The judicial process is modeled as a biform contest. A sufficient condition for a non-empty core is provided, and it is shown that the effect of the severity of charges on the core is ambiguous. The practice of plea bargaining actually applies Moulin’s (1984) mechanism which non-cooperatively implements the Kalai-Smorodinsky (1975) bargain solution, implying that equilibrium plea deals are regressive. Namely, defendant’s gain and social cost from a guilty plea deal both increase with the severity of the crime. The charge reduction rate is inversely related to the defendant’s, and positively related to the prosecutor’s, “fear of ruin” index.

1. INTRODUCTION

Plea bargaining involves negotiations between a defendant and a prosecutor regarding the conditions under which the defendant will enter 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 (Grossman and Katz, 1983; Franzoni, 1999; Wright, 2005; Ulmer et al., 2010), but the rate varies with the seriousness of the crime.  

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,

---

*I wish to thank Aviad Heifetz for very helpful comments on earlier versions of this work.

1 Plea bargains constituted 98% of convictions for embezzlement, immigration offenses, and gambling, about 80% for kidnapping and assault, but only 67% of those for murder (Reinganum, 2000).
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, 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 on the part 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 to 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 offense if he pleads guilty. According to the Federal Sentencing Guidelines (FSG), a guilty plea is not recognized as relevant for sentencing (Ulmer et al., 2010). Nevertheless, Wright (2005), King et al. (2005) and Ulmer et al. (2009) found significant differences between sentences imposed on trial-convicted defendants and those imposed on 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. That is, defendants who plead guilty enjoy charge reduction and additional sentence discounts. Charge reductions and trial penalties cause plea deals to be regressive; the defendant’s gain as well as the social cost from a guilty plea deal both increase with the severity of the crime.

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 the evidence. More recent studies shifted to the contest model (c.f. Katz, 1988; Hirshleifer and Osborne, 2001), but the lottery model has remained popular in plea bargaining analyses (c.f. Newman, 1966; Adelman, 1978; Grossman and Katz, 1983; Bechuk, 1984; Lott, 1987; Reinganum, 1988; Kobayashi, 1992; Reinganum, 2000; Baker and Mazetti, 2001; Kim, 2010). The few exceptions lack an explicit analysis of the judicial process as a contest, assume that litigants are risk-neutral and that the prosecutor is sincerely motivated (c.f. Landes, 1971; Lott, 1987; Franzoni, 1999; Yilankaya, 2002; Hay, 1995). These studies investigated, for example the

---

2 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).

3 For a survey of earlier literature, see Cooter and Rubinfeld (1989).
effect of expected trial outcome on litigants’ expenditures (c.f. Katz, 1988), or vice versa (c.f. Skaperdas and Gan, 1995). However, 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, that is, 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.

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 the conviction probability is exogenously given, according to evidence strength, and although more recent studies shifted to the contest model, plea bargaining analyses have 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. 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

---

4 See, for example, Kipnis (1976), Huff, Rattner and Sagarin (1986), Scott and Stunz (1992), Schulhofer (1992), Gazal and Bar-Gill (2004), and Wright (2005).
courtroom and enables the parties to estimate their chances. Thus, the plea agreement reflects the expected verdict more efficiently. 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 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.6

5 Ulmer et al. (2010) 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.

6 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. (2010), Bowers (2007) and other researchers have 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, the assumption of sincere judges and jurors ignores their incentive to save their own time and effort, and to encourage litigants to end the process as quickly as possible. The strategic bias of judges and jurors has been studied extensively in the literature. See, for example, Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1998), Dekel and Piccione (2000), and Coughlan (2000).
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. 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 \( x = \{x_1, x_2\} \) and assume that the winning probability of agent \( i \) is given by a Tullock (1980) type Contest Success Function,

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

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 \( d_i > 1 \), the prosecutor bears a heavier standard of proof. For the sake of simplicity, assume \( h_1(x_1) = x_1 \) and \( d = 1 \). This simplifying assumption does not affect the qualitative results of the model.\(^7\)

\(^7\) For axiomatization of Contest Success Functions, see Skaperdas (1996), Clark and Riis (1998), Epstein and Nitzan (2007) and Corbón and Dahm (2010).
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 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,

$$w'_i = A_i - x_i, \quad w''_i = A_i - x_i - v_i$$

(2)

$$w'_2 = A_2 - x_2, \quad w''_2 = A_2 - x_2 + v_2.$$

### 3.2. ATTITUDE TOWARD RISK

Assuming risk aversion significantly complicates the analysis (Hillman and Katz, 1984; Skaperdas and Gan, 1995; Konrad and Schlesinger, 1997; Cornes and Hartley, 2003). Therefore, the rent-seeking literature generally assumes risk-neutrality (see Epstein and Nitzan, 2007). 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 $R_i = -u''/u'$, 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 = u_i/ u_i' $. 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 a 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, $\lim_{v_i \to 0} q_i/v_i = u_i'/u_i$. 

---

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

9 The agent indifference condition is $u_i (w_i) = (1 - q_i) u_i (w_i + v_i) + q_i u_i (0)$. Normalizing $u_i (0) = 0$ and rearranging yields $q_i/v_i = [u_i (w_i + v_i) - u_i (w_i)]/v_i u_i (w_i + v_i)$, implying $\lim_{v_i \to 0} q_i/v_i = u_i'/u_i = 1/F_i$. 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.”

Review of Law & Economics, © 2012 by De Gruyter
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

\[
(3) \quad E u_i (w_i) = p_i (x) u_i (w'_i) + p_2 (x) u_i (w''_i), \quad i = 1, 2.
\]

First order conditions for interior solution are,

\[
(4) \quad p_i^* (x^*) \Delta u_i - E u'_i = 0 \quad \forall i
\]

where asterisks denote equilibrium values, \( p_i' = \partial p_i / \partial x \), \( \Delta u_i = u_i (w'_i) - u_i (w''_i) \) and \( E u'_i = p_i (x^*) u'_i (w'_i) + (1 - p_i (x^*)) u'_i (w''_i) \). With risk-neutral contestants, (4) is reduced to \( p_i (\hat{x}) = 1 \quad \forall i \), implying

\[
(5) \quad \hat{x} = \left\{ \frac{v_1^2 v_2}{(v_1 + v_2)^2}, \frac{v_1 v_2^2}{(v_1 + v_2)^2} \right\}, \quad \hat{p} = \left\{ \frac{v_1}{v_1 + v_2}, \frac{v_2}{v_1 + v_2} \right\}
\]

where \( \hat{x} \) and \( \hat{p} \) 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 \( (S,V,a) \), where \( S = S^1 \times S^2 \times ... \times S^n \) is a profile of strategies, \( V : A \times 2^n \to \mathbb{R} \) is a cooperative game coalitional function and \( a = \{ \alpha_i \}_{i=1}^n \) 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.10

---

10 For a more accurate definition and explanation of the confidence index, see Brandenburger and Stuart (2007) or Stuart (2005).
Similarly, a \textit{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 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 \((x, p, u, b)\) where \(x = \{x_i\}_{i=1}^n\) is the contestants’ rent-seeking expenditures vector, \(p\) is the contest success function, \(u\) is a vector of contestants’ utilities and \(b\) is a bargaining problem.\footnote{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, Esteban and Sákovics (2002), Bayindir-Upmann and Gerber (2003), and Skaperdas (2006).}

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 

\[
\begin{align*}
    w_i^1 &= A_i - p_i v_i^1 - x_i^1 \\
    w_i^2 &= A_i + p_i v_i^2 - x_i^2
\end{align*}
\]

respectively. Assuming risk-aversion and applying Jensen’s inequality implies that \(u_i(w_i^*) \geq Eu_i(w_i)\). Hence, there is a \(w_i < w_i^*\) satisfying \(u_i(w_i) = Eu_i(w_i)\), (see Figure 1). \(w_i^*\) is the \textit{certainty equivalent} wealth of agent \(i\).

A plea bargain aims at avoiding contest confrontation through a compromise function \(f^i(v)\), \(v = \{v_i, v_j\}\), that assigns the defendant a charge reduction of \(f_i^1(v)\) in return for pleading guilty, and consequently \(f_j^2(v)\) to the prosecutor. Skaperdas and Gan (1995) analyzed a compromise in a contest problem, emphasizing that “[a]n agreed division of the prize, though, does not imply the...
absence of effort on 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, which means to be well-equipped for the conflict, if necessary. If an agreement is achieved, the defendant’s and the prosecutor’s net wealth are \( w_i = A_i - p_2 f_1 (v) - x_i \) and \( w_2 = A_2 + p_2 f_2 (v) - x_2 \), respectively. A compromise agreement is achievable if it is mutually beneficial, namely, if it ensures each contestant \( u_i (w_i) \geq u_i (w_j) \).  

**4.1. The Core of a Plea Bargaining Biform Contest**

For simplicity, assume \( f(v) = \beta v \). 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

\[
\pi_1 (x) = u_1 (A_1 - p_2 \beta v_1 - x_1),
\]

and the prosecutor’s target function is

\[
\pi_2 (x) = u_2 (A_2 + p_2 \beta v_2 - x_2),
\]

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:

\[
\hat{x} = \beta \hat{x} \Rightarrow \hat{p} = \hat{p}
\]

where \( \hat{x} \) and \( \hat{p} \) are defined above in (5). That is, as pointed out by Skaperdas and Gan (1995), bargaining position optimization implies that contestants’ expenditures are proportional to risk-neutral agents’ expenditures, where \( \beta \) 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 they are easily obtained from (8):

\[
\hat{x} = \beta \hat{x} \Rightarrow \hat{p} = \hat{p}
\]

---

12 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. \( p_2 v_i \) is the penalty discount for the defendant which reflects its expected value, and \( f(v) \) is the rate of charge reduction.
The derivation of (9) was based on the assumption that in the first stage, the sharing rule, $\beta$, is exogenous and common knowledge. In real world plea bargaining negotiations, however, $\beta$ lies in the center of the dispute between the parties. Therefore, we first have to define the upper and lower limits of $\beta$.

In other words, we now turn to explore the core of the plea bargain game.

A sharing rule $\beta$ (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

$$g_i(\beta) = \pi_i(\beta, \tilde{x}) - Eu_i(w, x).$$

It can be verified that $\partial g_1/\partial \beta < 0$, $\partial g_2/\partial \beta > 0$, $\partial^2 g_1/\partial \beta^2 < 0$ and $\partial g_2/\partial \beta > 0$, 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 $B(\beta) = \{\beta | g_i(\beta) \geq 0, \forall i\}$.

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 $\beta_{min} = \inf \{\beta \in B(\beta)\}$ and $\beta_{max} = \sup \{\beta \in B(\beta)\}$ (corresponding to points $A$ and $B$ in Figure 2, respectively). In other words, the defendant is ready to plead guilty if $\beta \leq \beta_{min}$ while the prosecutor will sign the agreement if $\beta \geq \beta_{max}$.

**Proposition 1:**

(a) $\beta_{min} = \frac{\hat{c}_2}{\hat{c}_1}$, $\beta_{max} = \frac{\hat{c}_1}{\hat{c}_2}$.

(b) $B(\beta) \neq \emptyset \Leftrightarrow v_1 \geq v_2 - \frac{\partial X}{\partial \beta} \bigg|_{\beta = \hat{\beta}}$, where $\hat{\beta} = \frac{\hat{c}_2}{\hat{c}_1}$.

**Proof:** See appendix.
Part (a) of Proposition 1 states that the bounds of $B(\beta)$ are determined by $\frac{c_i}{\hat{c}_i}$, which is the ratio between agent $i$’s expected payoffs under risk-aversion and risk-neutrality assumptions, respectively. The ratio $\frac{c_i}{\hat{c}_i}$ is thus the bound of the risk-premium that a risk-averse contestant is willing to pay in order to avoid risky confrontation. Part (b) of Proposition 1 implies that the core is non-empty if the defendant’s stake $v_1$ is at least as large as the prosecutor’s stake $v_2$ minus the change in total expenditures ($X$) induced by a change in conviction probability (evaluated at $\hat{\beta} = \frac{c_1}{\hat{c}_1}$). Notice that 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 for $u_i(w_i) = -e^{-w_i}$. Figure 3 presents combinations of $v_1$ and $v_2$ for a non-empty core simulated for this specification.

![Figure 3](image)

4.2. COMPARATIVE STATICS OF THE CORE

A non-empty core may contain an infinite number of rational deals, but 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 $g$ 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. (See Figure 4. $d$ in Figure 4 denotes the disagreement point).
In a plea bargaining context, disagreement means going to court, implying \( d = (Eu_1, Eu_2) \). Hence, both \( B \) and \( d \) are endogenous, implying, as many authors have indicated,\(^\text{13}\) that comparative statics of the core in a biform contest framework are complicated. To see this, notice that an increase in the prosecutor’s first-stage efforts, for instance, shifts the disagreement point from \( d_1 \) to \( d_2 \) and, consequently, the whole bargain set from \( B_1 \) to \( B_2 \) (see Figure 5).

For example, consider the effect of “fear of disagreement” on the core. Denote the elasticity of agent’s expenditures with respect to \( F_i \) by \( \eta_{x_i,F_i} \), the elasticity of conviction probability with respect to \( F_i \) by \( \eta_{p_i,F_i} \). Differentiating \( \beta_{\min} \) and \( \beta_{\max} \) with respect to \( F \) and manipulating algebraically yields:

\[
\begin{align*}
(a) \quad \frac{\partial \beta_{\min}}{\partial F_i(A_i)} &\geq 0 \iff \frac{p_{i}^{2}v_{i}}{x_{i}} \geq \frac{\eta_{x_{i},F_{i}}}{\eta_{p_{i},F_{i}}} \\
(11) \quad \frac{\partial \beta_{\max}}{\partial F_i(A_i)} &\geq 0 \iff \frac{\eta_{x_{i},F_{i}}}{\eta_{p_{i},F_{i}}} \geq \frac{\eta_{x_{i},F_{i}}}{\eta_{p_{i},F_{i}}}.
\end{align*}
\]

\(^{13}\) For example, see McDonald and Solow (1981), Alexander (1992), Anbarci, Skaperdas and Syropoulos (2002), Bayindir-Upmann and Gerber (2003), and Skaperdas (2006).
Namely, an increase in agent’s “fear of disagreement” index shifts the relevant boundary of $B(\beta)$ rightwards (leftwards) if his expected payoff to expenditure ratio $(p^{\beta}_{i}/x_{i})$ is higher (lower) then his elasticities ratio. It requires some tedious algebra to show that $\text{sgn}(\partial \beta_{\text{max}}/\partial v_{i})$, and $\text{sgn}(\partial \beta_{\text{max}}/\partial v_{i})$ are indeterminate.

**Corollary 1:** If $v_{i}$ is positively correlated with the severity of the charges, the charges severity effect 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 lower and upper limits of these cores by $[\beta_{\text{min}}^{L}, \beta_{\text{max}}^{L}]$ and $[\beta_{\text{min}}^{U}, \beta_{\text{max}}^{U}]$, 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 choose strategy $a$ if $\alpha' u_{i}(\beta_{\text{min}}^{L}) + (1-\alpha') u_{i}(\beta_{\text{max}}^{L}) > \alpha' u_{i}(\beta_{\text{max}}^{L}) + (1-\alpha') u_{i}(\beta_{\text{max}}^{U})$.

![Figure 6](image)

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 (11) and Corollary 1, we know that the effect of the “fear of ruin” index on $\beta_{\text{min}}$ and $\beta_{\text{max}}$ is ambiguous. The confidence index is probably correlated with the “fear of ruin” index, and this correlation should be taken into account. Secondly, equation (8) implies that equilibrium expenditures and $\beta$ are interrelated. Namely, expenditures depend on the anticipated specific value of $\beta$, while the boundaries of the $\beta$ values depend on first stage equilibrium expenditures. The analysis of these interrelations requires a unique and precise bargaining solution.
4.4. THE PLEA BARGAINING SOLUTION

Nash (1950) suggested a unique bargaining 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 monotonicity axiom (M) by Kalai and Smorodinsky (1975), who suggested an alternative unique bargaining solution that satisfies IAT, PO, S and M.

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

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

Namely, the KS solution equalizes the proportional gains of every agent, relative to his utopian gain.\(^{14}\) In Figure 7, the KS solution is the intersection point of the $d m$ line and the efficiency frontier.

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.

**Round 1:** Suppose that $p_i > p_j$. Contestant $i$ makes a proposal of $\beta_i$ to $j$. If $j$ accepts the proposal, it is implemented by the court.\(^{15}\) If $j$ rejects the proposal, proceed to Round 2.

\(^{14}\) Alternatively, the KS solution is defined by $ks_i = \max_{\beta \in (\beta_{min}, \beta_{max})} \min_i (g_i(\beta)/m_i)$.

\(^{15}\) 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.
**Round 2:** Contestant $j$ makes a counter-proposal, $\beta_j$. 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.\(^{16}\)

The PBP game-form is illustrated in Figure 8.

![Figure 8](image)

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 $d_i = \hat{p}_i u_i (w_i (I)) + \hat{p}_j u_j (w_j (II))$. For convenience, let us normalize $d = 0$. Of course, legally $\beta \in (0, 1)$, thus $m_i = \pi_1 |_{\beta_0}$ and $m_2 = \pi_2 |_{\beta_1}$, implying that the KS plea bargaining solution is $\beta \in (0, 1)$, which solves:

$$\begin{align*}
\frac{\pi_1 (\beta)}{m_1} &= \frac{\pi_2 (\beta)}{m_2}.
\end{align*}$$

**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 $k_S = \hat{p}_j$. Figure 9 presents a simulation of $k_S$ for this function. Figure 9 reveals that $k_S$ declines with $v_1$ and increases with $v_2$, reflecting the regressive characteristic of plea deals, as formally established below in Proposition 3.

\(^{16}\) The rationale of Round 2 stems from the assumption of “almost” automatic approval of guilty plea agreements in Round 1, as explained in footnote 15. That is, 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 $p_i$. 
4.5. COMPARATIVE STATICS OF THE SHARING RULE

Non-cooperative subgame-perfect implementation of a bargaining solution, $f(v)$, implies that $\tilde{x}$ and $\tilde{p}$ are given by (8), 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 (13) is

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

Solving (14) for $\beta$, inserting $\hat{c}_1 = v_1 v_2 (2v_1 + v_2) / (v_1 + v_2)^2$ and $\hat{c}_2 = v_1^2 / (v_1 + v_2)^2$ and rearranging, yields:

$$k_s = \frac{v_1^2 F_1}{F_2 v_1 (2v_1 + v_2) + v_1^2 F_1}.$$
Proposition 3:

(a) The charge reduction rate is positively related to the defendant’s stake and inversely related to the prosecutor’s stake.

(b) The charge reduction rate is inversely related to the defendant’s, and positively related to the prosecutor’s, “fear of ruin” index.

Proof: See appendix.

Proposition 3 postulates that plea deals are regressive, as mentioned above. The regressive characteristic of plea deals is related to the monotonicity axiom. An increase in \( v_1 \) shifts the \( P(B) \) curve rightward except for the \( m_2 \) point. Hence, the intersection point of the \( dm \) line and the efficiency frontier in Figure 7 also shifts rightwards, implying that \( g_i \) increases unambiguously.

Define the defendant’s “discount function” by \( G_i = \hat{p}_i v_i (1-ks) \). 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 \( G_z = -\hat{p}_z v_z (1-ks) \). Plugging \( \hat{c}_i, \hat{c}_2, \hat{p}_2 \) and \( ks \) into \( G_i, i = 1,2 \) yields:

\[
G_i = \frac{v_i^2 v_j F_j (2v_i + v_j)}{(v_i + v_j)(F_i v_i + F_j v_j)}, \quad i = 1,2
\]

Proposition 4:

(a) \( G_i \) increases with \( v_i \), but the effect of \( v_2 \) on \( G_2 \) is ambiguous.

(b) The cross effect of \( v_i \) on \( G_j \) is ambiguous.

(c) \( G_i \) decreases with \( F_i \) and increases with \( F_j \).

Proof: See appendix.

Assuming that \( v_1 \) is positively correlated with the severity of the original charges implies (by Proposition 4 again) that plea deals are regressive, as by part (a) of Proposition 4 the defendant’s gain from a plea deal increases with his stake. Part (b) denies the common wisdom of an unambiguous negative cross effect of \( v_i \) on \( G_j \), while part (c) confirms the common wisdom about the expected effect of \( F_i \) on \( G_i \).

Figure 10 presents simulations for \( G_1 \) and \( G_2 \), respectively, assuming \( u_i = -e^{-u} \). 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.”

Figure 10

![Graphs showing the relationship between G, v, and v2.]

Define $G = G_1 + G_2$. It follows that:

$$G = \frac{(v_1 - v_2) v_1 v_2 F_1 (2v_1 + v_2)}{(v_1 + v_2) (F_2 v_1 (2v_1 + v_2) + F_1 v_2^2)},$$

implying Corollary 2, that provides sufficient condition under which plea-bargaining is a positive, zero, or negative sum game, and the effect of the “fear of disagreement” index on this sum.

17 As a measure of the social costs associated with plea deals, $G$ is downwards biased, because it ignores the societal costs associated with lenient sentences for ruthless criminals, as well as the societal costs associated with the increased risk of innocent convictions.
Corollary 2:

(a) \( G \gtrless 0 \Leftrightarrow v_i \lesssim v_2 \)

(b) \( \frac{\partial G}{\partial F_i} \lesssim 0 \Leftrightarrow v_i \lesssim v_j \)

(c) \( \text{sgn}\left(\frac{\partial G}{\partial v_i}\right) \text{ is indeterminate.} \)

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 \( v_i = 100 \) 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 (Millner and Pratt, 1991).\(^{18}\) Konrad and Schlesinger (1997) show that \( x'_i \gtrless \tilde{x}_i \Leftrightarrow Eu'_i(\tilde{x}) \lesssim \Delta u'_i(\tilde{x})/v_i \). However, equation (8) implies \( x \lesssim \tilde{x} \Leftrightarrow x \lesssim f\tilde{x} \). Hence, Konrad and Schlesinger’s condition is insufficient in a competitive environment.

Corollary 3: 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 \( \tilde{X} = k \tilde{x}X' \)). The upper row of Figure 12 contains combinations of \( v_1 \) and \( v_2 \), satisfying \( \Delta X \gtrless 0 \) for various values of \( F_1 \) and \( F_2 \), while the diagrams in the lower row

\(^{18}\) 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).
of Figure 12 contain combinations of $F_1$ and $F_2$, satisfying $\Delta X \geq 0$ for various values of $v_1$ and $v_2$ (simulated for $u_i(w_i) = -e^{-w_i}$).

The ambiguity of plea bargaining effect on expenditures undermines the common “public resources savings” justification for the plea bargaining practice.\textsuperscript{19} 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 “[w]hen 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 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., 2010).”\textsuperscript{20}

\textsuperscript{19} 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” (\textit{Santobello v. New-York} 404 U.S. 257, 261 (1971)).

\textsuperscript{20} For case studies of the effect of the abolition of plea bargaining, see, for example, Rubinstein and White (1979), Call, England and Talatico (1983) and Weininger (1987).
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 $R_i = -u''/u'$.

**Proposition 5**: $B(\beta) \neq \emptyset \Leftrightarrow R_i \geq 0, \forall i$.

**Proof**: See appendix.

Proposition 5 states that the core is non-empty if and only if both contestants are risk-aversers. Apparently, this trivial result contradicts Kahneman and Tversky’s (1979) prospect theory. According to 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 $U(\pi) = \sum_{i=1}^{l} \phi(p_i)u(w_i)$, where $w_i$ is a contingent outcome and $\phi(p_i)$ is the probability weighting function, 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, and the value function is S-shaped around it (see Figure 13), implying that individuals are not risk-aversers but loss-aversers, 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.21

---

21 Kahneman and Tversky based their theory on experimental evidence, but their methodology has evoked polemics. See, for example, Gigerenzer (1991) and Kahneman and Tversky’s (1996) reply. However, this polemic is beyond the scope of this article.
A defendant who wins in a trial is left with net wealth of $A_i - x_i$, and a defendant who loses is left with $A_i - x_i - v_i$. That is, 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 5 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 $w_i^m$. However, this interpretation is incompatible with the bulk of evidence on defendants’ behavior. On the one hand, some researchers (c.f. Tor et al., 2006) argue that defendants are subject to innocence bias, which causes them to over-estimate their winning probabilities and reject any guilty plea offers. Innocence bias implies that in Figure 1, the defendants’ reference point is to the right of $w_i^m$. On the other hand, reports from the Innocence Project organization revealed cases of innocent defendants falsely convicted after pleading guilty, implying that when the innocence bias is not binding, the defendants’ reference point is, indeed, $w_i^m$.

Motti Michaeli has 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 14, 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 14

---

22 These reports are available at http://www.innocenceproject.org/.
23 A Ph.D. student at the Center for the Study of Rationality (Hebrew University of Jerusalem), under the supervision of Professor Robert J. Aumann.
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, $k_\chi$ is positively correlated with the defendant’s “fear of disagreement,” and negatively correlated with the prosecutor’s “fear of disagreement.” However, plea deals are regressive, as $k_\chi$ is negatively correlated with the defendant’s stake and positively correlated with the 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.”

---

24 See also Wright and Engen (2007).
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 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 increase with the severity of the offense, implying that charge reduction weakens the innocence bias effect and encourages even innocent defendants to plead guilty.\textsuperscript{25} I conjecture that as the severity of the charges increases, deeper charge reduction is required by the $k_s$ solution, leading the law enforcement system to press legislators to supply more charging options for certain felony crimes.\textsuperscript{26}

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.”\textsuperscript{27} 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 sincere 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

\[
Eu_1 (w_1) = p_1 u_1 (A_i - x_i^*) + p_2 u_1 (A_i - z_i)
\]

(A1)

\[
Eu_2 (w_2) = p_1 u_2 (A_2 - x_2^*) + p_2 u_2 (A_2 + z_2)
\]

\textsuperscript{25} 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.

\textsuperscript{26} Bowers (2007) found that in petty cases, charge reduction is relatively rare and the plea bargain is centered mainly on the sentence.

\textsuperscript{27} Babylonian Talmud, tractate Gittin 45a.
where $z_1 = (v_1 + x_1^*)$ and $z_2 = (v_2 - x_2^*)$. On the other hand, if the two agents agree on a deal, $\beta(v_i)$, their expected utilities are

$$\pi_1(\beta, x) = u_1(A_1 - \beta \hat{c}_1)$$

(A2)

$$\pi_2(\beta, x) = u_2(A_2 + \beta \hat{c}_2)$$

where $\hat{c}_1 = \hat{p}_2v_1 + \hat{x}_1$, $\hat{c}_2 = \hat{p}_2v_2 - \hat{x}_2$, and $\hat{x}$ and $\hat{p}$ are defined in (5). Taking the Taylor expansion of (A1) yields:

$$Eu_i(w_i) = p_i^*\left[u_1(A_1) - x_1u'_1(A_1)\right] + p_2^*\left[u_1(A_1) - z_1^*u'_1(A_1)\right] + \delta_i$$

(A3)

$$Eu_2(w_2) = p_i^*\left[u_2(A_2) - x_2u'_2(A_2)\right] + p_2^*\left[u_2(A_2) + z_2^*u'_2(A_2)\right] + \delta_2$$

where $\delta_i$ is the remainder of the Taylor sequence. Similarly, the Taylor expansion of (A2) yields:

$$\pi_1(\beta, x) = u_1(A_1) - \beta \hat{c}_1u'_1(A_1) + \delta_1$$

(A4)

$$\pi_2(\beta, x) = u_2(A_2) + \beta \hat{c}_2u'_2(A_2) + \delta_2.$$

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

$$\overline{g}_i = \left(1 - p_i^* - p_2^*\right)F_i(A_1) - \beta \hat{c}_1 + p_i^*\hat{x}_1 + p_2^*\hat{z}_1 \geq 0$$

(A5)

$$\overline{g}_2 = \left(1 - p_1^* - p_2^*\right)F_2(A_2) - \beta \hat{c}_2 + p_1^*\hat{x}_2 + p_2^*\hat{z}_2 \geq 0$$

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

$$\beta_{\text{min}} = \frac{c_2^*}{\hat{c}_2}, \quad \beta_{\text{max}} = \frac{c_1^*}{\hat{c}_1}.$$

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

(A7) \[ \beta_{\text{min}} = \frac{p_2^*v_2 - x_2^*}{\hat{p}_2^*v_1}, \quad \beta_{\text{max}} = \frac{p_2^*v_1 + x_1^*}{\hat{p}_2^*(2 - \hat{p}_2)}. \]

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

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

The left-hand side of the inequality in (A8) is constant. Denote this constant by $\hat{\beta} = \hat{c}_2/\hat{c}_1$. It follows that (A8) implies that the “fear of ruin” affects $c_i^*$ (the denominator on the right side of (A8)) more than $c_j^*$ (the numerator). Suppose that the “fear of ruin” increases for both agents at the same rate, namely, $\partial F_i = \partial F_2 = \partial F$. It follows that

(A9) \[ \left( \frac{\partial p_2}{\partial F}v_2 - \frac{\partial x_2}{\partial F} \right)_{\hat{\beta}} \leq \left( \frac{\partial p_2}{\partial F}v_1 + \frac{\partial x_1}{\partial F} \right)_{\beta}. \]

Inserting $\partial X/\partial F = \partial x_1/\partial F + \partial x_j/\partial F$ into (A9) and rearranging yields:

(A10) \[ v_1 \geq v_2 = \frac{\partial X}{\partial p_2}_{\hat{\beta}}. \]

□ 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, $\beta_{\text{min}}$, to agent 2. By rejecting agent 1’s proposal and proposing $\beta_{\text{max}}$, agent 2’s expected utility is

(A11) \[ Eu_2 \geq \hat{p}_1u_2(\beta_{\text{max}}) + (1 - \hat{p}_1)u_2(d_2). \]

Going back to Round 1, the best response of agent 1 is to offer $\beta$, satisfying

(A12) \[ u_2(\beta) = \hat{p}_1u_2(\beta_{\text{max}}) + (1 - \hat{p}_1)u_2(d_2). \]

By rearranging, (A12) can be rewritten as
\( u_i(\beta) = \hat{p}_i u_i(\beta_{\text{min}}) + (1 - \hat{p}_i) u_i(d_i) \),

which by rearranging terms yields

\( u_i(\beta) - u_i(d_i) = \hat{p}_i \cdot \)

Combining (A14) and (A15) yields

\( u_i(\beta) - u_i(d_i) = \frac{u_i(\beta) - u_i(d_i)}{u_i(\beta_{\text{min}}) - u_i(d_i)} = \frac{u_i(\beta) - u_i(d_i)}{u_i(\beta_{\text{min}}) - u_i(d_i)} \).

By definition, the \( \beta^* \) which solves (A16) is the Kalai-Smorodinsky bargaining solution. Going back to the first stage, we obtain the subgame perfect equilibrium expenditures from (8).

\( \Box \quad QED \)

**Proof of Proposition 3**

Differentiating (15) with respect to \( v_i \) yields:

\[
\frac{\partial ks}{\partial v_i} = - \frac{v_i^2 F_i F_2 (4v_i + v_2)}{(v_i F_i (2v_i + v_2) + v_1^2 F_1)^2} < 0
\]

(A17)

\[
\frac{\partial ks}{\partial v_2} = \frac{v_i v_2 F_i F_2 (4v_i + v_2)}{(v_i F_i (2v_i + v_2) + v_2^2 F_1)^2} > 0
\]

\( \Box \quad QED(a) \)
Differentiating (15) with respect to $F_i$ yields,

$$\frac{\partial k_s}{\partial F_i} = \frac{v_i v_i^2 F_i (2v_i + v_2)}{(v_i F_i (2v_i + v_2) + v_i^2 F_i)^2} > 0$$

(A18)

$$\frac{\partial k_s}{\partial F_2} = -\frac{v_i v_i^2 F_i (2v_i + v_2)}{(v_i F_i (2v_i + v_2) + v_i^2 F_i)^2} < 0.$$ 

$\square$ QED(b)

PROOF OF PROPOSITION 4

Differentiating (16) with respect to $v_i$ yields,

$$\frac{\partial G_1}{\partial v_i} = \frac{v_i v_i^2 F_i (4v_i^3 v_2 F_i + 4v_i^3 F_i + 7v_i^3 F_i v_i + 4v_i^3 F_i v_2 + 2v_i^3 F_i + v_i F_i v_i^2)}{(v_i + v_2)^3 (v_i^2 F_i + 2v_i^2 F_2 + v_i F_i v_2)^2}$$

(A19)

$$\frac{\partial G_2}{\partial v_2} = \frac{v_i^2 F_i v_2 (8v_i^3 F_i + 12v_i^3 F_i v_2 - v_i^3 F_i + 6v_i F_i v_2^2 + F_i v_i^2)}{(v_i + v_2)^3 (v_i^2 F_i + 2v_i^2 F_2 + v_i F_i v_2)^2}.$$

The first derivative in (A19) is strictly positive, but the sign of the second is indeterminate.

$\square$ QED(a)

Similarly,

$$\frac{\partial G_1}{\partial v_2} = \frac{v_i^2 F_i (2v_i^3 v_2 F_i - 4v_i^3 F_i - 4v_i^3 F_i v_2 + 4v_i^3 F_i v_2 - v_i^3 F_i - v_i^2 v_2 F_i + v_i F_i v_i^2)}{(v_i + v_2)^3 (v_i^2 F_i + 2v_i^2 F_2 + v_i F_i v_2)^2}$$

(A20)

$$\frac{\partial G_2}{\partial v_1} = \frac{F_i v_i^2 (2v_i^3 v_2 F_i - 4v_i^3 F_i - 4v_i^3 F_i v_2 + 4v_i^3 F_i v_2 - v_i^3 F_i - v_i^2 v_2 F_i + v_i^2 F_i)}{(v_i + v_2)^3 (v_i^2 F_i + 2v_i^2 F_2 + v_i F_i v_2)^2}.$$

The signs of both derivatives in (A20) are indeterminate.

$\square$ QED(b)
Differentiating (16) with respect to \( F_i \) yields,

\[
\frac{\partial G_i}{\partial F_i} = -\frac{v_i v_i^2 (2v_i + v_j) F_2}{(v_i + v_j)(v_i^2 F_1 + 2v_i^2 F_2 + v_i v_j v_j)} < 0
\]

(A21)

\[
\frac{\partial G_i}{\partial F_2} = \frac{v_i v_i^2 (2v_i + v_j) F_1}{(v_i + v_j)(v_i^2 F_1 + 2v_i^2 F_2 + v_i v_j v_j)} > 0
\]

and

\[
\frac{\partial G_2}{\partial F_2} = -\frac{v_i v_i^2 (2v_i + v_j) F_1}{(v_i + v_j)(v_i^2 F_1 + 2v_i^2 F_2 + v_i v_j v_j)} < 0
\]

(A22)

\[
\frac{\partial G_2}{\partial F_i} = \frac{v_i v_i^2 (2v_i + v_j) F_2}{(v_i + v_j)(v_i^2 F_1 + 2v_i^2 F_2 + v_i v_j v_j)} > 0
\]

□ \( \text{QED(c)} \)

**Proof of Proposition 5**

If no deal is achieved and the agents turn to judicial competition in the court, their expected utilities are given by (A1), where \( z_1 = (v_i + x^*_i) \) and \( z_2 = (v_j - x^*_j) \). On the other hand, if the two agents agree on a deal, \( \beta(v) \), 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:

\[
E_u_i(w_i) = u_i - p_1 [x_i u_i' - \frac{1}{2} x_i^2 u_i''(A_i)] - p_1 [m_i u_i' - \frac{1}{2} m_i^2 u_i''] + \xi_i
\]

(A23)

\[
E_u_2(w_2) = u_2 - p_1 [x_j u_j' - \frac{1}{2} x_j^2 u_j''] + p_2 [m_j + \frac{1}{2} m_j^2 u_j''] + \xi_j
\]

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

\[
\pi_1(\beta, x) = u_i - \beta \hat{c}_i u_i' + \frac{(\beta \hat{c}_i)^2}{2} u_i'' + \xi_i
\]

(A24)

\[
\pi_2(\beta, x) = u_2 + \beta \hat{c}_2 u_j' + \frac{(\beta \hat{c}_2)^2}{2} u_2'' + \xi_2
\]
The gain of agent $i$ from a deal is defined by $g_i = \pi_i(\beta, x) - Eu_i(w_i)$. The core is non-empty if $g_i \geq 0 \ \forall i$. Thus, subtracting (A24) from (A23), dividing by $u'_i(A_i)$ and rearranging yields the following conditions:

$$
\bar{g}_i = -\frac{1}{2}(\beta \hat{c}_i)^2 R_i - \beta \hat{c}_i + \frac{1}{2} \left[ v_i(2x_i^* - v_i)R_i + 2v_i \right] p_i^* + m_i^* \left( 1 + \frac{1}{2} z_i^* R_i \right)
$$

(A25)

$$
\bar{g}_2 = -\frac{1}{2}(\beta \hat{c}_2)^2 R_2 - \beta \hat{c}_2 + \frac{1}{2} \left[ v_2(2x_2^* - v_2)R_2 + 2v_2 \right] p_2^* - m_2^* \left( 1 - \frac{1}{2} z_2^* R_2 \right)
$$

where $\bar{g}_i = g_i / u'_i(A_i)$. The roots of (A25) are:

$$
\beta_{\min,1,2} = \frac{1 \pm \sqrt{1 + \left[ v_2(2x_2^* - v_2) p_2^* + z_2^* \right]^2 R_2^2 + 2 \left( p_2^* v_1 - z_2^* \right) R_2}}{R_2 \hat{c}_2}
$$

(A26)

$$
\beta_{\max,1,2} = \frac{1 \pm \sqrt{1 + \left[ v_1(2x_1^* - v_1) p_1^* + z_1^* \right]^2 R_1^2 + 2 \left( p_1^* v_1 + z_1^* \right) R_1}}{R_1 \hat{c}_1}
$$

The relevant solutions are, of course, the non-negative values of $\beta$, where $\beta_{\max} = \max(\beta_{\max,1}, \beta_{\max,2})$ and $\beta_{\min} = \min(\beta_{\min,1}, \beta_{\min,2})$. As already mentioned above, the core is non-empty if and only if $\Delta \beta = \beta_{\max} - \beta_{\min} \geq 0$. Namely, the core is non-empty if and only if

$$
\beta_{\max} = \frac{-1 + \sqrt{1 + \left[ v_1(2x_1^* - v_1) p_1^* + z_1^* \right]^2 R_1^2 + 2 \left( p_1^* v_1 + z_1^* \right) R_1}}{R_1 \hat{c}_1}
$$

(A27)

$$
\geq \frac{1 - \sqrt{1 + \left[ v_2(2x_2^* - v_2) p_2^* + z_2^* \right]^2 R_2^2 + 2 \left( p_2^* v_1 - z_2^* \right) R_2}}{R_2 \hat{c}_2} = \beta_{\min}
$$

Clearly, (27) can hold for non-negative values of $\beta_{\min}$ and $\beta_{\min}$ if and only if $R_i > 0 \ \forall i$.

$\square \ QED$
References


