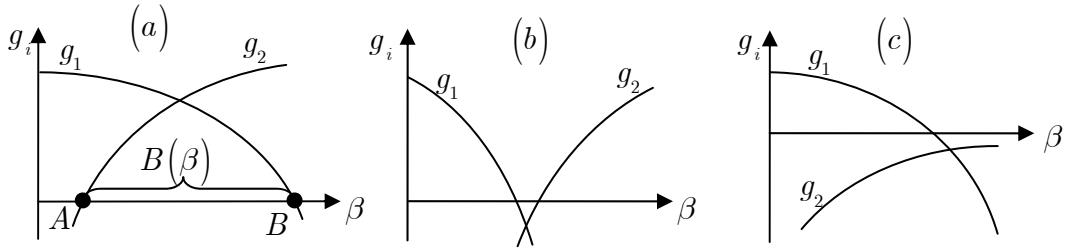
$$(12) \quad g_i(\beta) = \pi_i(\beta, \hat{\mathbf{x}}) - Eu_i(w_i, \mathbf{x}^*).$$

It can easily be verified that $\frac{\partial g_1}{\partial \beta} < 0$, $\frac{\partial g_2}{\partial \beta} > 0$, $\frac{\partial^2 g_1}{\partial \beta^2} < 0$ and $\frac{\partial^2 g_2}{\partial \beta^2} > 0$, implying that the g_1 curve is downward sloping and the g_2 curve is upward sloping (see Figure 2). The core is defined as

$$(13) \quad B(\beta) = \{\beta \mid g_i(\beta) \geq 0, \forall i\}.$$

In panel (a) of Figure 2, the core is represented by the segment AB . The other two panels of Figure 2 show examples of empty cores.

Figure 2



Define $\underline{\beta} = \inf \{\beta \in B(\beta)\}$ (point A in Figure 2), and $\bar{\beta} = \sup \{\beta \in B(\beta)\}$ (point B in Figure 2).

Proposition 1

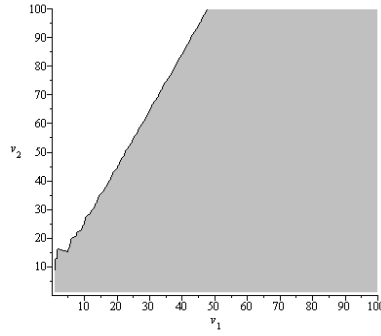
$$a. \quad \underline{\beta} = \frac{c_2^*}{\hat{c}_2}, \quad \bar{\beta} = \frac{c_1^*}{\hat{c}_1}.$$

$$b. \quad B(\beta) \neq \emptyset \Leftrightarrow (v_2 - v_1) \frac{\partial p_2}{\partial F} \Big|_{F=0} \leq \frac{\partial X}{\partial F} \Big|_{F=0}.$$

Proof: See appendix.

Proposition 1 implies that the core may be empty even when both agents are characterized with concave utility functions. The model is analytically solvable only if $u_i(w_i) = -e^{-w_i}$. Figure 3 presents combinations of v_1 and v_2 for a non-empty core simulated for this utility function specification.¹⁷

Figure 3



4.2. Comparative Statics of the Core

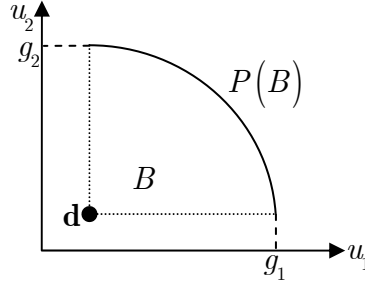
Although a non-empty core may contain an infinite number of rational deals, our attention is focused on Pareto-efficient deals. Denote by $P(B)$ the collection of all Pareto efficient deals in B . The above mentioned characteristics of g_i imply that

$\frac{\partial g_2}{\partial g_1} < 0$ and $\frac{\partial^2 g_2}{\partial g_1^2} > 0$. Namely, the efficiency frontier curve, $P(B)$, is downward

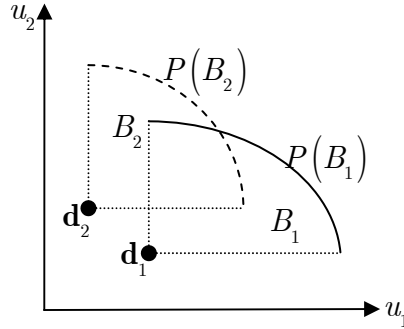
sloping on the utilities plane (u_1, u_2) , and the bargaining set, $B(\beta)$, is convex and compact, as demonstrated in Figure 4. (d in Figure 4 denotes the disagreement point).

¹⁷

Notice that for this utility function $R_i = 1$ and $F_i = -1$.

Figure 4

In a plea bargaining context, disagreement means going to trial, hence $\mathbf{d} = (Eu_1, Eu_2)$. Namely, both B and $\mathbf{d}$ are endogenous, implying, as many authors have indicated, that comparative statics of the core in a biform contest framework is complicated.¹⁸ For instance, an increase in the prosecutor's first-stage efforts shifts the disagreement point from $\mathbf{d}_1$ to $\mathbf{d}_2$ and, consequently, the bargain set from B_1 to B_2 (see Figure 5).

Figure 5

For example, consider the effect of "fear of disagreement" on the core. Differentiating $\bar{\beta}$ and $\underline{\beta}$ with respect to $F_i(A_i)$ and manipulating algebraically yields:

¹⁸

For example, see McDonald and Solow (1981), Alexander (1992), Anbarci, Skaperdas and Syropoulos (2002), Bayindir-Upmann and Gerber (2003) and Skaperdas (2006).

Corollary 1

$$\begin{aligned}
(14) \quad (a) \quad & \frac{\partial \bar{\beta}}{\partial F_1(A_1)} \geq 0 \Leftrightarrow -\frac{\eta_{p_2, F_1}^*}{\eta_{x_1, F_1}^*} \leq \frac{x_1^*}{p_2^* v_1} \\
(b) \quad & \frac{\partial \beta}{\partial F_2(A_2)} \geq 0 \Leftrightarrow \frac{\eta_{p_2, F_2}^*}{\eta_{x_1, F_2}^*} \geq \frac{x_2^*}{p_2^* v_2}
\end{aligned}$$

where $\eta_{a,b}$ denotes the elasticity of a with respect to b . It requires some tedious algebra to show that $\text{sgn}\left(\frac{\partial \bar{\beta}}{\partial v_i}\right)$, and $\text{sgn}\left(\frac{\partial \beta}{\partial v_i}\right)$ are indeterminate. Assuming that v_i is positively correlated with the severity of the charge leads to the following corollary.

Corollary 2

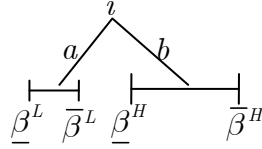
The effect of the severity of the charges on agreement availability is ambiguous.

4.3. The Confidence Index

The main drawback of the core is that, on one hand, it may be empty, and on the other, it may contain an infinite number of solutions. In the latter case, a confidence index is useful. Consider two different cores that result from two different strategic profiles, as demonstrated in Figure 6. Denote the upper and lower limits of these cores by $[\underline{\beta}^H, \bar{\beta}^H]$ and $[\bar{\beta}^L, \bar{\beta}^L]$, respectively. The confidence index is used to evaluate the weighted average utility of each player for a given core. Roughly speaking, the confidence index indicates how well player i anticipates doing in the resulting

cooperative games. For example, in Figure 6, player j will chose strategy a if $\alpha^i u_i(\underline{\beta}^L) + (1 - \alpha^i) u_i(\bar{\beta}^L) > \alpha^i u_i(\underline{\beta}^H) + (1 - \alpha^i) u_i(\bar{\beta}^H)$.

Figure 6



In a plea bargaining context, a confidence index is less useful. First, the confidence level is defined in terms of the critical level for preferring strategy a over strategy b (see Figure 6), but, from (14) and Corollary 2, we know that the effect of the "fear of ruin" index on $\underline{\beta}$ and $\bar{\beta}$ is ambiguous. The confidence index is probably correlated with the "fear of ruin" index, and this correlation should be taken into account. Secondly, equation (10) implies that equilibrium expenditures and β are interrelated. Namely, expenditures depend on the anticipated specific value of β , while the boundaries of the β values depend on first stage equilibrium expenditures. The analysis of these interrelations requires a unique and precise bargain solution.

Nash (1950) suggested a unique bargain solution which satisfies four basic axioms: Invariance to affine transformation (IAT), Efficiency or Pareto Optimality (PO), Symmetry (S) and Independence of Irrelevant Alternatives (IIA). The IIA axiom was criticized and replaced with the monotony axiom (M) by Kalai and Smorodinsky (1975), who suggested an alternative unique bargain solution that satisfies IAT, PO, S and M.

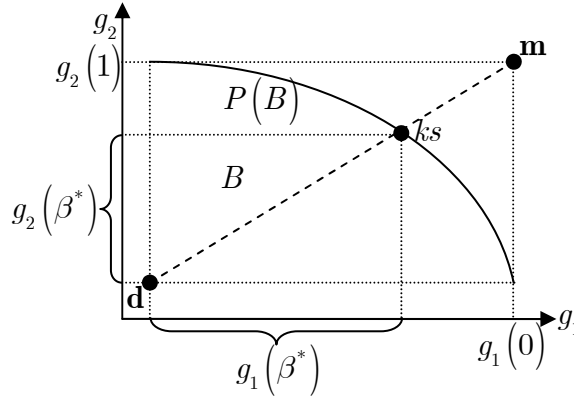
4.4. The Kalai-Smorodinsky Plea Bargaining Solution

Define $m_1 = g_1|_{\beta=\underline{\beta}}$ and $m_2 = g_2|_{\beta=\bar{\beta}}$ as the maximum feasible utilities for agents 1 and 2, respectively. $\mathbf{m} = (m_1, m_2)$ is known as the *utopia point*. The Kalai-Smorodinsky (henceforth KS) plea bargaining solution is the β which solves:

$$(15) \quad \frac{g_i(\beta(v_i))}{m_i} = \frac{g_j(\beta(v_j))}{m_j} \quad \forall i, j$$

Namely, the KS solution equalizes the proportional gains of every agent, relative to his utopian gain.¹⁹ In Figure 7, the KS solution is the intersection point of the $\mathbf{dm}$ line and the efficiency frontier.

Figure 7



Consider the following Plea Bargaining-Protocol (PBP):

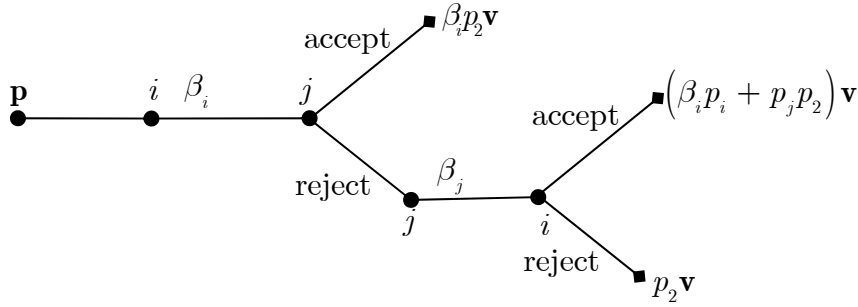
Round 0: The two parties invest x_1 and x_2 in order to establish a bargaining position.

¹⁹ Alternately, the KS solution is defined by $ks_i = \max_{\beta \in [\underline{\beta}, \bar{\beta}]} \min_i (g_i(\beta)/m_i)$.

Round 1: Suppose that $p_i > p_j$. Contestant i makes a proposal of β_i to j . If j accepts the proposal, it is implemented by the court.²⁰ If j rejects the proposal, proceed to Round 2.

Round 2: Contestant j makes a counter-proposal, β_j . If i accepts, the court implements the deal with probability p_i and the status quo with probability $1 - p_i$. If i rejects the counter-proposal, the status-quo point is implemented with certainty.²¹ (See Figure 8).

Figure 8



This protocol is a modification, for a competitive environment, of the static mechanism suggested by Moulin (1984). Notice that after Round 0, contestants' expenditures are sunk-costs, implying that this mechanism pegs the bargaining set and sets $d_i = \hat{p}_1 u_i(w_i(I)) + \hat{p}_2 u_i(w_i(II))$. Hence, for convenience, let us normalize $d = 0$. Of course, legally $\beta \in (0,1)$, thus $m_1 = \pi_1|_{\beta=0}$, and $m_2 = \pi_2|_{\beta=1}$, implying that the KS plea bargaining solution is $\beta \in (0,1)$, which solves:

²⁰ Legally, the court is not obliged to approve any plea agreement submitted by the litigants. However, since plea negotiations include fact bargaining, as explained above, the court is usually not exposed to "bad facts." Hence, the assumption that a "reasonable" plea agreement submitted by both sides will always be approved by the court is plausible.

²¹ The rationale of Round 2 stems from the assumption of "almost" automatic approval of guilty plea agreements in Round 1, as explained in note 20. Namely, as the bargaining process continues, flowing into Round 2, the public and the court are exposed to more "bad facts." Therefore, in Round 2 the approval probability is only p_i .

$$(16) \quad \frac{\pi_1(\beta)}{m_1} = \frac{\pi_2(\beta)}{m_2}.$$

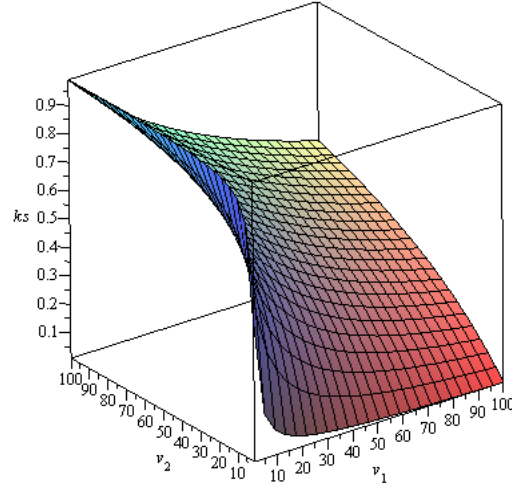
Proposition 2

The Plea Bargaining Protocol (PBP) non-cooperatively implements the Kalai Smorodinsky (1975) bargaining solution in subgame-perfect equilibrium.

Proof: See appendix.

Intuitively, this mechanism solves the prosecutor's credibility problem (Franzoni, 1999), since in case of disagreement the outcome is imposed by the court. Assuming $u_i(w_i) = -e^{-w_i}$ yields $ks = \hat{p}_2$. Figure 9 presents a simulation of ks for this function.

Figure 9



4.5. Comparative Statics of the Sharing Rule

Non-cooperative subgame-perfect implementation of a bargaining solution, $f(v_i)$, implies that $\tilde{x}$ and $\tilde{p}$ are given by (10), thus the PBP mechanism pegs the bargaining set, simplifies comparative static analysis and reduces ambiguity.

By common wisdom, an increase in F_1 increases the defendant's tendency to accept worse deals, and an increase in F_2 increases the prosecutor's generosity toward the defendant. This intuition is true. The Taylor expansion of (16) is

$$(17) \quad 1 - \frac{u'_1(A_1)}{u_1(A_1)} \beta \hat{c}_1 = 1 + \frac{u'_2(A_2)}{u_2(A_2)} (\beta - 1) \hat{c}_2.$$

Solving (17) for β and rearranging yields:

$$(18) \quad ks = \frac{\hat{c}_2 F_1}{\hat{c}_1 F_2 + \hat{c}_2 F_1}$$

Differentiating (18) yields:

$$(19) \quad \begin{aligned} \frac{\partial ks}{\partial F_1} &= \frac{\hat{c}_1 \hat{c}_2 F_2}{(\hat{c}_1 F_2 + \hat{c}_2 F)^2} > 0 \\ \frac{\partial ks}{\partial F_2} &= -\frac{\hat{c}_1 \hat{c}_2 F_1}{(\hat{c}_1 F_2 + \hat{c}_2 F)^2} < 0 \end{aligned}$$

And,

$$(20) \quad \begin{aligned} \frac{\partial ks}{\partial v_1} &= -\frac{v_2^2(4v_1 + v_2)F_1F_2}{4(v_1F_2(2v_1 + v_2) + v_2^2F_1)^2} < 0 \\ \frac{\partial ks}{\partial v_2} &= \frac{v_1v_2(4v_1 + v_2)F_1F_2}{4(v_1F_2(2v_1 + v_2) + v_2^2F_1)^2} > 0 \end{aligned}$$

Define the defendant's "discount function" by $G_1 = \hat{p}_2v_1(1 - ks)$. This function measures the expected trial penalty imposed on a defendant who insists on his constitutional right to a court trial. Similarly, define the prosecutor's "concession function" by $G_2 = \hat{p}_2v_2(1 - ks)$. By plugging $\hat{c}_1$, $\hat{c}_2$, $\hat{p}_2$ and ks into G_1 and G_2 , we obtain:

$$(21) \quad \begin{aligned} G_1 &= \frac{v_1^2v_2F_2(2v_1 + v_2)}{(v_1 + v_2)[F_2v_1(2v_1 + v_2) + F_1v_2^2]} \\ G_2 &= -\frac{v_1v_2^2F_2(2v_1 + v_2)}{(v_1 + v_2)[F_2v_1(2v_1 + v_2) + F_1v_2^2]} \end{aligned}$$

It can be easily verified that

$$(22) \quad \begin{aligned} \frac{\partial G_1}{\partial F_1} &< 0, \quad \frac{\partial G_1}{\partial F_2} > 0 \\ \frac{\partial G_1}{\partial v_1} &> 0, \quad \frac{\partial G_1}{\partial v_2} \geq 0 \Leftrightarrow \frac{F_1}{F_2} \leq \frac{2v_1^2(2v_1 + v_2)^2}{v_2^2(v_2^2 + 2v_1(v_2 + v_1))} \end{aligned}$$

However,

$$(23) \quad \frac{\partial G_2}{\partial v_2} \geq 0 \Leftrightarrow 0 \leq \frac{F_2}{F_1} \leq \frac{v_2^3}{6v_1v_2(2v_1 + v_2) + v_2^3 + 8v_1^3}$$

$$\frac{\partial G_2}{\partial v_1} \geq 0 \Leftrightarrow \frac{F_2}{F_1} \leq \frac{v_2^2(v_2^2 + 4v_2v_1 + 2v_1^2)}{v_1^2(2v_1 + v_2)^2}$$

Figure 10 presents our simulations for G_1 and G_2 , respectively, assuming $u_i = -e^{-w_i}$.

Figure 10

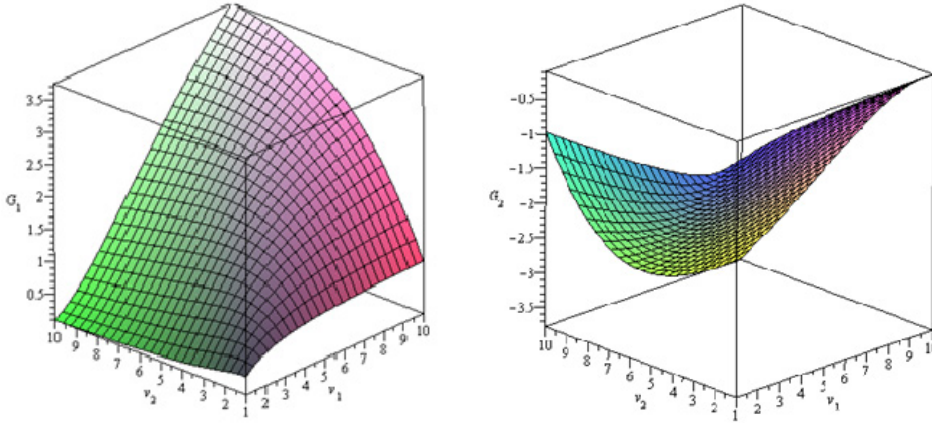


Figure 10 is compatible with Ulmer and Bradley (2006), who indicated that "In fact, under an organizational efficiency explanation, by far the most common in the sentencing literature, one might expect greater trial penalties among the relatively less serious cases here like robbery, because these might be seen as wasting the courts' time with unsuccessful trials when stakes (potential sentence severity) are lower compared to more severe cases. We find the opposite pattern – trial penalties increase with offense severity, at least for incarceration."

Define the total social cost of plea bargaining by $G = G_1 + G_2$. It follows that:

$$(24) \quad G = \frac{v_1 v_2 (v_1 - v_2) F_2 (2v_1 + v_2)}{(v_1 + v_2) [F_2 v_1 (2v_1 + v_2) + F_1 v_2^2]}.$$

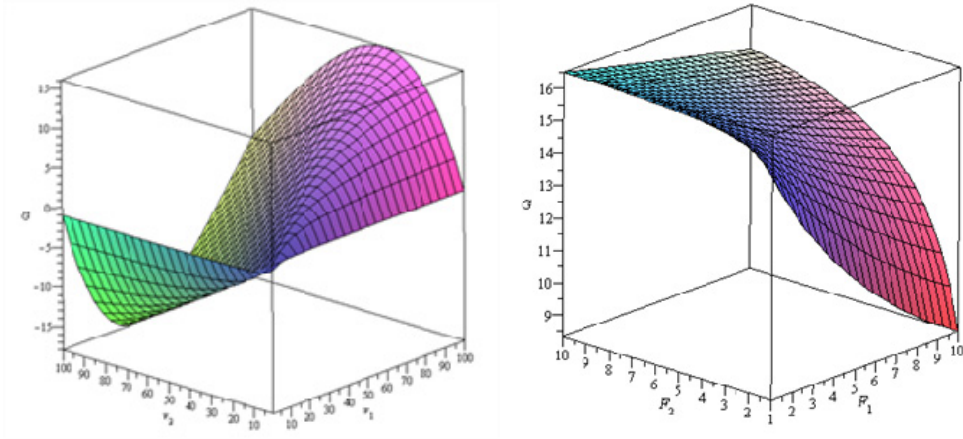
Corollary 3

$$(25) \quad G \geq 0 \Leftrightarrow v_1 \geq v_2,$$

$$\frac{\partial G}{\partial F_i} \geq 0 \Leftrightarrow v_i \leq v_j$$

Nevertheless, $\text{sgn}\left(\frac{\partial G}{\partial v_1}\right)$ and $\text{sgn}\left(\frac{\partial G}{\partial v_2}\right)$ are indeterminate. Figure 11 presents our simulations for G . (The left panel is calibrated by $F_1 = F_2 = 1$ and the right panel is calibrated by $v_1 = 100$ and $v_2 = 50$).

Figure 11



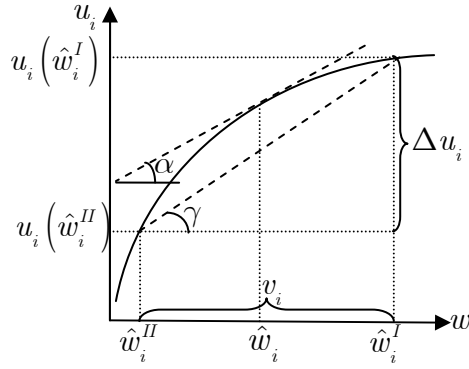
4.6. The Plea Bargaining Effect on Total Expenditures

Generally, the effect of risk-aversion on contestants' expenditures is ambiguous.²²

Konrad and Schlesinger (1997) show that $x_i^* \gtrless \hat{x}_i \Leftrightarrow Eu_i'(\hat{\mathbf{x}}) \gtrless \frac{\Delta u_i(\hat{\mathbf{x}})}{v_i}$. In terms of

Figure 12, $x_i^* \gtrless \hat{x}_i \Leftrightarrow \tan(\alpha) \gtrless \tan(\gamma)$.

Figure 12



However, from (10) we obtain $\mathbf{x}^* \gtrless \tilde{\mathbf{x}} \Leftrightarrow \mathbf{x}^* \gtrless \beta \hat{\mathbf{x}}$. Hence, the Konrad and Schlesinger condition is insufficient in a competitive environment.

Corollary 4

The effect of guilty plea deals on total expenditures is ambiguous.

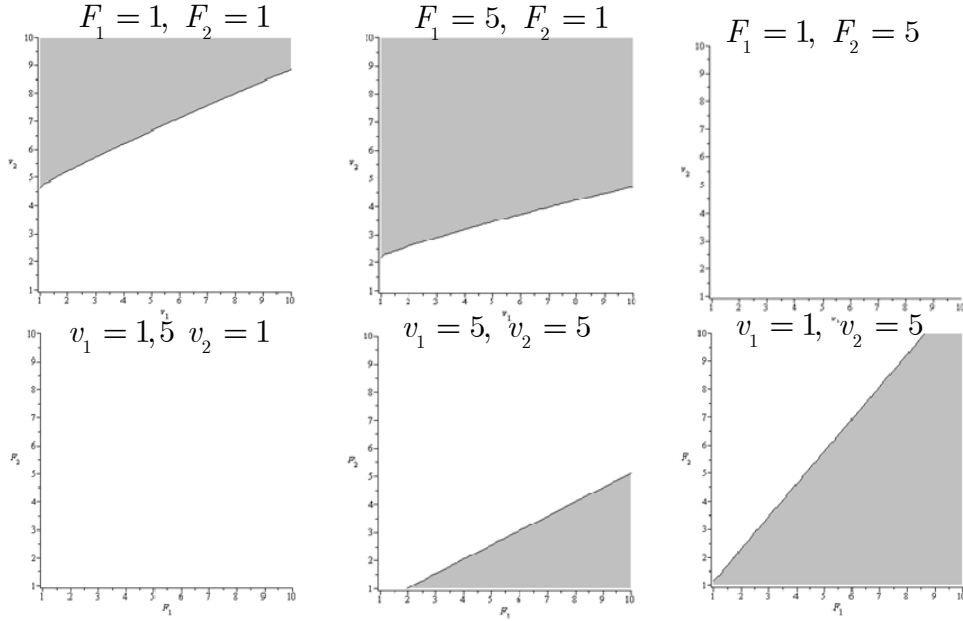
Define the total expenditures function by $X = x_1 + x_2$, and denote the plea bargaining effect on expenditures by $\Delta X = \tilde{X} - X^*$. (Recall that $\tilde{X} = ks\hat{X}$). The

²²

Millner and Pratt (1991). Hillman and Katz (1984) show that for "small" prizes, risk-aversion reduces rent-dissipation. (See also Hillman and Samet, 1987 and Nitzan, 1994). But as Konrad and Schlesinger (1997) indicated, generalizing this result is non-trivial (see also Cornes and Hartley, 2003).

upper row of Figure 13 contains combinations of v_1 and v_2 , satisfying $\Delta X \geq 0$ for various values of F_1 and F_2 , while the diagrams in the lower row of Figure 13 contain combinations of F_1 and F_2 , satisfying $\Delta X \geq 0$ for various values of v_1 and v_2 (simulated for $u_i(w_i) = -e^{-w_i}$).

Figure 13



The ambiguity of plea bargaining effect on expenditures undermines the common "public resources savings" justification for the plea bargaining practice.²³ Wright (2005) found that "a lighter workload for judges over the years produced fewer guilty pleas and more acquittals, but lighter workloads for prosecutors over time led to just the opposite result: more guilty pleas and fewer acquittals." Wright noted that this fact is "counterintuitive" and indicated that "When judges had more time to devote to criminal cases, the guilty plea rate tended to fall and the acquittal rate tended to rise: the judicial workload fell along with the guilty plea rate from the 1950's through the

²³

For example, Justice Burger indicated: "If every criminal charge were subjected to a full-scale trial, the States and the Federal Government would need to multiply by many times the number of judges and court facilities." *Santobello v. New-York* 404 U.S. 257, 261 (1971).

1970's. The gentle rise in the judicial criminal caseload since 1981 corresponded with the large increase in the acquittal rate" (see also Ulmer et al., 2009).²⁴

5. Risk-Seeking Agents

The analysis in the previous section used the "fear of ruin" index as the risk-aversion measure. As noted above, F_i may be positive even when $R_i = 0$, because this function measures "all or nothing" risk aversion, but not limited loss risk aversion. In particular, the F_i function does not distinguish between agents with concave or convex utility functions (risk-aversers and risk-seekers by common definitions). Hence, for the analysis of risk-seeking effect on the core, we use Pratt's (1964) absolute risk-aversion index, defined as $R_i = -u''_i/u'_i$.

Proposition 3

$$B(\beta) \neq \emptyset \Leftrightarrow R_i \geq 0, \forall i.$$

Proof: See appendix.

According to Kahneman and Tversky's (1979) "prospect theory," individuals classify outcomes relative to a certain *reference point* (for instance, the *status quo*). Outcomes that yield a higher utility level than the reference point are classified as *gains* while outcomes which yield lower utility are classified as *losses*. The expected utility of the individual is given by a *value function* such as $U(\bar{w}) = \sum_{l=1}^L \phi(p_l)u(w_l)$, where w_l is a contingent outcome and $\phi(p_l)$ is the *probability weighting function*,

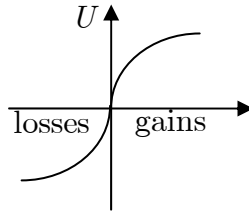
²⁴

For case studies of the effect of the abolition of plea bargaining, see, for example, Rubinstein and White (1979), Call, England and Talarico (1983) and Weininger (1987).

which is a non-linear transformation of the outcome's objective probability p_i that overestimates small probabilities and underestimates large probabilities.

The value function does not measure *wealth* but *changes* in wealth, namely *gains* and *losses*. Hence, the reference point is normalized to the origin, the reference point is normalized to the origin, and the value function is S-shaped around it (see Figure 14), implying that individuals are not *risk-averse* but *loss-averse*, as the negative impact of losses is greater than the positive impact of identical (in absolute values) gains. In other words, according to prospect theory, when outcomes are perceived as losses, individuals tend to behave as risk-seekers.²⁵

Figure 14



A defendant who wins in a trial is left with net wealth of $A_1 - x_1$, and a defendant who loses is left with $A_1 - x_1 - v_1$. Namely, the two contingencies are net losses, implying that according to prospect theory, defendants are expected to behave as risk-seekers and reject any plea deal proposal. Proposition 3 also predicts that if agents are risk-seekers, the core is empty. Apparently, both predictions are incompatible with the above mentioned data about frequent plea bargaining deals.

One may argue that in terms of Figure 1, given the high rate of convictions, the defendants' reference point is w_1^H . However, this interpretation is incompatible with the bulk of evidence on defendants' behavior. On one hand, some researchers argue that defendants are subject to *innocence bias*, which causes them to over-estimate their winning probabilities and reject *any* guilty plea offers (c.f. Tor, Gazal-Ayal and

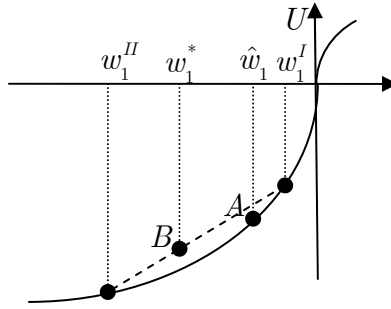
²⁵

Kahneman and Tversky based their theory on experimental evidence, but their methodology evoked a polemic. See, for example, Gigerenzer (1991) and Kahneman and Tversky's (1996) reply. However, this polemic is beyond the scope of this article.

Garcia, 2006). Innocence bias implies that in Figure 1, the defendants' reference point is to the right of w_1^I . On the other hand, reports of the Innocence-Project organization revealed cases of innocent defendants falsely convicted after pleading guilty,²⁶ implying that when the innocence bias is not binding, the defendants' reference point is, indeed, w_1^{II} .

Motti Michaeli²⁷ suggested that our results can be reconciled with Prospect Theory because through charge reduction, the expected utility and the utility from expected wealth no longer lie on the same plumb. (In Figure 15, point A represents higher utility than point B). Nevertheless, the prevalence of guilty plea deals indicates that this explanation should be empirically tested and the Prospect Theory should be theoretically reconsidered and empirically retested.

Figure 15



6. Summary and Discussion

In this article, I analyzed the judicial process as a biform contest, provided conditions for a non-empty core, and showed that the common plea bargain practice applies Moulin's (1984) mechanism which non-cooperatively implements the Kalai-Smorodinsky (1975) bargain solution in subgame-perfect equilibrium.

The PBP mechanism pegs the bargaining set and thus enables comparative static analysis of a compromise in a competitive environment. As expected, k_s is positively

²⁶ These reports are available at <http://www.innocenceproject.org/>.

²⁷ A Ph.D. student at the Center for the Study of Rationality (Hebrew University of Jerusalem), under the supervision of Robert J. Aumann.

correlated with the defendant's "fear of disagreement", and negatively correlated with the prosecutor's "fear of disagreement". However, k_s is negatively correlated with the defendant's stake and positively correlated with prosecutor's stake. If stakes monotonically increase with crime severity, the defendant's gain and social costs exhibit the same pattern. Contrary to the common justification for the plea bargaining practice, the effect of guilty pleas on expenditures is ambiguous.

Although the "fear of disagreement" effect on the core is ambiguous, risk-aversion is a necessary condition for a non-empty core. Apparently, this theoretical result and the common rate of trials that result in guilty pleas are incompatible with Kahneman and Tversky's (1979) Prospect Theory. Although Michaeli's suggestion for reconciling my results with Prospect Theory is nice, I hold that further research is required as this explanation should be empirically tested, and Prospect Theory should be both theoretically reconsidered and empirically retested.

The simplified analysis above reflects the common wisdom about criminal justice in countries where negotiated justice prevails: "criminal codes do not matter much." More important than the codes' provisions is the prosecutor's application of the criminal statutes. Wright and Engen (2006) examined charge reductions in North Carolina and found that charge reductions were common, "occurring in roughly half of all felony cases that resulted in convictions, and that the prosecutor's decision to reduce criminal charges has a large effect on average sentence severity." But they also found that "these effects do not apply equally, however, to all crimes. When a group of related crimes offer deeper charging options (that is, the number of charges that might apply to a given set of facts), the prosecution and defense agree more often to reduce the charges. Large distances between the sentences that attach to available charging options make it less likely that the prosecution and defense will agree on a particular charge reduction. Thus, plea bargaining is not an entirely free-market exercise that allows the parties to negotiate a customized outcome. Even in a world of charge-driven sentencing where prosecutorial discretion is a dominant feature, the substantive criminal law matters."²⁸

Following the above analysis, I suggest a reverse causality interpretation of Wright and Engen's findings. It is not the number of substitute charging options that

²⁸ See also Wright and Engen (2007).

cause charge reduction in guilty plea deals, but the law enforcement system's interest in encouraging these judicial deals. Without deciding whether innocence bias matters and to what extent, it seems plausible to assume that both over-optimism and the mental price of a false guilty plea increases with the offence severity, implying that charge reduction weakens the innocence bias effect and encourages even innocent defendants to plead guilty.²⁹ I conjecture that as the severity of the charges increases, deeper charge reduction is required by the KS solution, leading the law enforcement system to press legislators to supply more charging options for certain felony crimes.³⁰

The positivist view of the judicial procedure as a rent-seeking contest and the state prosecutor as a rent-seeker is, indeed, troubling. Normatively, justice should be equal for the poor and the rich. State prosecutors should not behave as contractors of convictions but seek the truth honestly and prosecute sincerely. Conviction should be based on reliable evidence beyond doubt and on due procedure. However, "*It is not the mouse that is the thief but the hole*" (Babylonian Talmud, tractate Gittin 45a). Namely, the built-in conflict of interests of the prosecutor's position, due to its principal-agent characteristics creates an incentive for reduced efforts. The plea bargain practice is the "hole" that enables the prosecutor to escape his duties as a public agent.

7. Appendix

Proof of Proposition 1

If no deal is achieved, the agents turn to judicial competition in the court. Their expected utilities are

²⁹ Of course, the defendant's innocence is his private information; hence, the bargaining outcome is inefficient as guilty defendants are also expected to enjoy charge reductions.

³⁰ Bower (2007) found that in petty cases, charge reduction is relatively rare and the plea bargain is centered mainly on the sentence.

$$\begin{aligned}
(A1) \quad Eu_1(w_1) &= p_1^* u_1(A_1 - x_1^*) + p_2^* u_1(A_1 - m_1) \\
Eu_2(w_2) &= p_1^* u_2(A_2 - x_2^*) + p_2^* u_2(A_2 + m_2)
\end{aligned}$$

Where $m_1 = (v_1 + x_1^*)$ and $m_2 = (v_2 - x_2^*)$. On the other hand, if the two agents agree on a deal, $\beta(v_i)$, their expected utilities are

$$\begin{aligned}
(A2) \quad \pi_1(\beta, \mathbf{x}) &= u_1(A_1 - \beta \hat{c}_1) \\
\pi_2(\beta, \mathbf{x}) &= u_2(A_2 + \beta \hat{c}_2)
\end{aligned}$$

where $\hat{c}_1 = \hat{p}_2 v_1 + \hat{x}_1$, $\hat{c}_2 = \hat{p}_2 v_2 - \hat{x}_2$, and $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ are defined in (7). Taking the Taylor expansion of (A1) yields:

$$\begin{aligned}
(A3) \quad Eu_1(w_1) &= p_1^* [u_1(A_1) - x_1^* u_1'(A_1)] + p_2^* [u_1(A_1) - m_1^* u_1'(A_1)] + \mathcal{R}_1 \\
Eu_2(w_2) &= p_1^* [u_2(A_2) - x_2^* u_2'(A_2)] + p_2^* [u_2(A_2) + m_2^* u_2'(A_2)] + \mathcal{R}_2
\end{aligned}$$

Where $\mathcal{R}_i$ is the remainder of the Taylor sequence. Similarly, the Taylor expansion of (A2) yields:

$$\begin{aligned}
(A4) \quad \pi_1(\beta, \mathbf{x}) &= u_1(A_1) - \beta \hat{c}_1 u_1'(A_1) + \mathcal{R}_1 \\
\pi_2(\beta, \mathbf{x}) &= u_2(A_2) + \beta \hat{c}_2 u_2'(A_2) + \mathcal{R}_2
\end{aligned}$$

The gain of agent i from a deal is defined by $g_i = \pi_i(\beta, \mathbf{x}) - Eu_i(w_i)$. The core is non-empty if $g_i \geq 0 \quad \forall i$. Thus, combining (A3) and (A4), dividing by $u'_i(A_i)$ and rearranging yields the following conditions for a non-empty core:

$$(A5) \quad \begin{aligned} \bar{g}_1 &= (1 - p_1^* - p_2^*)F_1(A_1) - \beta\hat{c}_1 + p_1^*\hat{x}_1 + p_2^*m_1 \geq 0 \\ \bar{g}_2 &= (1 - p_1^* - p_2^*)F_2(A_2) - \beta\hat{c}_2 + p_1^*\hat{x}_2 + p_2^*m_2 \geq 0 \end{aligned}$$

where $\bar{g}_i = \frac{g_i}{u'_i(A_i)}$. Recall that by definition $1 - p_1^* - p_2^* = 0$, and it follows that

$$(A6) \quad \underline{\beta} = \frac{c_2^*}{\hat{c}_2}, \quad \bar{\beta} = \frac{c_1^*}{\hat{c}_1} \quad \textbf{QED(a)}.$$

By inserting c_i^* and $\hat{c}_i$ into (A6), we easily obtain

$$(A7) \quad \underline{\beta} = \frac{p_2^*v_2 - x_2^*}{\hat{p}_2^2v_1}, \quad \bar{\beta} = \frac{p_2^*v_1 + x_1^*}{\hat{p}_2v_1(2 - \hat{p}_2)}.$$

Define $\Delta\beta = \bar{\beta} - \underline{\beta}$. Non-emptiness of the core implies $\Delta\beta \geq 0$. Hence, by (A7),

$$(A8) \quad \Delta\beta \geq 0 \Leftrightarrow \frac{\hat{p}_2v_2 - \hat{x}_2}{\hat{p}_2v_1 + \hat{x}_1} \geq \frac{p_2^*v_2 - x_2^*}{p_2^*v_1 + x_1^*}.$$

The left hand side of the inequality in (A8) is constant. In other words, (A8) implies that the "fear of ruin" affects c_1^* (the denominator on the right side of (A8)) more than c_2^* (the numerator). Suppose that the "fear of ruin" increases for both agents at the same rate, namely, $\partial F_1 = \partial F_2 = \partial F$. It follows that

$$(A9) \quad \left(\frac{\partial p_2}{\partial F} v_2 - \frac{\partial x_2}{\partial F} \right)_{F=0} \leq \left(\frac{\partial p_2}{\partial F} v_1 + \frac{\partial x_1}{\partial F} \right)_{F=0}$$

Rearranging (A9) and inserting $\frac{\partial X}{\partial F} = \frac{\partial x_1}{\partial F} + \frac{\partial x_2}{\partial F}$ yields:

$$(A10) \quad \left(v_2 - v_1 \right) \frac{\partial p_2}{\partial F} \Big|_{F=0} \leq \frac{\partial X}{\partial F} \Big|_{F=0} \quad \mathbf{QED(b)}.$$

Proof of Proposition 2:

A subgame perfect equilibrium is computed by backward induction. Suppose without loss of generality that $p_1 > p_2$, so in Round 1, agent 1 makes an offer, $\underline{\beta}$, to agent 2. By rejecting agent 1's proposal and proposing $\bar{\beta}$, agent 2's expected utility is

$$(A11) \quad Eu_2 \geq \hat{p}_1 u_2(\bar{\beta}) + (1 - \hat{p}_1) u_2(d_2)$$

Going back to Round 1, the best response of agent 1 is to offer β , satisfying

$$(A12) \quad u_2(\beta) = \hat{p}_1 u_2(\bar{\beta}) + (1 - \hat{p}_1) u_2(d_2)$$

By rearranging, (A12) can be rewritten as

$$(A13) \quad \frac{u_2(\beta) - u_2(d_2)}{u_2(\bar{\beta}) - u_2(d_2)} = \hat{p}_1$$

Now suppose that $p_2 > p_1$, thus in Round 1 agent 2 makes an offer to agent 1. Similar argumentation leads to the conclusion that agent 2's best response strategy is to propose β satisfying

$$(A14) \quad u_1(\beta) = \hat{p}_1 u_1(\underline{\beta}) + (1 - \hat{p}_1) u_1(d_1),$$

which by rearranging terms yields

$$(A15) \quad \frac{u_1(\beta) - u_1(d_1)}{u_1(\underline{\beta}) - u_1(d_1)} = \hat{p}_1.$$

Combining (A14) and (A15) yields

$$(A16) \quad \frac{u_1(\beta) - u_1(d_1)}{u_1(\underline{\beta}) - u_1(d_1)} = \frac{u_2(\beta) - u_2(d_2)}{u_2(\bar{\beta}) - u_2(d_2)}.$$

By definition, the $\beta^*(v_i)$, which solves (A16), is the Kalai-Smorodinsky bargaining solution. Going back to the first stage we obtain the subgame perfect equilibrium expenditures from (10). ***QED.***

Proof of Proposition 3:

If no deal is achieved and the agents turn to judicial competition in the court, their expected utilities are given by (A1), where $m_1 = (v_1 + x_1^*)$ and $m_2 = (v_2 - x_2^*)$. On the other hand, if the two agents agree on a deal, $\beta(v_i)$, their expected utilities are given by (A2), where $\hat{c}_1 = \hat{p}_2 v_1 + \hat{x}_1$ and $\hat{c}_2 = \hat{p}_2 v_2 - \hat{x}_2$. Taking the Taylor expansion of (A1) and rearranging yields:

$$\begin{aligned} Eu_1(w_1) &= u_1 - p_1^* \left[x_1^* u_1' - \frac{1}{2} x_1^{*2} u_1''(A_1) \right] - p_2^* \left[m_1^* u_1' - \frac{1}{2} m_1^{*2} u_1'' \right] + \mathcal{R}_1 \\ (A17) \quad Eu_2(w_2) &= u_2 - p_1^* \left[x_2^* u_2' - \frac{1}{2} x_2^{*2} u_2'' \right] + p_2^* \left[m_2^* + \frac{1}{2} m_2^{*2} u_2'' \right] + \mathcal{R}_2 \end{aligned}$$

Where u_i , u_i' and u_i'' are evaluated at A_i . Similarly, the Taylor expansion of (A2) yields:

$$\begin{aligned} \pi_1(\beta, \mathbf{x}) &= u_1 - \beta \hat{c}_1 u_1' + \frac{(\beta \hat{c}_1)^2}{2} u_1'' + \mathcal{R}_1 \\ (A18) \quad \pi_2(\beta, \mathbf{x}) &= u_2 + \beta \hat{c}_2 u_2' + \frac{(\beta \hat{c}_2)^2}{2} u_2'' + \mathcal{R}_2 \end{aligned}$$

The gain of agent i from a deal is defined by $g_i = \pi_i(\beta, \mathbf{x}) - Eu_i(w_i)$. The core is non-empty if $g_i \geq 0 \quad \forall i$. Thus, subtracting (A18) from (A17), dividing by $u_i'(A_i)$ and rearranging yields the following conditions:

$$\begin{aligned}
\bar{g}_1 &= -\frac{1}{2}(\beta\hat{c}_1)^2 R_1 - \beta\hat{c}_1 + \frac{1}{2}\left[v_1(2x_1^* - v_1)R_1 + 2v_1\right]p_1^* + m_1^*\left(1 + \frac{1}{2}m_1^*R_1\right) \\
\bar{g}_2 &= -\frac{1}{2}(\beta\hat{c}_2)^2 R_2 + \beta\hat{c}_2 + \frac{1}{2}\left[v_2(2x_2^* - v_2)R_2 + 2v_2\right]p_1^* - m_2^*\left(1 - \frac{1}{2}m_2^*R_2\right)
\end{aligned}
\tag{A19}$$

where $\bar{g}_i = \frac{g_i}{u'_i(A_i)}$. The roots of (A19) are:

$$\bar{\beta}_{1,2} = \frac{-1 \pm \Delta_1}{R_1\hat{c}_1} \quad \underline{\beta}_{1,2} = \frac{1 \pm \Delta_2}{R_2\hat{c}_2}
\tag{A20}$$

where

$$\begin{aligned}
\Delta_1 &= \sqrt{1 + \left[v_1(2x_1^* - v_1)p_1^* + m_1^{*2}\right]R_1^2 + 2(p_1^*v_1 + m_1^*)R_1} \\
\Delta_2 &= \sqrt{1 + \left[v_2(2x_2^* - v_2)p_1^* + m_2^{*2}\right]R_2^2 + 2(p_1^*v_1 - m_2^*)R_2}
\end{aligned}
\tag{A21}$$

The relevant solutions are, of course, non-negative values of β , where $\bar{\beta} = \max(\bar{\beta}_1, \bar{\beta}_2)$ and $\underline{\beta} = \min(\underline{\beta}_1, \underline{\beta}_2)$. As already mentioned above, the core is non-empty if and only if $\Delta\beta = \bar{\beta} - \underline{\beta} \geq 0$. Namely, the core is non-empty if and only if

$$\bar{\beta} = \frac{-1 + \Delta_1}{R_1\hat{c}_1} \geq \frac{1 - \Delta_2}{R_2\hat{c}_2} = \underline{\beta}
\tag{A22}$$

Clearly, (A22) can hold for non-negative values of $\bar{\beta}$ and $\underline{\beta}$ if and only if $R_i > 0 \forall i$. *. QED.*

References

- Adelstein R. P. (1978). The Plea Bargain in Theory: A Behavioral Model of the Negotiated Guilty Plea. *Southern Economic Journal*, 44, 488-503.
- Alexander C. (1992). The Kalai-Smorodinsky Bargaining Solution in Wage Negotiations. *The Journal of Operational Research Society*, 43(8), 779-786.
- Anbarci N., Skaperdas S. and Syropoulos C. (2002). Comparing Bargaining Solutions in the Shadow of Conflict: How Norms Against Threat Can Have Real Effects. *Journal of Economic Theory*, 106, 1-16.
- Arrow K. J. (1965). *Aspects of the Theory of Risk-Bearing*. Helsinki: Yrjö Jahnssonin Säätiö.
- Arrow K. J. (1971). The Theory of Risk Aversion. *Essays in the Theory of Risk-Bearing*.
- Aumann R. J. and Kurz M. (1977). Power and Taxes. *Econometrica*, 45(5), 1137-1161.
- Austen-Smith D. and Banks J. S. (1996). Information Aggregation, Rationality and the Condorcet Jury Theorem. *American Political Science Review*, 90(1), 34-45.
- Baker S. and Mazzetti C. (2001). Prosecutorial Resources, Plea Bargaining and the Decision To Go To Trial. *Journal of Law, Economics and Organization*, 17, 149-167.
- Bayindir-Upmann T. and Gerber A. (2003). *The Kalai-Smorodinsky Solution in Labor Market Negotiations*. CESifo working paper No. 941, Category 4: Labour Markets.
- Bebchuk L. A. (1984). Litigation and Settlement under Imperfect Information. *RAND Journal of Economics*, 15(3), 404-15.
- Becker G. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76, 169-217.

- Ben-Yashar R. and Nitzan S. (1997). The Optimal Decision Rule for Fixed Size Committees in Dichotomous Choice Situations – The General Result. *International Economic Review*, 38(1), 175-87.
- Bowers J. (2007). Grassroots Plea Bargaining. *Marq. L. Rev.*, 91.
- Brandenburger A. and Stuart H. (2007). Biform Games. *Management Science*, 53(4), 537-549.
- Call J. E., England D. E. and Talarico S. M. (1983). Abolition of Plea Bargaining in the Coast Guard. *Journal of Criminal Justice*, 11, 351-358.
- Carney F. J. and Fuller A. L. (1969). A Study of Plea Bargaining in Murder Cases in Massachusetts. *Suffolk University Law Review*, III(2), 292-307.
- Clark D. J. and Riis C. (1998). Contest Success Functions: An Extension. *Economic Theory*, 11, 201-204.
- Cooter R. D. and Rubinfeld D. L. (1989). Economic Analysis of Legal Disputes and Their Resolution. *Journal of Economic Literature*, XXVII, 1067-1097.
- Corhón L. and Dahm M. (2010). Foundations for Contest Success Functions. *Economic Theory*, 43, 81-98.
- Cornes R. and Hartley R. (2003). Risk Aversion, Heterogeneity and Contests. *Public Choice*, 117, 1-25.
- Coughlan P. J. (2000). In Defense of Unanimous Jury Verdicts: Mistrials, Communication and Strategic Voting. *American Political Science Review*, 94 (2), 375-393.
- Dekel E. and Piccione M. (2000). Sequential Voting Procedures in Symmetric Binary Elections. *Journal of Political Economy*, 108(1), 34-54.
- Eesteban J. and Sákovics J. (2002). Endogenous Bargaining Power. *Universitat Pompeu Fabra Working Paper No. 644*.
- Epstein G. S. and Nitzan S. (2007). *Endogenous Public Policy and Contests*. Berlin, Heidelberg: Springer-Verlag.
- Feddersen T. and Pesendorfer W. (1998). Convicting the Innocent: The Inferiority of Unanimous Jury Verdict. *American Political Science Review*, 92(1), 1-25.
- Foncel J. and Treich N. (2005). Fear of Ruin. *Journal of Risk and Uncertainty*, 31(3), 289-300.
- Franzoni L. A. (1999). Negotiated Enforcement and Credible Deterrence. *The Economic Journal*, 109, 509-35.

- Gazal O. and Bar-Gill O. (2004). Plea Bargaining Only for the Guilty. *University of Michigan Law School, The John M. Olin Center for Law & Economics Working Paper Series, Paper 30*.
- Gigerenzer G. (1991). How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases". *European Review of Social Psychology*, 2, 83-115.
- Grossman G. and Katz M. I. (1983). Plea Bargaining and Social Welfare. *The American Economic Review*, 73(4), 749-57.
- Grossman S. J. and Hart O. D. (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*, 94, 691-719.
- Hart O. and Moore J. (1990). Property Rights and the Nature of the Firm. *Journal of Political Economy*, 95(6), 1119-1158.
- Hay B. L. (1995). Effort, Information, Settlement, Trial. *Journal of Legal Studies*, 24, 29-62.
- Hillman A. L. and Katz E. (1984). Risk-Averse Rent-Seekers and the Social Cost of Monopoly Power. *Economic Journal*, 94, 104-110.
- Hillman A. L. and Samet D. (1987). Dissipation of Rents and Revenues in Small Number Contests. *Public Choice*, 54, 63-82.
- Hirshleifer J. and Osborne E. (2001). Truth, Effort, and the Legal Battle. *Public Choice*, 108, 169-195.
- Huff C., Rattner A. and Sagarin E. (1986). Guilty Until Proved Innocent: Wrongful Convictions and Public Policy. *Crime and Delinquency*, 32 (4), 518-544.
- Kahneman D. and Tversky A. (1996). On the Reality of Cognitive Illusions. *Psychological Review*, 103(3), 582-591.
- Kahneman D. and Tversky A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-292.
- Kalai E. and Smorodinsky M. (1975). Other Solutions to Nash's Bargaining Problem. *Econometrica* (43), 513-518.
- Katz A. (1988). Judicial Decisionmaking and Litigation Expenditure. *International Review of Law and Economics*, 8, 127-143.
- Kim J. Y. (2010). Credible Plea Bargaining. *European Journal of Law and Economics*, 29, 279-293.

- King N. J., Soulé D. A., Steen S. and Weidner R. R. (2005). When Process Affects Punishment: Differences in Sentences After Guilty Plea, Bench Trial, and Jury Trial in Five Guidelines States. *Columbia Law Review*, 105, 959-1009.
- Kipnis K. (1976). Criminal Justice and the Negotiated Plea. *Ethics*, 86(2), 93-106.
- Kobayashi B. H. (1992). Deterrence with Multiple Defendants: An Explanation for Unfair Plea Bargains. *RAND Journal of Economics*, 23(4), 507-17.
- Kobayashi, B. H. (1992). Deterrence with Multiple Defendants: An Explanation for Unfair Plea Bargains. *RAND Journal of Economics*, 23(4), 507-17.
- Konrad Kai A. and Schlesinger H. (1997). Risk Aversion in Rent-Seeking and Rent-Augmenting Games. *The Economic Journal*, 107, 1671-1683.
- Landes, W. M. (1971). An Economic Analysis of the Courts. *Journal of Law and Economics*, 14, 61-107.
- Lott J. R., Jr. (1987). Should the Wealthy Be Able to 'Buy Justice'? *Journal of Political Economy*, 95(6), 1307-16.
- McDonald I. and Solow R. M. (1981). Wage Bargaining and Employment. *The American Economic Review*, 71(5), 896-908.
- Miceli T. J. (1990). Optimal Prosecution of Defendants Whose Guilt is Uncertain. *Journal of Law, Economics and Organization*, 6, 189-201.
- Millner E. L. and Pratt M. D. (1991). Risk Aversion and Rent-Seeking: An Extension and Some Experimental Evidence. *Public Choice*, 69, 81-92.
- Miyagawa E. (2001). Subgame-Perfect Implementation of Bargaining Solutions. *Columbia University, Department of Economics Working Paper No. GEB#703031.1*.
- Moulin H. (1984). Implementing the Kalai-Smorodinsky Bargaining Solution. *Journal of Economic Theory*, 33, 32-45.
- Muto S., Nakayama M., Potters J. and Tijs S. (1998). On Big-Boss Games. *The Economic Studies Quarterly*, 39(4), 303-321.
- Nash J. (1950). The Bargaining Problem. *Econometrica*, 18(2), 155-162.
- Newman D. J. (1966). *Conviction: Determination of Guilt and Innocence Without Trial*. Boston: Little, Brown and Co.
- Nitzan S. (1994). Modeling Rent-Seeking Contests. *European Journal of Political Economy*, 10, 41-60.
- Nti, K. O. (1999). Rent-Seeking with Asymmetric Valuations. *Public Choice*, 98, 415-430.

- Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, 32, 122-136.
- Reiganum J. (1988). Plea Bargaining and Prosecutorial Discretion. *American Economic Review*, 78(4), 713-28.
- Reiganum J. (2000). Sentencing Guidelines, Judicial Decretion and Plea Bargaining. *RAND Journal of Economics*, 31(1), 62-81.
- Rubinstein A. (1982). Perfect Equilibrium in a Bargaining Model. *Econometrica*, 50 (1), 97-109.
- Rubinstein M. L. and White T. J. (1979). Alaska's Ban on Plea Bargaining. *Law and Society*, 367-383.
- Schulhofer S. J. (1992). Plea Bargaining as Disaster. *Yale Law Journal*, 101, 1979-2009.
- Scott R. E. and Stunz W. J. (1992). Plea Bargaining as a Contract. *Yale Law Journal*, 101, 1909-1968.
- Skaperdas S. (1996). Contest Success Functions. *Economic Theory*, 7, 283-290.
- Skaperdas S. (2006). Bargaining Versus Fighting. *unpublished*.
- Skaperdas S. and Gan L. (1995). Risk Aversion in Contests. *The Economic Journal*, 105, 951-962.
- Stuart H. W., Jr. (2005). Biform Analysis of Inventory Competition. *Manufacturing & Service Operations Management*, 7(4), 347-359.
- Svejnar J. (1986). Bargaining Power, Fear of Disagreement, and Wage Settlements: Theory and Evidence from U.S. Industry. *Econometrica*, 54(5), 1055-1078.
- Tijs S. (1990). Big Boss Games, Clan Games and Information Market Games. (I. Tatsuru, A. Neyman, & Y. Tauman, Eds.) *Game Theory and Applications*, pp. 410-412.
- Tor A., Gazal-Ayal O. and Garcia M. (2006). Innocence, Fairness and the Willingness to Accept Plea Bargain Offers. *Available at SSRN*: <http://ssrn.com/abstract=880506>.
- Ulmer J. T. and Bradley M. S. (2006). Variation in Trial Penalties Among Serious Violent Offences. *Criminology*, 44(3), 631-670.
- Ulmer J. T., Eisenstein J. and Johnson B. D. (2009). Trial Penalties in Federal Sentencing: Extra-Guidelines Factors and District Variation. *Justice Quarterly*, 1-33.

- Weininger R. A. (1987). The Abolition of Plea Bargaining: A Case Study of El Paso County, Texas. *UCLA Law Review*, 35: 265, 265-313.
- Wright R. and Miller M. (2002). The Screening/Bargaining Tradeoff. *Stanford Law Review*, 55.
- Wright R. F. and Engen R. L. (2007). Charge Movement and Theories of Prosecutors. *Marquette Law Review*, 101-128.
- Wright R. F. and Engen R. L. (2006). The Effect of Depth and Distance in a Criminal Code on Charging, Sentencing and Prosecutor Power. *N.C.L. Rev.*, 1935.
- Wright R. F. (2005). Trial, Distortion and the End of Innocence in Federal Criminal Justice. *University of Pennsylvania Law Review*, 154(79), 79-156.
- Yilankaya O. (2002). A Model of Evidence Production and Optimal Standard of Proof and Penalty in Criminal Trials. *Canadian Journal of Economics*, 35(2), 385-409.