



Queen's Economics Department Working Paper No. 1001

## Mysterious Bargaining

Dan Usher  
Queen's University

Department of Economics  
Queen's University  
94 University Avenue  
Kingston, Ontario, Canada  
K7L 3N6

10-2001

# Mysterious Bargaining

Dan Usher  
Economics Department  
Queen's University  
Revised, June 16, 2008

## Abstract

We know that people strike bargains and that civilized life could not proceed otherwise. We do not know how bargains are struck. This paper is a study of propositions about bargaining: theorists' derivations of equilibrium bargains based upon a supposed sense of fairness or an imposed bargaining procedure, lawyers' notion of a transaction cost of bargaining to be minimized in the choice of laws, economists' analogy between social and technical production, the transformation of impediments to bargaining into a virtue when democracy is preserved by checks and balances between branches of government, and the common leap from the existence of equilibria within formal bargaining models to the presumption that there must exist a bargaining equilibrium - comparable to the equilibrium of prices, quantities and assignment of goods to people in competitive markets - when each bargainer acts in this own interest exclusively in the light of how others behave. The central thesis of this paper is that our models of bargaining are too far from our experience of bargaining to justify inferences about the terms of bargains or to guarantee that some bargain must be struck. Such confidence as we have in the determinacy of bargaining must derive from experience rather than from theorems. Bargaining is at once ubiquitous and mysterious.

JEL Classification C7

Keywords: bargaining, fairness, transaction cost, conflict

There would arise a general demand for a *principle of arbitration*.

And this aspiration of the commercial world would be but one breath in the universal sigh for articles of peace. For almost every species of social and political contract is affected with an indeterminateness.....an evil which is likely to be much more felt when, with the growth of intelligence and liberty, the principle of *contract* shall have replaced both the appeal to force and the acquiescence of custom.....in the general absence of a mechanism like perfect competition, the same essential indeterminateness prevails; in international, in domestic politics; between nations, classes, sexes.

The whole creation groans and yearns, desiderating a principle of arbitration, an end of strifes.

F. Y. Edgeworth

*Mathematical Psychics*, 1881, page 51

“What do Economists Economize?”. This question was the title of an old article by D.H. Robinson (1956). His answer was that economists economize love. The task of the economist as Robertson saw it was to organize as much of the world as possible under the domain of greed so that humanity’s small stock of love could be entirely directed to those aspects of life where love, and love alone, is necessary to keep society afloat. By greed, Robinson meant self-interested behaviour of the sort normally postulated by economists. By love, he meant altruism, cooperation or a willingness to act in accordance with some notion of the public good. The epitome of the economist’s invocation of greed is the general equilibrium of a competitive economy where the combined effect of entirely self-interested behaviour by each and every participant yields a outcome that no planner, however wise, powerful or benevolent, can alter to make everybody better off.

This paper is about a related question, “Is bargaining appropriately placed under the heading of greed or of love?” To place bargaining under the heading of greed is to say that there is a bargaining equilibrium when each bargainer acts in his own self-interest exclusively. To place bargaining under the heading of love is to say that bargainers are, at least sometimes, fair-minded and reasonable, prepared to make concessions on the understanding that others will do so

too, and where no calculation of rational self-interested behaviour can determine what the outcome of bargaining will be. This paper is a study of bargaining models with special reference to the economist's quest to explain explaining as much a possible of the world as the outcome of purely self-interested behaviour. It is argued that the objective has, at best, been attained imperfectly. The theories to be discussed are insightful, but their assumptions are far from the circumstances of negotiation and the motivations of the negotiators. The bargaining equivalent of perfect competition remains elusive.

As these words are being written, Sunnis and Shiites in Iraq are attempting to forge a viable government in which the interests of both are taken into account. We do not yet know whether they will be able to do so. If the argument in this paper is correct, the outcome will depend on something more than a mechanical resolution of competing powers and interests. Comparable, if less lethal, uncertainties arise every day when the Senate and the House of Representatives pass different versions of a bill and must resolve their differences if a final bill is to emerge. A central theme of this paper is that democracy involves bargaining as well as voting and that bargaining is not determinate in the way that a competitive economy is determinate as a consequence of the distribution among people of property rights.

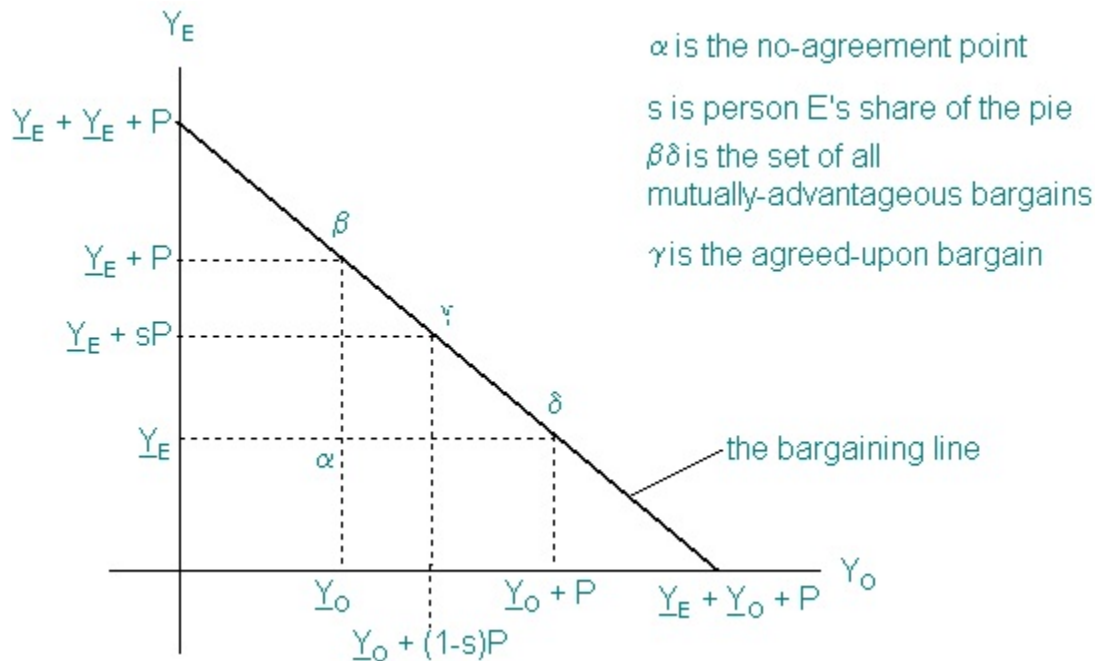
From here on, this article proceeds as follows. It begins with an illustration of the fundamental indeterminacy in the paradigmatic bargaining situation where two parties are jointly entitled to a pie if and only if they can agree between themselves how large each person's slice is to be. There follows a list of modifications in the paradigmatic bargaining situation. Next, is a brief section on the interplay between voting and bargaining, on how models of public policy formation by voting are incomplete and indeterminate without an imposed assumption about how bargains are struck. Then the innards of the bargaining process are discussed, with emphasis on the Nash bargaining solution based upon a common sense of fairness and the Staahl-Rubinstein bargaining solution based upon an imposed or agreed-upon process of bargaining. These are discussed with special reference to the grounds for confidence that bargains will in practice be struck. There follows a discussion of transaction cost and conflict. Transaction cost is an aspect of bargaining not accounted for in the paradigmatic bargain, but prominent in the economic analysis of law where the emphasis is on the choice of laws to minimize transaction cost, regardless and almost oblivious of how total transaction cost is allocated between bargainers. Models of conflict draw an implicit analogy between social interaction where outcomes depend upon the efforts of bargainers and technical interactions within a production function where output depends upon inputs of factors of production. There is finally a discussion of bargaining within democratic politics, at once absolutely necessary and potentially dangerous.

Bargaining is necessary because voting in the absence of deal-making has no determinate outcome. Bargaining is dangerous in facilitating the exploitation of minorities by majorities. Democratic government is designed in part to foster one aspect of bargaining while blocking the other. There are two appendices, one containing a formal proof of the Staahl-Rubinstein bargaining solution and the other examining three models of how bargaining fails.

### The Paradigmatic Bargain

A bargain is a division of the spoils. Two or more people are entitled to something collectively, but they cannot appropriate or make use of it until they agree about how it is to be shared. Bargaining may be over the allocation of things or of money. A bargain must make all participants better off than if no bargain were struck, but a conflict of interest remains, for a particular allocation must be chosen from the set of all possible allocations, some relatively advantageous to one party, some relatively advantageous to others.

**Figure 1: The Paradigmatic Bargain**



The paradigmatic bargain is illustrated in Figure 1. Two people bargain over the sharing of a sum of money. The money is dispersed to the bargainers if and when they come to an agreement about how much of the money each is to receive. Without such an agreement, the money is not disbursed at all. The bargain itself is a mutually agreed-upon assignment of shares.

The money at stake is  $\$P$  (mnemonic for pie). For reasons that will become evident later on, the bargainers are called E and O. The income,  $Y_E$ , of person E is shown on the vertical axis, and the income,  $Y_O$ , of person O is shown on the vertical axis. Prior to the bargain or if the bargainers fail to reach an agreement, the incomes of the bargainers are  $\underline{Y}_E$  and  $\underline{Y}_O$ , represented by the point  $\alpha$ , called the no-agreement point. A bargain is an agreed-upon assignment of a share  $s$  of the pie to person E and a share  $(1-s)$  to person O. A bargain raises person E's income from  $\underline{Y}_E$  to  $\underline{Y}_E + sP$ , and raises person O's income from  $\underline{Y}_O$  to  $\underline{Y}_O + (1-s)P$ . All feasible allocations are represented by points on the downward-sloping line cutting both axes at a distance  $\underline{Y}_E + \underline{Y}_O + P$  from the origin of the figure. Allocations for which both bargainers are better off are represented by points on the segment of that line between  $\beta$  and  $\delta$ . The chosen bargain is represented by the point  $\gamma$ . An increase in  $s$  moves  $\gamma$  toward  $\beta$  and a decrease in  $s$  moves  $\gamma$  toward  $\delta$ .

The paradigmatic bargain in Figure 1 captures central features in all bargaining, but some aspects of bargaining are not emphasized or not taken into account.

i) Surrounding Property Rights: In principle, the entire national income might be allocated among people in one vast bargain, but actual bargaining is circumscribed by property rights that all bargainers are presumed to respect. Bargains bridge the gaps in property rights, converting joint ownership of the pie over which people bargain into single ownership by each person of his agreed-upon share. Bargaining requires a prior understanding about who is (and who is not) entitled to a share of the pie and a recognition of the bargainers' untouchable property rights apart from the object of the bargain. Bargaining also requires an understanding that the deal is final, so that no bargainer can return to the table demanding a portion of the other's agreed-upon share. Otherwise, a person could appropriate another's property bit by bit, demanding something and acquiring a portion of what he demands, over and over again. Specification of property rights is the business of law and civil rights. One can imagine a specification so complete that there would never be anything to bargain about. That is more than can be expected in this imperfect world. What can be hoped for is that peoples' rights will be well enough defined that bargains are narrowly framed, with little or no dispute as to who is entitled to bargain and what the bargain is about.

ii) Utility vs. Money: The bargain in Figure 1 is about the allocation of a sum of money. Actual bargains may be about the allocation of things - such as territory in international disputes or parental rights after divorce - that are not tradable at market prices, or about the choice among alternative rules - such as whether or not one may sleep on park benches - with different impacts on the different sub-groups of the population. The outcome of bargaining and the specification of the no-agreement point may only be representable in utils rather than in money. Despite the ordinality and interpersonal incomparability of utility, models of bargaining have been designed to yield solutions where tradeoffs between people's utilities must be taken into account.

iii) Bargaining Cost: There is no such thing as bargaining cost in the paradigmatic bargain in Figure 1. The pie remains undiminished no matter how it is apportioned, and nobody devotes time and resources in the attempt to augment his share. Actual bargaining, may be different. The bargaining process may be time-consuming and expensive. The pie may shrink over time as would be the case when the total profit in a business venture diminishes steadily the longer the venture is delayed. Mere passage of time may be more or less costly to a bargainer depending on his rate of discount. One might expect that it would be relatively easy to strike small bargains and relatively difficult to strike large ones. That is not always so. People sometimes strike large bargains easily, while long and mutually-advantageous associations may be dissolved over trifles. These considerations will be discussed in connection with specific bargaining models and with the analysis of conflict and transaction later on in the paper.

iv) Bargains about Actions as well as Sharing: A bargain may be about what each participant must do as well as about the division of the spoils, but only the latter is emphasized in the paradigmatic bargain in figure 1. For example, an agreement between firms to engage in a joint venture might specify each firm's investment of time and money in the project. Recognition of efforts as well as rewards makes little difference to the bargain as long as each as all promises are sure to be kept to the letter. Problems would arise if neither firm can observe the other's investment or if such observations are unverifiable in court. Bargains that are mutually advantageous when there is full observability and verifiability of each party's actions might cease to be mutually advantageous otherwise, and forms of industrial organization that would be optimal may cease to be so.

v) Interpretation of the No-agreement Point: The no-agreement point can be looked upon as representing the incomes (or expected incomes) of the bargainers as they would be a) if the pie over which they bargain did not exist or b) if a dispute that must be resolved one way or another is in fact resolved by some costly alternative to bargaining. The first possibility is exemplified by

bargaining over the sharing of the profit from a new venture by two parties whose cooperation is required for the venture to be successful. The second possibility is exemplified by disputes between neighbours that would have to be resolved by costly litigation if no bargain is struck. The pie in this case is the combined cost saving to both parties of “bargaining in the shadow of the law”. The second possibility is also exemplified by the resort to war when nations cannot resolve their differences by peaceful negotiation. Bargaining itself may be costly, but less so than the alternative if bargaining is to be worthwhile. The bargaining models to be discussed below apply most readily to the first possibility.

vi) A Multiplicity of No-agreement Points: Regardless of how the no-agreement point is interpreted, it is implicit in the description of the paradigmatic bargain in Figure 1 that everybody’s perception of the no-agreement point is the same. In practice, that need not be so. Bargainers may disagree about the location of the no-agreement point and the size of the pie, they may have private information that they cannot or are unwilling to share, and there may be a first-strike advantage. Private settlement may be less costly than resort to the courts for the resolution of disputes, but “bargaining in the shadow of the law” may be thwarted if plaintiff and defendant are both relatively optimistic about their prospects in the event of a trial. Starting from either party’s no-agreement point, there may be a range of mutually-advantageous bargains, but there need be no commonly-recognized range when perceptions differ significantly. By contrast, there is a first strike advantage when the no-agreement point is relatively favourable to whichever party gives up on bargaining first.. In conflicts between nations, a first strike advantage may block the way to peace, even when both nations would be better off in expectation if a promise not to strike first were binding.

vii) The Number of Bargainers: Bargaining need not be restricted to two participants. A group of  $n$  people may have the collective right to a pie that cannot be touched until an agreement is reached assigning each person’s share. With three bargainers - A, B and C - a pie, P, must be divided by agreement into three shares,  $s_A$ ,  $s_B$  and  $s_C$ , where

$$s_A + s_B + s_C = 1 \quad (1)$$

With  $n$  bargainers, a bargain is an agreement on a set of shares  $\{s_1, s_2, s_3, \dots, s_n\}$  where  $s_i$  is the share of person  $i$ .

As long as agreement requires unanimity, the representation of a three-person bargain becomes a straightforward three-dimensional extension of Figure 1. The no-agreement point is



lodged in a three-dimensional space, the set of all mutually-advantageous bargains is expanded from a segment on a line to a triangle on a plane, and the bargaining problem is to agree upon some point within the triangle. With more than three bargainers, the representation of the bargaining problem becomes an n-dimensional extension of Figure 1.

viii) Coalitions: With more than two bargainers, an agreement of only a portion of the bargainers may be required to cement a deal. With n bargainers, agreement among only m of them may be sufficient as long as m is at least half the number of bargainers -  $n > m > (n + 1)/2$  - to avoid situations where two coalitions are both entitled to impose shares. There now emerge two distinct bargaining problems: who forms a coalition with whom, and how the pie is divided among the members of the coalition. The latter problem is a straightforward extension of the paradigmatic bargaining problem in figure 1. The former is endemic in the politics of majority rule voting.

If the pie consists of the entire national income and if only a minimal coalition is required to determine an allocation of the pie, then the bargaining problem is automatically transformed into the exploitation problem in majority rule voting where a bare majority of the population may grab the lion's share of the entire national income for itself. Members of the majority coalition may be bound together perhaps by some identifiable characteristic - such as race, ethnicity, geography, religion or income - or perhaps by nothing more than the prospect of gain from inclusion in a majority coalition. Such bargaining may also give rise to the classic paradox of voting where every possible allocation of the pie can be defeated in a pair-wise vote by some other allocation.

ix) A Sequence of Bargains: Bargaining today may be influenced by bargains struck yesterday or by the prospect of new bargains tomorrow. A bargain today may set a precedent, or it may affect the location of the no-agreement point in a bargain to come. Precedent is exemplified by on-going dealings between firms where shares of profit once agreed upon may persist for a long time. Influence upon the location of no-agreement points in future bargains can be important in disputes between countries over territory. The more territory a country acquires today, the stronger it may become tomorrow, and the more it may claim in the next round of negotiation.

x) Externality: As described in figure 1, the paradigmatic bargain is beneficial to the bargainers themselves, and, by implication since nothing is said about the matter, of no consequence one way or another to anybody else. Not all bargains are like that. Bargains may

convey externalities to non-participants. Bargains may therefore be classified as *benign*, *predatory* or *neutral* according to whether non-participants are helped, harmed or unaffected. Benign bargains are exemplified by a deal among scientists with different skills to invent a new product which they will patent; consumers gain from the appearance of the new product regardless of the price. Predatory bargains are exemplified by cartels; the cartel's gain is always less than its cost to the rest of the community. Neutral bargains are exemplified by wage-setting by workers and employers; outsiders' incomes may be the same regardless of whether wages are high or low.

### **Bargaining Assumptions in Models of Democratic Politics**

As a preface to the presentation of formal models of bargaining, this section reviews three instances where models of democratic government are completed by the imposition of assumptions about bargaining. In the first, voting is only determinate after the imposition of arbitrarily-chosen *rules* for splitting the difference in disagreements between houses in a bicameral legislature and between legislature and executive. In the second, shares of a pie are allocated among legislators or their constituents by an agreed-upon or imposed *process*. In the third, a judicial injunction rather than a model, the Court required parties are to strike a bargain in full confidence that the parties could actually do so.

In *Partisan Politics, Divided Government and the Economy* (Cambridge University Press, 1995), Alesina and Rosenthal consider a society where all political outcomes can be represented by points on a left-right continuum. Politics is about the choice of a number,  $x$ , on a scale from 0 at the extreme left to 1 at the extreme right. Every voter has a favourite position on that scale, and his only concern is to minimize the distance between his favourite positions and the political outcome as determined by voting and by bargaining among politicians elected to office. There are two parties, left and right, with different ideal points on the continuum. Preferences of politicians within each political party are the same. All politicians in the "left" party have the same first preference  $x_L$ , and all politicians in the "right" party have the same first preference  $x_R$ , where, of course,  $x_L < x_R$  meaning that the preferred outcome of the "right" party is to the right of the preferred outcome of the "left" party. Each party, if it could have its own way, would arrange a political outcome in accordance with its first preference.

Citizens elect legislators and a president. Voting for legislators is by proportional representation. Citizens vote for parties rather than for legislators directly, and then seats in the

legislature are allocated to the parties in accordance with the number of votes received.

The final outcome is a point on the left-right continuum, determined simultaneously by two costless bargains, one within the legislature and another between the legislature and the president. Bargaining within the legislature yields a legislative preference,  $x_Q$ , where

$$x_Q = \alpha x_L + (1-\alpha)x_R \quad (2)$$

where  $x_L$  and  $x_R$  are the first preferences of the “left” and “right” parties and where  $\alpha$  is the left party’s share of the seats in the legislature. This is a very strong assumption. It might have been assumed that the majority party gets its way completely. Instead, all legislators’ preferences are assumed to be weighted equally in a compromise where each party’s political power is proportional to its membership in the legislature. Bargaining between the legislature and the executive yields the final political outcome,  $x$ , where

$$x = \beta x_Q + (1-\beta)x_P \quad (3)$$

where  $x_P$  is the first preference of the president (which must be either  $x_L$  if the president is from the “left” party or  $x_R$  if the president is from the “right” party) and where  $\beta$ , which must be between 0 and 1, is the legislature’s bargaining power in its dealings with the president. Nothing within the model determines the magnitude of  $\beta$ . It is treated as an unexplained fact of political life.

Alesina and Rosenthal’s model yields interesting and insightful propositions about voting and the formation of public policy. It explains, for example, how rational voters’ choices between Republican and Democratic candidates for the legislature are influenced by whether the President is Republican or Democratic. More importantly from our point of view, it makes a persuasive case for the proposition that bargaining and voting are intertwined, that bargaining is an inextricable component of majority-rule voting, that there can be no electoral equilibrium apart from a capacity for bargaining among our politicians, and that, without bargaining, democratic government would be impossible.

For both of these objectives, a simple and perhaps unrealistic model of bargaining is quite sufficient. Bargains are struck in accordance with the parties’ bargaining power. Within the legislature, each party’s bargaining power is assumed to be proportional to its number of seats. Between legislature and president, bargaining power depends on a parameter pulled out of thin air. Though entirely ad hoc and ungrounded in any persuasive explanation of how rational and

self-interested people come to agree, these assumptions are reasonable enough as long as we do not allow ourselves, on the strength of the theory, to suppose that bargaining is more predictable and determinate than is really the case. Democratic government needs bargaining, but a need for bargaining does not render bargaining determinate.

Note finally that the Alesina and Rosenthal model shrinks all politics to a single dimension. A thousand dimensions of public choice - in defense, health care, education, tax policy, the redistribution of income and so on, each with a thousand options to choose from - are condensed onto a single value of  $x$  on a left-right scale. The procedure may well be justified on the grounds that, even then, voting and bargaining are intertwined. If bargaining must supplement voting in that simple environment, then it must surely do so in more realistic circumstances. We return to this question in the last section of the paper.

A very different picture of legislative decision-making is drawn by Baron and Frerejohn in "Bargaining in Legislatures" (1989). The models differ in two respects that have a particular bearing on the concerns in the present paper. First, the postulated task of politics, represented by Alesina and Rosenthal by the choice of one number from a continuum, is represented by Baron and Frerejohn as the allocation of shares of total expenditure to each and every legislator or his constituents. Second, the mechanism by which determinacy is acquired, represented by Alesina and Rosenthal by an imposed sharing rule, is represented by Baron and Frerejohn by a voting procedure modeled as a non-cooperative game with an imposed sequence of bills and amendments, and where entitlement to propose bills and amendments is assigned by lot. The striking feature of the Baron and Frerejohn model is its solution to a problem that has vexed theorists of democracy since the very origins of political speculation in ancient Greece: how to divide a pie by majority rule voting, or, more generally, whether voting can be made to yield an equilibrium distribution without prior assignment of property rights.

The task of the legislature in the Baron and Frerejohn model is to allocate a nation-wide benefit - that could be the national income but might be something less all-encompassing such as seats in the cabinet or local expenditures - among the legislators or their constituents. A bill is passed or defeated by majority rule voting. A voting equilibrium has proved elusive in this context because the paradox of voting is particularly troublesome under majority rule. Any bill with some given allocation of total expenditure can always be defeated by some other bill with a different allocation. The sought-for equilibrium is created by strong procedural rules. The right to propose a bill is assigned by a flip of a coin. Subsidiary rules guide amendments to bills, the introduction of alternative bills and the introduction of new bills when a preceding bill has been

voted down. The paper is largely about the consequences of different rules.

These two political models can be thought of as applications of the two principal bargaining solutions to be discussed below, the Nash bargaining solution and the Staahl-Rubinstein bargaining solution. How much faith can reasonably be placed in these models, and whether they are really models of bargaining at all, will be considered after the two solutions have been described. More will be said about the Baron and Frerejohn model within the discussion of the Staahl-Rubinstein bargaining solution.

Confidence in the determinacy of bargaining is illustrated in a recent case before the Canadian Supreme Court. Over the last half century, the dominant political issue in Canada has been the threat of the separation of Quebec. A significant minority in the province of Quebec would like to transform Quebec into an independent country with French as the only official language, but two referendums in the province have failed to produce a majority for secession. Canadians outside of Quebec would have to take the prospect of separation very seriously if a majority of the people of Quebec voted for separation in another referendum, but there is no consensus whatsoever about how large a majority would be required for Quebec to secede or about the exact terms of secession.

To clarify the matter, the Federal government asked the Canadian Supreme Court for a judgment on the several questions, the principal question being: “Under the Constitution of Canada, can the National Assembly, legislature or government of Quebec effect the secession of Quebec from Canada unilaterally?”. The Court’s answer (in *Reference re: Secession of Quebec*, 1998) was that “The secession of a province from Canada must be considered, in legal terms, to require an amendment to the constitution which perforce requires *negotiation*” and the Court went on to say that, “constitutional rules themselves are amenable to amendment, but only through a process of *negotiation* which ensures that there is an opportunity for the constitutionally defined rights of all parties to be respected and reconciled.” (Italics added). How the negotiation is to proceed, who is to be a party to the negotiation, what to do if one party is intransigent and how to recognize intransigence are questions the Court did not even consider.

The judgment in this case exemplifies the commonly-held view that any disagreement can be resolved if the right sort of people are put in a room and told to get on with solving it. What this judgment has in common with the literature discussed above is a presumption that bargains can and will be struck. In this judgment, the emergence of a deal is treated without much explanation as a fact of life. Confidence in deal-making is relied upon as a substitute for explicit

rules that might have been promulgated instead. Such confidence is not entirely misplaced. Negotiation does yield a determinate outcome much of the time. Yet negotiation does sometimes break down into non-agreement or outright violence. Together these models of politics and the Supreme court case suggest three questions about bargaining.

i) whether we have a plausible model of bargaining corresponding closely enough to bargaining as we know it.

ii) whether (and if so when and to what extent) we can rely on bargaining to solve actual disputes, and

iii) whether there exists an electoral equilibrium, comparable to the general equilibrium in a competitive economy, with political outcomes entirely dependent upon the preferences of voters and upon electoral procedures, but without the need for bargaining to fill in the gaps where politics would otherwise be indeterminate.

The first of these questions is the principal subject of the rest of this paper. The second will be discussed superficially and with a sense of relief that bargaining seems to work better in practice than in theory. The third will be discussed in the final section of the paper. Our very tentative answers to these questions are: no, yes and no.

## **Bargaining Models**

Implicitly or explicitly, confidence in the determinacy of bargaining - as exemplified by the judgment of the Canadian Supreme Court - may be fortified by the existence of equilibrium in explicit models of bargaining. Three such models will be examined briefly in turn, models based upon a common sense of fairness, a sequence of concessions and a fixed bargaining procedure.<sup>1</sup> For each, it will be argued that the model, though interesting and instructive, does not in the end supply the confidence we seek.

### A) A Shared Sense of What is Fair.

---

<sup>1</sup>For an excellent survey of contemporary bargaining theory, see Abhinay Muthoo, *Bargaining Theory with Applications*, Cambridge University Press, 1999. For a useful collection of earlier literature including papers by Bishop, Cross, Nash and Schelling to be cited below, see Oran R. Young, *Bargaining: Formal Theories of Negotiation*, University of Illinois Press, 1975.

To say that bargains are struck in accordance with a shared sense of what is fair implies that people i) have a common understanding of what is meant by a fair bargain and ii) strike bargains accordingly. These requirements will be discussed in turn.

Begin with the working assumption (to be examined presently) that, for the simple paradigmatic bargain over the apportionment of a fixed sum of money between two people, a “fair” bargain is a fifty-fifty split. Even so, the notion of fairness would be vague and perhaps of little use unless it could be extended in some natural way from bargaining over dollars to bargaining over the apportionment of things - like family heirlooms or authority over children in the event of divorce - for which market prices are irrelevant. There may be situations where people’s gains from a bargain can only be represented as utilities. Utility supplants money as the object of bargaining if a sense of fairness leads bargainers to take account of disparities in their incomes.

The difficulty in bargaining about utilities is that utility is ordinal, defined up to a linear transformation, but not comparable from one person to the next. When bargaining is over a sum of money, P, a bargain may be deemed fair when it is a fifty-fifty split, when  $\Delta Y_E = \Delta Y_O$  where  $\Delta Y_E = sP$  and  $\Delta Y_O = (1-s)P$  so that person E’s share, s, of the pie is exactly one half. When bargaining is over the assignment of utilities, we would like to equate  $\Delta u^E$  and  $\Delta u^O$  where  $\Delta u^E$  is the impact of the bargain on the utility,  $u^E(Y_E)$ , of person E and  $\Delta u^O$  is the impact of the bargain on the utility,  $u^O(Y_O)$ , of person O. We would like to define a fair allocation of the pie as one for which  $\Delta u^E = \Delta u^O$

where 
$$\Delta u^E = [u^E(\underline{Y}_E + sP) - u^E(\underline{Y}_E)] \tag{4}$$

$$\Delta u^O = [u^O(\underline{Y}_O + (1 - s)P) - u^O(\underline{Y}_O)] \tag{5}$$

and where  $\underline{Y}_E$  and  $\underline{Y}_O$  are incomes of persons E and O as they were prior to the bargain. That is not feasible because utilities are incommensurate.

There is a way around this difficulty, leading to a rule called the Nash bargaining solution<sup>2</sup>. When  $\Delta Y_E$  is equated to  $\Delta Y_O$ , the product  $\Delta Y_E \Delta Y_O$  is automatically maximized subject to the constraint that  $\Delta Y_E + \Delta Y_O = P$ , and the product  $s(1 - s)$  is maximized as well. This property of fair allocation can be extended from income to utility even though the simple equality

---

<sup>2</sup>John Nash, “The Bargaining Problem”, *Econometrica*, 1950, 155-62.

of shares cannot. The Nash bargaining solution is to choose  $s$  to maximize the product

$$\Delta u^E \Delta u^O = [u^E(\underline{Y}_E + s P) - u^E(\underline{Y}_E)] [u^O(\underline{Y}_O + (1 - s) P) - u^O(\underline{Y}_O)] \quad (6)$$

This maximand is well-defined despite the fact that the utilities themselves are incommensurate. The value of  $s$  that maximizes this expression is unchanged by a linear transformation of either utility of income function. This procedure boils down to a fifty-fifty split of the pie in the special case where both bargainers' utility functions are linear in income.

The Nash bargaining solution is a theorem derived, like all theorems, from a set of axioms. In high school geometry, axioms were presented as "self-evident truths". In the social sciences, we have no such luxury. Our axioms must be substantive, and theorems can be no stronger than the axioms from which they are derived. Among Nash's axioms is this: "If  $S$  is symmetric and  $u_1$  and  $u_2$  display this then  $c(S)$  must lie on the line  $u_1 = u_2$  ." (axiom # 8 ) where  $S$  is the set of all possible outcomes and  $c(S)$  is the set of all possible fair outcomes. To postulate that is to mandate a fifty-fifty split of the pie in the simple paradigmatic bargain when utility is proportional to income. Nash's program is to extend the notion of fair allocation from money to utils, but not to justify the fair allocation itself. Bargainers who for one reason or another refuse to accept a fifty-fifty split as a fair allocation of the pie - bargainers who for one reason or another do not conform to axiom #8 - would have no difficulty in refusing to accept the shares assigned in the Nash bargaining solution.

The Nash bargaining solution assigns the larger slice to the person with the smaller income if bargainers' utility of income functions are the same and if the common utility of income function is concave, but the Nash bargaining solution is not always redistributive.<sup>3</sup> Also, to agree on an allocation of the pie in accordance with the Nash bargaining solution, bargainers would need to know one another's utility of income functions. Without such knowledge, a resort to a fifty-fifty split might be the only feasible procedure.

---

<sup>3</sup>Suppose i) that persons E and O are bargaining over the allocation of  $\$P$ , ii) that their utility functions are  $U_E = (Y_E)^{1/2}$  and  $U_O = (Y_O)$  where  $Y_E$  and  $Y_O$  are their incomes, and iii) that, to keep the arithmetic simple, the initial income of person E is 0. On these assumptions, the value of  $\Delta u^E \Delta u^O$  in equation (13) reduces to  $[(s P)^{1/2}][(1 - s)P]$  which is maximized when  $s = 1/3$ . The fair share of person E is either  $1/2$  or  $1/3$  depending on whether fairness is defined with reference to dollars or to utils. This is true despite the fact that person E could well be very much less well off than person O.



For bargains over the allocation of money, a fifty-fifty split might be acceptable not just because it is fair, but because it is a focal point, the only readily-recognizable rule. If a fifty-fifty split were customary, then all bargainers would know exactly what to do. A general convention that people in a dispute ought to split the difference evenly would be relatively easy for everybody to follow. Other conventions would be difficult to maintain. For instance, a convention supplying two-thirds of the pie to person E and the remaining third to person O is meaningless without a prior understanding about who is to play the role of person E and who is to play the role of person O. Perhaps such a convention might be founded on class structure, but that would require a substantial modification of the assumptions about the paradigmatic bargain in Figure 1.

A convention to divide the pie equally might be enforced by an understanding that anybody who deviates from the convention will be punished not by the state as one would be punished for robbery, but by his fellow citizens who would refuse to deal with him again.<sup>4</sup> On the other hand, a convention enforced by sanction is the antithesis of what most people would think of as a bargain. There is a sense in which a law punishing people for robbery can be seen as a nation-wide convention not to steal, but such a convention is not a bargain in the sense described in Figure 1. Nor is a convention punishing people by ostracism for refusing to accept an equal division of the pie. Bargaining only takes place in the absence of coercion.

There are two incentive problems. The first has to do with the creation of disputes. If disputes are to be resolved by a fifty-fifty split of the pie, it becomes in everybody's interest to create disputes at other people's expense. I assert that a third of what you claim to be your land is really mine, and we agree that I get one sixth. A state of affairs where what anybody chooses to call a dispute is resolved by a fifty-fifty split is untenable in the long run. The moral of the story, as already mentioned in the discussion of the paradigmatic bargain, is that the notion of a fair bargain is meaningless except in a context of well-established property rights. A bargain to split what is initially jointly-owned or collective property may be fair. A bargain to split what is initially your exclusive property is automatically unfair. Unfortunately, the line between individually-owned and collectively-owned property is not always as sharp as we would like. Disputes over the redistribution of income can be framed as being between people who see the present distribution of property as inviolate and people who see the present distribution of property as the residue of ancient theft or who look upon the entire national income as collective

---

<sup>4</sup>Such a mechanism is analyzed by Robert Axelrod in *The Evolution of Cooperation*, Basic Books, 1984.

property to be allocated in the service of the common good.

The other problem has to do with bargainers' motivation. When we speak of "explaining" bargains, what we really have in mind is a bargaining outcome comparable to the outcome in general equilibrium in competitive markets where people act in their own interest exclusively, responding to market-determined prices but not to one another. Nobody in perfect competition is "fair". Every person is unremittingly greedy, cooperating with his fellow man if and only if it is personally advantageous to do so. To accept a fair bargain because it is fair is a different order of behaviour altogether. It is the incorporation of uncoerced good-will into the core of the market. It is an admission of failure in the great project of explaining outcomes in the economy by self-interest alone. Indeed, if people could be relied upon to be "fair" voluntarily, the market itself might prove unnecessary except perhaps to identify each person's appropriate behaviour in any given situation. It is precisely our failure to subsume bargaining under the heading of self-interest that justifies its description as mysterious.

Bargaining is, almost by definition, indeterminate. A bargain is the resolution of a dispute. If bargainers can be relied upon to respect a notion of fairness, to agree on a fifty-fifty split of the pie, or to accept shares mandated by the Nash bargaining solution, then bargaining is just playacting, for there is no real dispute and nothing left to bargain about.

### B) A Sequence of Concessions

In the early nineteen-thirties, J. R. Hicks and F. Zeuthen developed models of bargaining between companies and unions when both have a degree of monopoly power.<sup>5</sup> Bargaining in these models consisted of a series of concessions dependent on the harm to each party from a failure to agree and upon each party's expectation that the other would concede instead. Hicks draws what he calls an "employer's concession curve" and a "union's resistance curve". The crossing of these curves identifies the agreed-upon wage. In the Zeuthen model, the failure of employees and owners to agree leads to "conflict", the exact meaning of which is not spelled out in detail. Zeuthen's principal assumption is that each bargainer's concession to the other is proportional to his expected harm from conflict as it would be if antagonism between the bargainers rises to the point where the entire pie is wasted through a failure to agree.

---

<sup>5</sup>See John R Hicks, *The Theory of Wages*, 1932, chapter VII and F. Zeuthen, *Problems of Monopoly and Economic Welfare*, 1930, chapter IV. See also John Harsanyi, "Approaches to the Bargaining Problem Before and After the Theory of Games: A Critical Discussion of Zeuthen's Hick's and Nash's Theories", *Econometrica*, 1956, 144-57.

Both models allocate the surplus in proportion to harms that do not actually occur because they are averted by timely concessions. Strikes in Hick's model are imagined strikes. Conflict in Zeuthen's model is imagined conflict. Neither model contains an explanation of when, if at all, bargaining breaks down and the unfortunate alternative to agreement is realized. Nor is it explained how bargaining in the midst of a strike or bargaining in the midst of conflict differs from bargaining in anticipation of these events. Neither party is bloody-minded, insisting on favourable terms come hell or high water. This consideration is especially problematic because, if one bargainer is really and truly adamant, it is usually in the interest of the other party to back down. More will be said about this presently. Bargaining is made determinate within these models, but only by ignoring essential features of the world where bargains are struck. There is no satisfactory explanation of the timing and the magnitude of concessions, and no allowance for the possibility that the final agreement is conditioned by the history of bidding as well as by the initial values of the bargainers' harms from conflict.

Genuine concessions are modeled by Cross in "A Theory of the Bargaining Process" (*American Economic Review*, 1965). Both parties' concessions are rendered determinate by the principle that delay is costly so that, if you do not concede quickly, then I must. At least three kinds of harm might be identified. a) loss by both parties of what would otherwise be their shares of the pie, b) delay which may be more costly for one party than for another depending on their rates of discount, and c) actual harm inflicted as when a labour union goes on strike or when the firm locks out its employees. Cross attempts to derive the sequence of concessions as the outcome of rational, self-interested behaviour, transporting this aspect of bargaining from the domain of psychology - where people may act stubbornly, vindictively or irrationally - into the domain of economics - where each person does what is best for himself in the light of his best guess of what others will do. Yet the model contains no persuasive explanation of why bargainers do not proceed to the ultimate deal all at once if the ultimate deal is predictable from the initial conditions, as Cross assumes it to be.

In the light of subsequent literature, these models would seem to be open to the objection that the bargainers are neither entirely fair-minded, as in the Nash bargaining solution, nor entirely self-interested in any rational and calculating way. Bargainers are seen as making concessions, but their concessions do not arise naturally from the maximization of an objective function in response to given constraints. It is difficult to decide how much weight to attach to this objection. Want of strict rationality may account for the eclipse of these models in economic literature, but, in their defense, it may be argued that bargaining is not really as rational a process as more recent models would suggest.

### C) Mutually-agreed upon Procedures

A bargaining solution may arise not just from a common sense of fairness or as the outcome of a sequence of concessions, but as the outcome of an alternating sequence of offers by one party to be accepted or rejected by the other.

Begin with the simplest possible case. Persons E and O are bargaining over the allocation of a pie that emerges just for an instant and disappears if it is not shared at once. The pie appears for just long enough for one person say, "I offer you such-and-such a share and I will take the rest.", and for the other person to reply either "yes" or "no". No other speech is admitted. Suppose, no matter why, it is person E who is entitled to make the offer. If the person O's response is "yes", the pie is shared accordingly. If person O's response is "no", the pie vanishes and nobody gets anything.

It is obvious what happens. Person E offers person O a penny, keeping all the rest of the pie for himself. Recognizing that a penny is better than nothing, person O accepts the offer, and the pie is allocated accordingly. If the original pie was \$100, person O ends up with one penny and person E ends up with \$99.99. In effect, the person entitled to make the take-it-or-leave-it offer gets to keep the entire pie.

If that seems a bit harsh, and much too far from anything we would ordinarily call bargaining, we can even out the allocation by allowing the pie to disappear over two time periods rather than just one. Suppose i) that the pie appears at sunrise of day 1 and disappears in two stages, half at sunset on day 1 and the other half at sunset on day 2, ii) that offers to share of the pie (or what remains of it when the offer is made) are at noon each day, by person O on day 1 and, if person O's offer is rejected, by person E on day 2, iii) that every offer is an assignment of shares, iv) the person to whom the offer is made must accept or reject the offer immediately, and v) nothing else may be said by either person.

Again it is obvious what must happen. At noon on day 1, person O offers person E half the pie, and person E accepts. Why? If person E rejected person O's offer on day 1, the most person E could expect would be half the original pie because nothing more would be left on day 2 when it is person E's turn to make an offer. Except for the switch in roles and the size of the pie, both parties find themselves in the same situation at noon on day 2 as in the one period take-it-or-leave-it bargain, and they act accordingly. Since person E can assure himself half of the pie (less a penny) by waiting until his turn to make an offer comes round, he would never accept less than

half of the pie in any offer from person O on day 1, and person O has no incentive to offer more.

There is, of course, nothing inevitable about the equal sharing of the pie or about the restriction of bargaining to two periods. The pie may disappear over any number of days, and the disappearances each day need not be the same. Suppose the pie diminishes over four days,  $1/10$  at sunset on day 1,  $2/10$  at sunset on day 2,  $3/10$  at sunset on day 3 and the remaining  $4/10$  at sunset on day 4. If so, then at noon on day 1, person O offers  $2/5$  of the pie [ $1/10$  plus  $3/10$ ] to person E, leaving the remaining  $3/5$  of the pie [ $2/10$  plus  $4/10$ ] for himself, and person E accepts. Person O would accept nothing less. Person E need offer nothing more. The logic of this allocations is backward induction.

Begin by supposing that no deal has been struck by noon on day 4, the last day when any of the pie remains. Since the day 4 is an even day, it is person E's turn to make an offer. As in the one period case, person E would offers just a penny to person O, keeping the remainder - which is only  $4/10$  of the original pie - for himself. Now step backward from the day 4 to day 3 when  $7/10$  of the pie remains and when person O is entitled to make the offer. Person O cannot expect person E to accept anything less than  $4/10$  of the pie, for that is what person E could acquire by waiting for his turn to make an offer, but person O need not offer more. Person O offers  $4/10$  of the pie to person E, keeping the remaining  $3/10$  of the pie for himself. Step backward one more day to day 2 when  $9/10$  of the pie remains and person E is entitled to make the offer. Person E cannot expect person O to accept anything less than  $3/10$  of the pie, for that is what person O could acquire by waiting, but person E need not offer more. Person E offers  $3/10$  of the pie to person O, and keeps the remaining  $6/10$  of the pie for himself. Finally, person O is entitled to make the offer in day 1 before any of the pie has vanished. Person O cannot expect person E to accept anything less than  $6/10$  of the pie which is what person E could acquire by waiting, but person O need not offer more. Person O offers  $6/10$  of the pie to person E, keeping the remaining  $4/10$  of the pie for himself.

When the pie diminishes over a number of days, each bargainer captures the sum of the diminutions of the pie on the evenings of all the days when he is entitled to make the offer. Person E obtains a share equal to the sum of the diminutions in all even-numbered days, and person O obtains the sum of the diminutions on all odd-numbered days.

Generalizing slightly, when time is graduated in years rather than days, when a pie of size P diminishes spontaneously over the course of n years, when person E is entitled to make an offer in all even years and when person O is entitled to make an offer in all odd years, then an

acceptable offer would be made in the very first year of bargaining with a slice  $P_E$  to person E and a slice  $P_O$  to person O where

$$P_E = \sum_{t \text{ even}} p_t \quad \text{and} \quad P_O = \sum_{t \text{ odd}} p_t \quad \text{and} \quad P_E + P_O = P \quad (7)$$

where  $p_t$  is the size of the slice of the pie that disappears on the  $t^{\text{th}}$  year and where  $P_E$  and  $P_O$  must sum to  $P$ . This is an equilibrium bargain because it is in the interest for each person to accept a share of the pie equal to the sum of the disappearances on all of the times when he would be entitled to make the take-it-or-leave-it offer. None of the pie is lost in the process of bargaining because the bargain is struck in the of the first year before any of the pie has disappeared.

An interesting extension of this model replaces disappearance by discounting. Suppose i) that the pie over which people bargain lasts undiminished forever, or would do so unless a bargain is struck, but ii) that the bargainers value present income over future income, each in accordance with his own rate of discount, and iii) each bargainer is entitled to make offers in alternative years. It can be shown that, once again, a bargain is struck as soon as bargaining begins, but that now the equilibrium shares of the pie are inversely proportional to the bargainers' discount rates. Specifically,

$$P_E = r_O / (r_E + r_O) \quad \text{and} \quad P_O = r_E / (r_E + r_O) \quad (8)$$

where  $r_E$  and  $r_O$  are the discount rates of persons E and O. If my discount rate is high, my share of the pie is correspondingly low. To have a high discount rate is analogous to sacrificing a large share of the pie if one refuses the other bargainer's offer, so that one's equilibrium share of the pie is correspondingly reduced. Equation (14) is called the Staahl-Rubinstein bargaining theorem.<sup>6</sup> A simple, and hopefully intuitive, proof of the theorem is presented in Appendix 1.

To induce a deal as soon as bargaining begins, the present value of the pie must be made to shrink when the deal is delayed. Two equally effective processes have been discussed: physical

---

<sup>6</sup>The earliest bargaining solution of this type was proposed by Ingolf Staahl in *Bargaining Theory*, 1972. A more tractable form of the model was proposed by Ariel Rubinstein in "Perfect Equilibrium in a Bargaining Model", *Econometrica*, 1982, 97-109. For a short and simple presentation of the Staahl-Rubinstein bargaining model, see J. Sutton, "Non-Cooperative Bargaining Theory: An Introduction", *Review of Economic Studies*, 1986, 709-24, and, for a thorough treatment of the subject, see Martin J. Osborne and Ariel Rubinstein, *Bargaining and Markets*, Academic Press, 1990.

contraction over time, and reduction in present value due to discounting. The processes are analytically similar, but the latter has the distinct advantage that it is based upon the characteristics of bargainers (their rates of discount) rather than upon the imposed conditions in which bargaining takes place. Rates of discount are attached to people. Physical shrinkage of the pie is not.

The explanation based upon bargainers' discount rates has serious problems of its own. As shown in the appendix, equation (14) is strictly valid as a bargaining equilibrium if and only if the bargainers are immortal and the pie lasts forever in the event that no bargain is struck. The Staahl-Rubinstein bargaining solution requires that bargainers E and O must be prepared to carry on making offer and counter-offer in the year 3008 if no agreement had been reached before that time. Without that assumption, equation (14) is just an approximation, though it becomes more and more accurate the longer the time before the pie finally disintegrates.

Nothing so extreme is required for the explanation based on physical diminution of the pie. Bargaining opportunities arise from time to time, and then disappear. In business and politics, it is rare for today's opportunities to remain available in five years time, and it is not unreasonable to suppose that physical shrinkage of the pie might have more impact on the outcome of bargains than bargainers' rates of discount.

The combined effect of shrinkage of the pie and discounting by bargainers is easy enough to compute. Consider a "bargaining" regime with alternating offers up to some odd year T where T is assumed to be odd, but with an arbitrarily apportioned in the year T - yielding a share S to person E and a share (1 - S) to person O - if no deal has been struck voluntarily by that time. It is shown in the appendix that a bargain is struck immediately in the year 1, yielding a share  $s(E, 1)$  to person E where

$$s(E, 1) = (1 - x)[r_O / \{r_E + r_O + r_E r_O\}] + xS \quad (9)$$

where

$$x = (\delta_O \delta_E)^{(T-1)/2} \quad (10)$$

and where  $\delta_E$  and  $\delta_O$  are discount factors,  $\delta_E = 1/(1+r_E)$  and  $\delta_O = 1/(1+r_O)$ . Person E's share is a weighted average of his shares as they would be if bargainers did not discount future income and as they would be if with discounting but the pie lasts undiminished forever unless a deal is struck. It follows immediately from equation (9) that person E's share varies steadily from S when  $T = 1$  to  $r_O / \{r_E + r_O + r_E r_O\}$  which is the Staahl-Rubinstein bargaining solution when T

approaches infinity.

For example, if the pie must be shared equally ( $S = \frac{1}{2}$ ) unless a deal is struck before the seventh year of bargaining (i.e.  $T = 7$ ), if person E's discount rate is 9% ( $r_E = .09$ ) and person O's discount rate is 1% ( $r_O = .01$ ), then pie is divided voluntarily in the year 1 with a share of 45% to person E and the remaining 55% to person O. Person O would offer person E a share of 45% and person E would accept. By contrast, if the pie would remain undiminished forever, person E's share would be only 10% in accordance with the pure Staahl- Rubinstein bargaining solution in equation (14) or in equation (15) as  $T$  approaches infinity. Person E's share of 45% in this deal is much closer to the 50% he would obtain if the bargainers did not discount future income than the 10% he would obtain if bargainers discounted future income but the pie lasted forever. One example proves nothing, but it does highlight the significance of the assumption in the Staahl - Rubinstein model that the pie would remain undiminished, and ready to be allocated, forever unless some bargain is struck.

An equilibrium bargain can be computed for any arbitrarily-assigned pattern of disappearance of the pie or apportionment of the pie between bargainers at an assigned time some years ahead. What cannot be altered if there is to be an equilibrium at all is the imposed sequence of offers to be accepted or rejected by the other party. Nothing works unless bargainers respect the required sequence of speech. The procedure itself may be agreed-upon by bargainers or externally-imposed. If the procedure originates from a prior agreement between the bargainers, and in so far as its outcome may be predicted from the characteristics of the bargainers and of the procedure itself, then the outcome of bargaining is foretold in the chosen procedure, and there is really nothing to bargain about. What we are calling bargaining would, once again, be play-acting, with no real give and take between bargainers, and there would be is some question about whether what is being called bargaining theory is really about bargaining at all. Nor would the procedure correspond to what we normally think of as a bargain if it were externally imposed, for, once again, the outcome of bargaining would be pre-determined before the bargainers ever meet.

There is even some question about whether the process leading up to the Staahl-Rubinstein bargaining solution can properly be described as bargaining. The process has some of the features we normally associate with bargaining, but it differs in several significant respects. It is a sequence of speech, but more ordered and confined than ordinary speech. There is no counterpart in actual bargaining to the implicit gags required to yield the Staahl-Rubinstein solution. Neither "sub-game perfection" nor the postulated infinite life of the unallocated pie are



to be found in bargaining as we know it. A process is not a bargain because we choose to call it so. But if the Staahl-Rubinstein process is not bargaining, what can it be?

The question is highlighted by the Baron and Frerejohn article already discussed briefly in our examination of bargaining assumption in models of democratic politics. Called “Bargaining in Legislatures”, it is an extension of the Staahl-Rubinstein process to the proposal, amendment, passage or rejection of bills in the legislature. Two bargainers are replaced by many legislators. Splitting of the pie is replaced by its allocation among legislators and their constituents. Alternation of offers is replaced by random selection of proposers of bills and amendments. Agreement between bargainers is replaced by majority rule (with certain interesting exceptions). Common to both models is the critical role of discounting in the choice between what can be obtained today (by accepting an offer or by voting for a bill) and what one can expect to obtain tomorrow. Common to both models are sub-game perfection and strictly self-interested behaviour. Nobody today can promise to act tomorrow - to vote or to refuse an offer - in a way that would not be advantageous at that time. Members of a majority coalition cannot promise one another to vote as agreed upon by the coalition as a whole, refusing any bribe, no matter how attractive, to act otherwise.

Remarkable as it is that any procedure can be found to generate an equilibrium sharing of a pie among many people, there remains some question about what exactly the pie is presumed to be. This presents no difficulty in ordinary bargaining - as when businessmen share the profit from a joint venture - but it is problematic for legislation. One’s first thought is to identify the pie with the entire national income. Legislators could be imagined as voting on an allocation of the entire national income among themselves and their constituents. But that cannot be right. Legislation may influence the distribution of income, for example through the progressivity of the income tax, but the entire national income cannot be assigned democratically with no reference whatsoever to the prior assignment among people of property rights. Looked upon as a political allocation of the national income, the model would seem to be exchanging people’s endowments of labour and capital for a sequence of lotteries determining who is entitled to propose bills and amendments, with a fresh assignment of incomes every time the legislature convenes. The consequences of such an assignment of incomes would be horrendous. Incentives to work and save would be all but eliminated by the severing of the connection between one’s contribution to and benefit from the national income as a whole. Replacement of individual effort with mere chance in the determination of each person’s income would quickly destroy the consensus in society to abide by the outcome of the vote. People would never tolerate the public choice of who is to be rich and who is to be poor based upon nothing more substantial than the flip of a coin. A

democracy that tried to allocate the entire national income politically would not remain a democracy for long.

The model might be more appropriately applied to such matters as the allocation of seats in the cabinet or the distribution of pork, of add-ons to bills of special projects for the benefit of one locality or to one interest group exclusively, projects that must inevitably be politically assigned.

A terminological question has a special bearing on the subject of this paper. Is Baron and Frerejohn's analysis best described as a model of bargaining or as a model of parliamentary procedure? The process of proposing, amending and voting on bills in the Baron and Frerejohn paper seems a great deal more like what we ordinarily think of as parliamentary procedure. Senator McCain and Senator Feingold knocked their heads together to forge the McCain-Feingold Campaign Reform Bill which was then voted upon and possibly amended in the Senate. We would normally think of the original formation of the bill as bargaining, and of its passage as parliamentary procedure. To extend the term bargaining to voting about the passage of the bill is an unusual usage of words. One might by analogy think of the legislature as a great bargaining machine for the passage of bills. Such usage of language is not wrong per se, but there are circumstances where it can be obfuscating. Were that usage adopted, it would sometimes be necessary to distinguish between bargaining as the search for an agreement by compromise and bargaining as the outcome of universally self-interested behaviour within well-defined rules, such as the rules of parliamentary procedure.<sup>7</sup>

Which brings us back to the Staahl-Rubinstein bargaining solution. It is at least arguable that the model from which it is derived is more like a peculiar variant of parliamentary procedure than like bargaining as most people understand the term. A bargaining solution has been constructed by redefinition. Clean and interesting as it may be, the Staahl-Rubinstein solution

---

<sup>7</sup>Whether what goes on in legislatures - what we normally think of as parliamentary procedure - can rightly be called bargaining is not a question that can be answered definitively, for there are no God-given definitions of words. It may be useful in some contexts to refer to "an animal with fins and scales and that swims in the sea" as a type of cat, but, if so, it would be useful in other contexts to distinguish between land cats and sea cats. A sequence of bills in the legislature as chosen by an agenda-setter may, by analogy, be designated as a bargain, but, if so, a distinction would have to be drawn between legislative bargaining and person-to-person bargaining, and the title of this article would have to become "Mysterious Person-to-person Bargaining". A stronger case than I have so far seen would be required to justify this peculiar usage of words.

cannot be trusted as a foundation for a belief - such as that of the Canadian Supreme Court in the separation case - that differences among people or groups of people will be resolved peacefully, and that amicable deals will be struck. Restrictions on speech in the Staahl-Rubinstein bargaining solution are simply too stringent. To say of a model that it is a model of bargaining, does not make it so.

#### D) Conversations and Threats

Whatever else it may be, bargaining is a conversation. Bargainers talk to one another, make offers, tell stories about why their offers ought to be accepted, appeal to one another's sense of fairness, reject offers, and so on. The model of bargaining as fair division ignores this aspect of bargaining altogether, for outcomes emerge directly from the initial conditions with no room for speech at all. The Staahl-Rubinstein solution acquires a certain plausibility from its resemblance to conversation, but the conversation is artificial in two respects: It is artificial because, as discussed above, the equilibrium deal is struck before any actual conversation takes place. It is also artificial because bargainers are severely restricted in what they can say and when they can say it. Speech is limited to three, and only three, utterances: "I offer ...", "Yes" and "No", with a switch in the bargainers' roles at each stage of the conversation, until a deal is struck. There is a prescribed spacing between utterances and a prescribed order of speech, neither of which are to be found in actual conversation or negotiation between firms, between employer and employees or between the buyer and seller of a house. Actual bargaining is far less orderly and coherent than the model would suggest. There is no fixed order of speech. People interrupt one another. People try to persuade one another of their good faith and of their unwillingness to accept one penny less than some offered amount. *Ex post*, negotiation may have been a sequence of offers, first by one person, then by the other. *Ex ante*, there is no prescribed order of speech, no restriction on the content of speech and, most importantly, no prescribed time between utterances. And it is the *ex ante* sequence, or absence of sequence, that matters in actual bargaining. Nobody enforces the prescribed sequence of offers or the rule of silence in the intervals between one offer and the next. Talk is unrestricted. There are in practice no gags, and, without gags, it is virtually impossible to predict what the outcome of bargaining will be. The ordering of speech in the Staahl-Rubinstein model is more than a convenient simplification. It is an essential part of the model without which the model falls apart completely.

More importantly, the restriction on the sequence and content of speech includes a ban on non-negotiable threats. By restricting bargainers' speech to "I offer ...", "Yes" and "No", the model automatically forbids anybody to say "Under no circumstances will I ever accept less

than....”. Bargainers are forbidden to say such things not just because the words are banned from the vocabulary of the model, but, more importantly, because of the postulate, called *sequential rationality* or *sub-game perfection*, that nobody can commit himself now to an action later on if that action would not be in his own interest at that time. Suppose, for example, that the equilibrium bargain under the Stahl-Rubinstein procedure supplies 10% of the pie to person E and the remaining 90% to person O. As long as person O is rational and if person E can commit himself while person O cannot, person E can appropriate any share of the pie - say 60% - by committing himself to accept nothing less. Person E might make the commitment binding by means of a binding side contract with a third party to pay the third party a substantial sum if he accepts less than 60% of the pie in his bargain with person O. Person O must give way, for the alternative is to lose the entire pie and to acquire nothing. To be sure, person E has no monopoly on threats. If person E can make threats backed up by side contracts or by a need to preserve his reputation as a tough and astute bargainer, then so too can person O. If they threaten one another and if their threats are incompatible, adding up to more than the value of the pie to be shared, there can be no agreement and both end up with nothing.

The risk of this outcome makes bargainers cautious but does not abolish threats altogether. Abandon the rigid sequence, and the outcome of bargaining comes to depend on who gets to make the first threat, on the credibility of threats, on the parties' concern for their reputations, on how stubborn they choose to be. Abandon the rigid sequence, and the neat bargaining equilibrium disintegrates. The postulate of sub-game perfection preserves the sequence automatically. There may be times when this postulate is reasonable and accurate, but there are surely other times when it is not. Bargainers may have an incentive to hang tough. “Make yourself into a force of nature” is an old maxim of bargaining. Adolph Hitler is alleged to have said while bargaining that “one of us has got to be reasonable, and it isn't going to be Hitler.”

The *locus classicus* on threats and blackmail is Schelling's “An Essay on Bargaining” (1956). It is not, strictly-speaking, a theory of bargaining, for it supplies no formal prediction of how shares of a pie will actually be allocated among the claimants. The article is an examination of relevant considerations, placing considerable stress on commitment and on the importance of binding oneself to refuse anything less than some large share of the pie. Perhaps, the lion's share of the pie goes to whoever is the first to commit himself and to communicate that commitment to the other bargainer, but it is virtually impossible to say *a priori* who that will turn out to be.

Concerns for reputations may influence bargaining in two opposite ways. On the one

hand, you want a reputation for being reasonable and accommodating to induce prospective partners to join with you in new ventures. Nobody wants to become your partner if you are expected to be too rigid whenever conflicts of interest arise. On the other hand, costly intransigence today may pay off tomorrow as a warning to partners in future bargains that you are tough. Your partners might be induced to concede to your demands if you acquire a reputation for being stubborn enough to resist conceding to their's. You want to appear soft to prospective partners and hard afterwards. The postulate of sequential rationality, or sub-game perfection assumes such behaviour away. Seduced by the elegance of these assumptions - possibly even by the connotations of the words "rationality" and "perfection", for who can object to anything that is at once rational and perfect - a vast range of behaviour is swept out of sight. Schelling's essay remains as a corrective, even a reproach, to much of the more recent literature on bargaining.

#### E) Transaction Cost

A key feature of both the Nash bargaining solution and the Staahl-Rubinstein bargaining theorem is that bargaining is costless. The exclusive focus of these models is upon the allocation of the pie between bargainers, with no allowance for the possibility that part of the pie is used up in the process of deciding who gets what.<sup>8</sup> Actual real-live bargaining is almost always resource consuming. Bargaining as modelled is not. The exact opposite is true of the notion of transactions cost.

Transaction cost is typically defined to encompass more than the cost of bargaining, but, in so far as it refers to bargaining, it includes the full cost to all parties concerned with no apportionment of total cost to each party separately and no explanation of how bargainers' shares of the residual are determined. Transaction cost has been defined in various ways:

- "encompasses all those costs that cannot be conceived to exist in a Robinson Crusoe economy....a spectrum of institutional costs including those of information, of negotiation, of drawing up and enforcing contracts, of delineating and policing property rights, of monitoring

---

<sup>8</sup> The Staahl-Rubinstein model allows for the prospect, but not the realization, of the wastage of all or part of the pie. Agreements are provoked by the bargainers' fear that the entire pie may be lost through a failure to agree or that the present value of the of the pie may be diminished because a deal is not struck at once. But, at least in the simple versions of the models, these fears turn out to be groundless and the pie is always allocated intact. Waste enters the story with the introduction of complications from asymmetric information or through miscalculation in the model of bargaining as a confrontation of simultaneous demands.

performance, and of changing institutional arrangements. In short, they comprise all those costs not directly incurred in the physical process of production.”<sup>9</sup>

- “executing contracts effectively, attenuating opportunism”<sup>10</sup>

- “the cost of establishing and maintaining property rights”<sup>11</sup>

- “the cost of effecting a transfer of rights”<sup>12</sup>

Transaction cost features prominently, though by no means exclusively, in the literature of law and economics. A few almost randomly chosen quotations from Richard Posner’s *Economic Analysis of Law* (fifth edition, 1998) conveys a sense of how the concept is employed:

- The Coase theorem is “that, if transactions are costless, the initial assignment of property rights will not affect the ultimate use of the property” (8).

- “If there are significant elements of bilateral monopoly,...transaction costs may be quite high. Negotiations to settle a lawsuit are an example. Because the plaintiff can settle only with the defendant, and the defendant only with the plaintiff, there is a range of prices within which each party will prefer settlement to the more costly alternative of litigation. Ascertaining this range may be costly, and the parties may consume much time and resources in bargaining within the range.” (68)

---

<sup>9</sup>Steven Cheung, “Economic Organization and Transaction Costs” *The New Palgrave: A Dictionary of Economics*, volume 2, 56.

<sup>10</sup>Oliver E. Williamson, “Transaction Cost Economics: The Governance of Contractual Relations”, *Journal of Law and Economics*, 1979, 233-61. Williamson distinguishes between market and idiosyncratic transactions. Market transactions are defined as buying and selling as postulated in models of perfect competition where every good has a well-specified market price and people interact with the market rather than with one another. Idiosyncratic transactions are dealings among people who respond to one another in circumstances where no externally-given market prices can be defined. Transaction cost arises in idiosyncratic transactions.

<sup>11</sup>Douglas W. Allen (“What are Transaction Costs?”, *Research in Law and Economics*, 1991, 1-18)

<sup>12</sup>Richard Posner’s *Economic Analysis of Law* (fifth edition, 1998) , 39.

- “Efficiency requires that the driver drive more slowly. But because transaction costs with potential victims such as yourself are prohibitive, he will not do so unless the legal system steps in, as by holding him liable for damages... should an accident occur.” (180)

- “A doctor chances upon a stranger lying unconscious on the street, treats him, and later demands a fee...The cause of the high transaction cost in that case is incapacity. In other cases it may be time (e.g., the stranger is conscious but bleeding profusely and there is no time to discuss terms). In such cases, the law considers whether, had transaction cost not been prohibitive, the parties would have come to terms, and if so what (approximately) the terms would be.”(151)

The different definitions of transaction cost and the quotations from Posner on how the concept is employed raise questions about the meaning of the concept itself.

What exactly is transaction cost? The term could be defined broadly to encompass such things as the notary’s fees for transferring property or even the wholesaler’s cost of bringing food from farmer to market, but it could be defined narrowly to include only the cost of bargaining together with the loss of surplus when bargainers fail to strike a deal. The narrower definition is the more appropriate in the context of this article. In Hart’s model of dealings between two firms, transaction cost is the difference between their combined profit as it would be if all actions by both firms could be verifiably observed and their combined profit when this is not so. In Fearon’s model of bargaining and war (or in commercial or political relations where failure to strike a bargain causes potential income to be wasted), it is the cost of war weighted by war’s probability occurrence. Bargaining is costless in Alesina and Rosenthal’s model of politics, but transaction cost could be introduced by expanding the model to allow for the possibility that negotiation breaks down from time to time. An enlarged model might include a civic equivalent of Fearon’s descent into war whenever deals within the legislature or between legislature and executive cannot be arranged.

The notion of transaction cost relies upon an implicit analogy between social and physical technology. An ordinary cost is the amount of money that must be spent to acquire something. We speak of the cost of oranges at the grocery store. We speak of the cost of production of an automobile as an amount of money dependent on the price of inputs and the technology of automobile manufacture. Transaction cost is not like that. The dollar value of transaction cost cannot be inferred from the mechanics of negotiation. Bargains are not constructed like automobiles in accordance with some interpersonal equivalent of the laws of physics. Some bargains get struck quickly, easily and costlessly, other bargains only emerge after costly

litigation and delay, and we have no basis for predicting which is which. There may be no underlying social technology with which bargains are struck. The implicit analogy between social and physical technology is inherently dubious.

Though one can often imagine what transaction cost might be, the notion of transaction cost comes unequipped with a mechanism for deciding whether it is large or small in every particular case. It is as though we are expected to recognize transaction cost when we see it and to know in our hearts whether it is large or small. Those who employ the notion of transaction cost in the design of the law must rely on gut feelings or experience in deciding when transaction cost is likely to be high and when it is likely to be low. Legislators passing laws and judges interpreting laws are expected to minimize transaction cost, and the rest of us have only their word for it that they know what they are doing. That may be the best we can hope for in this imperfect world, but there remains a nagging suspicion that what is unrecognized in theory may turn out to be mischievous in practice.

For what purpose is transaction cost identified? This much is clear from the usage of the term: transaction cost is something to be minimized or, if possible, circumvented altogether in the choice of laws. Law A is better than law B insofar as the common purpose of both laws is attained by law A at a lower transaction cost. There are, however, two possible versions of this doctrine. The first and more modest version is that, whatever the objective of law or public policy, that objective should be attained at the lowest possible transaction cost. For example, if the state chooses to redistribute income from rich to poor or to engage in programs to that effect, public policy should be conducted in such a way that a given benefit to the poor is attained with the least possible harm to the rich. Unavoidable deadweight loss in taxation could easily create conditions where a gain of \$5 to the poor can only be procured at a cost of \$10 to the rich. That would be no violation of the minimization of transaction cost as long as no alternative policy could procure the \$5 gain to the poor at a cost of only \$9 to the rich.

The other version of this doctrine is less benign. In this version, the minimization of transaction cost is an aspect of the search for efficiency in the economy as a whole, where efficiency is linked to the maximization of the national income or some other measure of the sum total of the incomes of everybody in the economy.<sup>13</sup> Good law is seen as whatever minimizes transaction cost, regardless of how that cost is apportioned among plaintiff, defendant and the

---

<sup>13</sup>as distinct from a utilitarian criterion or from a measure of the national income weighted by a measure of the degree of equality in the distribution of income.



state. Transaction cost becomes the difference between the national income as it might be and the national income as it is. In this version, all dollars are equal to whomever they may accrue and from whomever they may be taken. A law that augments one person's income by \$10 and diminishes two persons' incomes by \$4 each is a good law, at least in so far as there is no alternative making all three people better off. There is much to be said for this principle in many of the contexts where it is invoked. The minimization of transaction cost may be the appropriate criterion in the design of laws for a new and as yet unsettled country where prospective settlers have no idea what roles they will play once the new country is established. It may be appropriate in commercial law where businessmen have equal chances of appearing before the courts as plaintiff or defendant.

Taken to its logical extreme, this interpretation of the minimization of transaction cost would block all redistribution of income, for there is inevitably some transaction cost in any transfer of income through the public sector from rich to poor. As a criterion for the choice of laws, the minimization of transaction cost would, presumably, be attractive to the rich and unattractive to the poor. The matter boils down to what public policy is thought to maximize. It is one thing to say that efficiency is the only relevant criterion for people who do not know whether they will be rich or poor at the moment when the law has a direct impact on their lives, though, even in this case, risk aversion creates a bias for equality in the distribution of income. It is quite another thing to say that efficiency, as the minimization of transaction cost, is the only criterion for a community of people who know perfectly well whether they are rich or poor.<sup>14</sup>

The notion of transaction is divorced from any determinate apportionment among bargainers of whatever it is they are bargaining about. A doctrine emphasizing the total cost of bargaining without reference to how the bargain is struck or how the pie is ultimately divided is only half a theory, and there is some question as to whether one half can be trusted without the other. Transaction cost may depend on who gets what in the splitting of the pie. We may not know the cost of splitting the pie until we know how and to whose advantage it is split.

---

<sup>14</sup>In texts of law and economics, we learn that the choice between negligence and strict liability in tort law turns on the costs associated with each rule, the sum of the cost of care and the expected cost of accidents which nonetheless occur. In the history of law, we learn that the very distinction between negligence and strict liability originated in the law of industrial accidents where, in practice, a negligence rule placed the burden of industrial accidents upon workers (because owners were typically not deemed negligent) and a regime of strict liability placed the burden upon owners. See Lawrence M. Friedman, "Torts", *A History of American Law*, 1986.

Be that as it may, a great virtue of the notion of transaction cost in the context of bargaining theory is as a corrective to the presumption, introduced into the models implicitly or derived as a theorem, that bargaining is costless, that bargaining divides the pie without diminishing it.

#### E) Social Technology in Rent-seeking and Conflict

As applied to bargaining, the models of rent-seeking and conflict succeed in accommodating transaction cost and the apportionment of the pie under one roof. These models were not designed to render bargaining determinate. The original rent seeking problem was the allocation of an import quota among interested importers where a restriction on total imports lifts the domestic price above the world price, generating a surplus for importers fortunate enough to receive a share, and where all or part of the potential surplus is dissipated as would-be quota holders compete for the attention of the government agency entrusted to assign shares. The entire surplus is wasted when there is no limit on the number of entrants to the rent-seeking competition. Part of the surplus is preserved for the rent-seekers when the number of entrants is limited.

Rent seeking becomes conflict when the prize, surplus or object of competition is endogenous. The paradigmatic conflict is outright war with all spoils to the victor, but the conflict model has been claimed relevant in a range of non-military contexts - labour disputes, legal disputes and commercial disputes - where people do not actually attack one another and where bargaining is normally thought of as an integral part of the resolution of disputes.<sup>15</sup> Formally, conflict differs from rent seeking in that the combined income of the competitors becomes the prize over which they compete.

---

<sup>15</sup>In “The Technology of Conflict as an Economic Activity”( *American Economic Review*, May, 1991, 130-34), Jack Hirshleifer described conflict as follows: “Conflict, as opposed to mere failure of cooperation, comes about when one or more parties calls upon a special technology. To wit, a technology where some or all contenders for resources incur cost in an attempt to weaken or disable competitors. This definition is broad enough to encompass not only war but strikes and lockouts, lawsuits, sibling rivalries within families and redistributive politics. But, for concreteness here, I use military language.”(page130) .

The main ingredient of both models is the “conflict success function”.<sup>16</sup> With only two rent seekers (or two contestants), person E and person O, the conflict success function becomes

$$s = C(F_E, F_O) \quad (11)$$

where  $F_E$  and  $F_O$  (with F mnemonic for fighting) are rent seekers’ the expenditures and where s can be interpreted either as person E’s share of the prize or as person E’s probability of acquiring the entire prize. For the function C to serve its purpose in this context, person E’s share must be an increasing function of  $F_E$  and a decreasing function of  $F_O$ . So defined, the conflict success function is a social analogue to the production function. Inputs  $F_E$  and  $F_O$  yield output s.

Models of rent seeking and conflict are alike in that  $F_E$  and  $F_O$  are chosen in a Nash equilibrium, but they differ in the nature of the prize. In rent seeking, the prize is a fixed sum of money, P. Person E chooses  $F_E$  to maximize  $N_E$ , his net revenue from the contest, where net revenue is his expected portion of the prize less his expenditure to acquire it.

$$N_E = sP - F_E = C(F_E, F_O)P - F_E \quad (12)$$

and person O chooses  $F_O$  to maximize his net revenue

$$N_O = (1-s)P - F_O = [1 - C(F_E, F_O)]P - F_O \quad (13)$$

In conflict, the stakes are endogenous and much higher. Competition is over the contestants’ combined income net of their combined fighting expenditures. The contestants are endowed with incomes  $\underline{Y}_E$  and  $\underline{Y}_O$  of which they devote  $F_E$  and  $F_O$  to fighting. Their combined net income,  $(\underline{Y}_E + \underline{Y}_O - F_E - F_O)$ , replaces the fixed prize, P, as the object over which they compete. Person E chooses  $F_E$  to maximize expected net income

$$N_E = C(F_E, F_O)(\underline{Y}_E + \underline{Y}_O - F_E - F_O) \quad (14)$$

---

<sup>16</sup>A conflict success function was implicit in Winston Bush’s model of anarchy, “Individual Welfare in Anarchy” in Gordon Tullock ed., *Explorations in the Theory of Anarchy*, Center for the Study of Public Choice, 1972. An explicit function was employed by Gordon Tullock in “Efficient Rent-Seeking”, in Buchanan, Tollison and Tullock eds. *Toward a Theory of the Rent-Seeking Society*, 1980. The term “conflict success function” was introduced and alternative forms were compared in Jack Hirshleifer, “Conflict and Rent-Seeking Success Functions: Ratio vs, Difference Models of Relative Success”, *Public Choice*, 1989, 10-1-12.

and person O chooses  $F_O$  to maximize his expected net income

$$N_O = [1 - C(F_E, F_O)](\underline{Y}_E + \underline{Y}_O - F_E - F_O) \quad (15)$$

It is not always obvious whether rent seeking and conflict is the more appropriate concept in civilian disputes where bargaining is thought to be involved. The rent seeking model is perhaps closer to everyday economic activity because the prize,  $P$ , originally a gift from the government, can easily be reinterpreted as the potential surplus in labour disputes or profit accruing if and only if those entitled to the profit can agree about how it is to be shared. On the other hand, if surplus or profit is ill-defined, or if competing parties may injure one another gratuitously to induce compliance, the conflict model with combined net income as the ultimate prize may be a better representation of what is at stake. The conflict model is perhaps closer to politics, especially when legislation has a major impact on the size and distribution of the national income.

Models of rent-seeking and conflict stand or fall upon the strength of the analogy - already discussed in connection with transaction cost - between social and physical technology. The core of the analogy is between the conflict success function and an ordinary production function, both of which make output of one kind or another dependant upon inputs of social or technical factors of production. Though the form of the ordinary production function is arbitrarily imposed, our reason for believing in the existence of some such function is that production reflects the laws of physics. The production function may be an entirely arbitrary representation of physical technology, but one is at least confident that the technology exists. Similarly, our confidence in the existence of a conflict success function in war rests ultimately on what we think we know about military technology, and our confidence in the existence of a conflict success function in classical rent seeking rests ultimately on what we think we know about the peculiarities of civil servants granting favours in the disposition of public largess.

There is also some question about use of the Nash equilibrium. There is said to be a Nash equilibrium when each person within a group of people chooses his own action on the assumption that everybody else's action is invariant in the special sense that others' actions would remain what they are regardless of what he does. The assumption is entirely reasonable within a competitive market where each person is a minute part of the whole. It is much less reasonable in two-person interactions such as bargaining. In choosing  $F_E$ , person E is assumed to look upon person O's choice of  $F_O$  as invariant regardless of how he, person E, chooses to behave; in choosing  $F_O$ , person O is assumed to look upon person E's choice of  $F_E$  as invariant regardless of how he, person O, chooses to behave. Necessary though it may be to generate a

determinate outcome, this assumption is far from innocuous, for each person's action would normally be influenced by his assessment of how the other would respond to his behaviour, especially if bargaining is stalled and a resumption of bargaining might be part of that response. The presumption in ordinary bargaining that, if I concede a bit, you might concede a bit too is ruled out by the Nash equilibrium and has no counterpart in models of rent seeking and conflict.

### **Bargaining, Voting and Parliamentary Procedure**

Democratic politics is infused with bargaining in two very different ways, one benign and the other predatory. Bargaining is required to fill in the gaps when outcomes under the rules of parliamentary procedure are less than completely determinate. Bargaining is also an essential part of the process by which a majority in the legislature can employ the power of the vote to exploit the corresponding minority, directing to itself as much of the national income and of the fruits of office as it is in the majority's interest to take. A principal objective in the design of democratic government and parliamentary procedure is to foster one aspect of bargaining and to thwart the other.

It is well-known that the outcome of majority-rule voting depends critically on how voting is conducted. Voting is frequently intransitive. Voting about the apportionment of income among themselves is always intransitive. Coupled with sincere voting, this intransitivity empowers the agenda-setter to lead the legislature by the nose to whatever outcome he pleases. Parliamentary procedure can be seen as a set of rules to draw order out of this chaos, rules about who can propose bills or amendments and about the ordering of votes among bills, amendments and the status quo. An example of such rules is that the last vote in any sequence must be between the bill as amended and the status quo, ensuring that whatever the legislature decides does not leave a majority worse off than before voting began.

Ideally, passage of bills in the legislature would be entirely determinate as long as each legislator votes in his own interest or in the interests of his constituents within the rules of the legislature as they happen to be, with nothing left to negotiate about.<sup>17</sup> That may be too much to

---

<sup>17</sup>To ask whether bargaining is inextricable in democratic politics is equivalent to asking about the existence of a political equilibrium. The subject is discussed, pessimistically, by several authors in *Political Equilibrium* (1982). An political equilibrium without reliance on bargaining is supplied by models of "probabilistic voting" and of the "citizen-candidates". I have raised questions about the applicability of these models to politics as we know it. See Usher (1994) and Usher (2005).

ask of political institutions. Situations arise where deals must be struck, benefits must be shared and accommodation must be forged among people whose interests conflict to some extent. Platforms of political parties must emerge as the outcome of genuine compromise among party delegates. Joint committees of the Senate and the House of Representatives must hammer out a common bill after similar but not identical bills have been passed in both houses of the legislature. The distinction here is between i) reaching an agreement when each participant acts in his own interest exclusively within the bounds of well-specified rules identifying a fair bargain or prescribing the sequence of votes and ii) reaching an agreement without such rules and relying instead upon mere compromise, where everybody must accept something less than his first preference so that an agreement can be struck. The distinction is between greed and love as these terms are interpreted at the beginning of this paper, where we may seek to subsume as much as possible of the world's business under the domain of greed, but where love - bargaining as unconstrained compromise - cannot be dispensed with entirely.

Much of our unease about the Nash and Staahl-Rubinstein bargaining solutions is that, by rendering bargain determinate, they draw attention away from circumstances where it is not. Attention is drawn away from the need for genuine compromise ungoverned by any rules of procedure, as in Iraq today where peace would seem to require compromise between Shiites and Sunnis with no apparent formula by which such compromise can be obtained. To model bargaining as a species of rule-bound maximization by all participants is to overlook the need for genuine compromise - out of love, duty or public-spirit - if democratic government is to proceed.

There is another consideration. To model bargaining as rule-bound maximization is to draw attention away from predatory bargaining, from an aspect of genuine bargaining (in the restricted sense of the term) that voting rules are in part designed to suppress. Throughout most of recorded history, thoughtful people doubted whether democracy could ever be made to work. Whatever its merits as an ideal, government by majority-rule voting was thought to be impossible in practice. Sooner or later, democracy would self-destruct. A majority of voters would employ its authority over the government and the army to deprive the minority of income, property and civil rights, redirecting income and privilege to members of the majority coalition and, in the end, eroding the willingness of citizens to accept the will of the majority peacefully.<sup>18</sup> Anticipating

---

<sup>18</sup>To the general proposition that democracy can be threatened by a willingness of majorities to exploit the corresponding minority, there is one very important exception. A majority of the poor may safely plunder the minority of the rich by means of progressive income taxation. Unlike exploitation of the adherents of one religion by the adherents of another, progressive income taxation need pose no threat to democracy because there is a limit - well

its fate at the hands of a predatory majority, the minority may rebel, abolishing democracy in the process. Or, anticipating its fate at the hands of its successors, an unpopular party in office might use the power of the state to squelch the opposition; better to rule tyrannically than to be dominated by a predatory majority in a democratic state.<sup>19</sup> Recognition of a certain validity in the anti-democratic argument has led political theorists to identify opposing forces in a democratic society and, more importantly, to design institutions that hold predatory majorities in check. Fear of predatory majorities was a central concern of the authors of the Constitution of the United States.

Among the constraints upon a predatory majority is the maintenance of well-specified property rights. Respect for property rights places a limit on what a minority stands to lose at the ballot box. Property rights are respected because a majority's gain from the expropriation of the property of the corresponding minority can be expected to fall short of its loss from disorganization of the economy in any attempt to tamper seriously with property rights. Not all capitalist societies are democracies, but all democracies are capitalist, at least to the extent of maintaining private ownership of a significant portion of the means of production.<sup>20</sup> The rule of

---

short of 100% - to how high a tax it would be in the interest of the poor to impose. The constraint is deadweight loss in taxation - through the trade-offs between labour and leisure, between consumption and investment, between tax payment and tax evasion - causing the tax base to shrink as the tax rate rises, and placing an upper ceiling on the tax rate it is in anybody's interest to impose. Deadweight loss in taxation supplements impediments to bargaining in the preservation of democracy.

<sup>19</sup>In *The Dark Side of Democracy*, (Cambridge University Press, 2005), Michael Mann extends this line of reasoning to ethnic cleansing. A tyrannical ruler cares little about the ethnic composition of his subjects. Government by majority rule voting may supply one ethnic group with the incentive to exploit or, in extreme circumstances, exterminate another. On this theme, see also Alvin Rabushka and Kenneth Shepsle, *Politics in Plural Societies: A Theory of Democratic Instability*, 1972. The problem was recognized by Aristotle. Referring principally to democracies, Aristotle wrote that a "state cannot be constructed from any chance body of persons...Most of the states that have admitted persons of another stock... have been troubled by sedition..."(*The Politics of Aristotle*, p210). The problem was recognized by James Madison. "...a pure democracy, by which I mean a society consisting of a small number of citizens, can admit no cure from the mischief of factions....such democracies have ever been spectacles of turbulence and contention; have been found incompatible with personal security or the rights of property; and have, in general, been as short in their lives as they have been violent in their deaths.." (*The Federalist Papers*, #10, 1789).

<sup>20</sup>On the logic of the connection between capitalism, see Usher (1981).

law plays a similar role. A ban on ad hominem legislation or unequal treatment of different people by the courts places a floor on what one stands to lose if one's party fails to win the election. Laws must not reward the supporters of the party in power or punish its enemies. Actual governments violate this principle to some extent, but there is a limit to how far one can go without placing democracy in jeopardy.

Over and above these constraints upon the content of legislation and policy are constraints in the design of government. A venerable line of political thought sees democracy as supported by the separation of powers, by checks and balances between the different branches of government and by countervailing power, three principles that are sometimes distinguished but remain very much alike. Preservation of democracy is said to require a balance among legislature, executive and judiciary and between legislatures in bicameral government. In the words of Thomas Jefferson, "An elective despotism was not the government we fought for, but one which should not only be founded on free principles but in which the power of government should be so divided and balanced among the several bodies of magistracy, as so no one could transcend their legal limits, without being effectively checked and restrained by the others." (Quoted by James Madison in *The Federalist Papers*, #48, 1789).<sup>21</sup>

Jefferson's checks and restraints can be interpreted as the frustration of predatory bargaining by constitutional rules. Exploitation of minorities by majorities requires a two-fold bargain. Members of the majority coalition must recognize one another, and they must agree upon a division of the spoils. Both facets of this predatory bargain become more of an impediment under the constitutional arrangements that Jefferson and Madison propose. With two legislatures, a predatory majority in one may be thwarted by a predatory majority (or by the absence of any such majority) in the other. To exploit their minorities, the two factions would need to compromise. That in itself may be difficult. If successful, it would be tantamount to the formation of a larger majority in society as a whole with a correspondingly smaller minority and correspondingly smaller gains from exploitation. Beyond that, a predatory majority in the legislature may be blocked by the executive or by the judiciary. Allocation of income and the spoils of office among members of the majority coalition becomes correspondingly difficult.

---

<sup>21</sup>"In the compound republic of America, the power surrendered by the people is first divided between two distinct governments, and then the portion allotted to each is subdivided among distinct and separate departments. Hence a double security arises to the rights of people. The different governments will control each other at the same time as it will be controlled by itself". (James Madison, *The Federalist Papers* #51)



Predatory bargains are easier to arrange when there is a natural split in society between clearly-identifiable tribes, such as the Sunni and the Shia in Iraq, and correspondingly difficult in a very diverse society. To cite James Madison, “Whilst all ... will be derived from, and dependant on the society, the society itself will be broken into many parts, interests and classes of citizens, that the rights of individuals, or of a minority, will be in little danger from interested combinations of the majority. ... security of civil rights...consists... in the multiplicity of interests and.... in the multiplication of sects.”<sup>22</sup> (*The Federalist Papers* , #51)

It is at least arguable that the separation of powers - between central and state governments and between the different houses of the legislature and among president, legislature and judiciary - is explicitly designed to make government inefficient, and that the source of inefficiency in divided government is the need for bargaining among the branches of government if anything constructive is to be done. Constitutions of democratic countries are designed with bargaining as friction to stop governments from working too well. Powers of government are divided between a central government and local governments with partly distinct and partly overlapping spheres of authority. None of this would block a predatory majority if bargaining were as determinate and as costless as some of our bargaining models would suggest. The division of powers and the corresponding checks and balances supports democratic government because and only because bargaining is costly and indeterminate. The hope is that friction and indeterminacy in bargaining drive up the cost of exploitation by majority rule voting to the point where the manoeuvre is no longer advantageous.

### **Concluding Observations**

Bargaining theories are like the blind men and the elephant in the old Hindu tale. One man touched the elephant’s leg and compared the elephant to a tree. Another touched its tail and compared it to a snake. Another touched its tusk and compared it to a sword. None of the blind men were altogether wrong. One of our theories prescribes a fair bargain. Another sees bargaining as emerging from an imposed sequence of offers and refusals. Another stresses the cost of striking a bargain. Another stresses threats and focal points. Another infers the bargain from the optimal efforts of the bargainers. None of the theories are altogether wrong, but they do not fit well together. There is no central switching mechanism assigning particular bargains to

---

<sup>22</sup>“If there were only one religion in England, we should have no fear of despotism; if there were two, they would cut each other’s throats; but there are thirty, and live in peace and happiness”. Voltaire, quoted in Scott Gordon, *Controlling the State* , p.230.

one mechanism or another, and there is no mixing process by which the different sketches are combined into a well-drawn portrait of how bargains are struck.

Bargaining may not be determinate at all. As mentioned at the beginning of this paper, the economist's project is to explain as much as possible of the world as the outcome of greed, just as the outcome in a competitive economy is explained as the consequence of many people acting out of pure self-interest in response to prices but not directly to one another. An essential ingredient of bargain as we experience it is precisely the response of people to one another. Bargainers must temper self-interest with compromise, a degree of concern for others, and some recognition of how others can be expected to respond to one's actions. Examples discussed in this paper - the formation of a government of Sunnis and Shiites, joint committees of the Senate and the House of Representatives to hammer out a common legislation, deals between levels of government, formation of platforms of political parties - require something over and above pure self-interest in response to externally-given prices or rules. Bargaining fill the gaps left open when countervailing self-interest would otherwise be indeterminate or chaotic. Bargaining supplies the necessary element of cooperation and accommodation, of love as Robertson called it, in a world where greed is predominant, frequently productive but not always sufficient in itself to get the world's work done. Bargaining - a capacity of politicians and others to strike deals in the absence of any identifiable mechanism or procedure by which deals are struck - is a necessary ingredient of democratic societies, though it can turn predatory in the exploitation of minorities. How bargaining works is explained to some extent by our models but remains largely mysterious.

## References

- Alesina, A. And Rosenthal, H., *Partisan Politics, Divided Government and the Economy*, Cambridge University Press, 1995
- Allen, Douglas, W., “What are Transaction Costs?”, *Research in Law and Economics*, 1991, 1-18
- Aristotle, *The Politics of Aristotle*, edited and translated by Ernest Barker, Clarendon Press, 1946
- Axelrod, Robert, *The Evolution of Cooperation*, Basic Books, 1980
- Baron, David and Frerejohn, John, “Bargaining in Legislatures”, *American Political Science Review*. 1989, 1181-1206)
- Bush, Winston, “Individual Welfare in Anarchy” in Gordon Tullock ed., *Explorations in the Theory of Anarchy*, Center for the Study of Public Choice, 1972
- Cheung, Steven, “Economic Organization and Transaction Costs” *The New Palgrave: A Dictionary of Economics*, volume 2, page 56
- Cross, John, G., “A Theory of the Bargaining Process”, *American Economic Review*, 1965, 67-94
- Edgeworth, F. Y., *Mathematical Psychics*, 1881, reprinted by A. M. Kelley, 1961.
- Ellingsen, Tore, “The Evolution of Bargaining Behaviour”, *Quarterly Journal of Economics*, 1997, 581-601
- Fearon, James, D., “Rational Explanations for War”, *International Organization*, 49, 1995
- Friedman, Lawrence, M., “Torts”, *A History of American Law*, 1986.
- Garfinkel, Michele, “Arming as Strategic Investment in a Cooperative Equilibrium”, *American Economic Review*, 1990, 50-68.
- Gordon, Scott, *Controlling the State: Constitutionalism from Ancient Athens to Today*, Harvard University Press, 1999

Harsanyi, John, "Approaches to the Bargaining Problem Before and After the Theory of Games: A Critical Discussion of Zeuthen's, Hick's and Nash's Theories", *Econometrica*, 1956, 144-57

Hart, Oliver, *Firms, Contracts and Financial Structure*, Clarendon press, 1995

Hicks, John, R., *The Theory of Wages*, 1932, chapter VII

Hirshleifer, Jack, "Conflict and Rent-Seeking Success Functions: Ratio vs, Difference Models of Relative Success", *Public Choice*, 1989, 10-1-12.

Hirshleifer, Jack, "The Technology of Conflict as an Economic Activity", *American Economic Review*, May, 1991, 130-34.

Intrilligator, Michael and Brito, Dagobert, "Can Arms Races lead to the outbreak of War?" *Journal of Conflict Resolution*, 1984, 63-84.

Madison, James, *The Federalist Papers*, #10, #48 and #51, 1788

Mann, Michael, *The Dark Side of Democracy*, (Cambridge University Press, 2005

Muthoo, Abhinay, *Bargaining Theory with Applications*, Cambridge University Press, 1999

Nash, John, "The Bargaining Problem", *Econometrica*, 1950, 155-62

Nash, John, "Two-person Cooperative Games", *Econometrica*, 1953, 128-140.

Osborne, Martin J. and Rubinstein, Ariel, *Bargaining and Markets*, Academic Press, 1990.

Priest, George L. and Klein, Benjamin, "The Selection of Disputes for Litigation", *Journal of Legal Studies*, 1984, 1-55.

Rabushka, Alvin and Shepsle, Kenneth, *Politics in Plural Societies: A Theory of Democratic Instability*, Merrill, 1972

Richard Posner, Richard, *Economic Analysis of Law*, fifth edition, Little Brown, 1998

- Robertson, D. H., "What do Economists Economize?", in *Economic Commentaries*, Staples Press, 1956.
- Rubinstein, Ariel, "Perfect Equilibrium in a Bargaining Model", *Econometrica*, 1982, 97-109.
- Schelling, Thomas, "An Essay on Bargaining", *American Economic Review*, June, 1956
- Sutton, J., "Non-Cooperative Bargaining Theory: An Introduction", *Review of Economic Studies*, 1986, 709-24
- Staahl, Ingolf, *Bargaining Theory*, 1972.
- Supreme Court of Canada, *Reference re: Secession of Quebec*, 2 SCR 217, 1998.
- Tullock, Gordon in "Efficient Rent-Seeking", in Buchanan, Tollison and Tullock eds., *Toward a Theory of the Rent-Seeking Society*, 1980
- Usher, Dan, *The Economic Prerequisite to Democracy*, 1981.
- Usher, Dan, "The Significance of the Probabilistic Voting Theorem", *Canadian Journal of Economics*, 1995, 433-45.
- Usher, Dan, "Assessing the Citizen-candidate Model", *Public Choice*, 2005, 43-65.
- Williamson, Charles, E., "Transaction Cost Economics: The Governance of Contractual Relations", *Journal of Law and Economics*, 1979, 233-61.
- Young, H. P., "An Evolutionary Theory of Bargaining", *Quarterly Journal of Economics*, 1993, 145-168
- Young, Oran, R., *Bargaining: Formal Theories of Negotiation*, University of Illinois Press, 1975
- Zeuthen, F., *Problems of Monopoly and Economic Welfare*, 1930,

## Appendix 1: Deriving the Staahl-Rubinstein Bargaining Solution

The purpose of this appendix is two-fold: to construct the Staahl-Rubinstein bargaining solution straightforwardly and in a way that reveals the assumptions on which it is based, and to show how the “disappearance effect” giving rise to the bargaining solution in equation (13) and the “discounting effect” giving rise to the Staahl-Rubinstein bargaining solution in equation (14) can both be thought as special cases of a more general formula.

Person E and person O (mnemonic for even and odd) bargain over the division of a pie worth  $\$P$ . Bargaining consists of an alternating sequence of offers, the first by person O during the year 1, the next by person E during the year 2, and so on, until an offer has been accepted or the pie has disappeared entirely. Each offer specifies a division of the pie, so much for person O and so much for person E. No speech is allowed other than the sequence, “I offer you a share of.....”, answered by “I agree” or “I don’t agree”, occurring once and only once each year until a deal is struck or the pie vanishes altogether.

For convenience of exposition, it will initially be assumed that, if no bargain has been struck by the very end of the year T, the pie will at that time be allocated arbitrarily, a share S assigned to person E and the remaining share  $(1 - S)$  assigned to person O. The Staahl-Rubinstein bargaining solution emerges when T approaches infinity. In principle, any pattern of disappearance could be accounted for. An example below allows half of the pie to disappear in each of two consecutive years.

Let T be an odd number, and let the rates of interest of persons E and O be  $r_E$  and  $r_O$  respectively so that their discount factors are  $\delta_E$  and  $\delta_O$  where  $\delta_E = 1/(1 + r_E)$  and  $\delta_O = 1/(1 + r_O)$ .

When both bargainers are rational, a bargain is struck as soon as bargaining begins in the year 1. Person O, who is entitled to make the offer in that year, chooses the lowest share for person E that person E is prepared to accept. The offer must be such that person E is no worse off by accepting person O’s offer than he could become by waiting a year until it is his turn to make an offer instead. The magnitude of the offer is discovered in a process of “backward induction”.

Since the ultimate bargain is determined as the end product of a hypothetical series of offers to be accepted or rejected, it is convenient to define two time series,  $s(E, t)$  and  $s(O, t)$  where

i) for odd years when it is person O's turn to make the offer,  $s(E, t)$  is the share that would be offered by person O to person E, and  $s(O, t)$ , equal to  $1 - s(E, t)$ , is the share person O keeps for himself, and

ii) for even years when it is person E's turn to make the offer,  $s(O, t)$  is the share that would be offered by person E to person O, and  $s(E, t)$ , equal to  $1 - s(O, t)$ , is the share person E keeps for himself.

Start by supposing that no deal has been struck before the year T when the pie is due to be arbitrarily divided with a share S for person E and a share  $(1 - S)$  for person O. Person O is entitled to make an offer in that year, but no offer to person E of less than S would be accepted, and it would not be in person O's interest to offer more. Thus,

$$s(E, T) = S \quad \text{and} \quad s(O, T) = 1 - S \quad (\text{A1})$$

Now move backward to the year T-1 when the entire pie is still intact and when person E is entitled to make the offer. Since person O can acquire a share  $(1 - S)$  by waiting for his turn to make an offer in the year T, he would accept nothing less than  $\delta_O S$  where  $\delta_O$  is person O's discount factor. That is the least person O would accept, and person E need offer him nothing more. Thus, if the pie is allocated by bargaining in the year T-1, the bargainers' equilibrium shares become

$$s(O, T-1) = \delta_O(1 - S) \quad \text{and} \quad s(E, T-1) = 1 - s(O, T-1) = 1 - \delta_O + \delta_O S \quad (\text{A2})$$

Move backward one more year to the year T - 2 when it is once again person O's turn to make an offer. Since person E can acquire  $s(E, T-1)$  by waiting, person O's offer must be

$$s(E, T-2) = \delta_E s(E, T-1) = \delta_E \{1 - \delta_O + \delta_O S\} = (1 - \delta_O)\delta_E + (\delta_O \delta_E)S \quad (\text{A3})$$

The meaning of equation (A3) is that person E's share of the pie today can be derived from person E's share as it will become two years ahead. In the equation, today is the year T - 2 and two years ahead is the year T, but the equation remains valid when today is the year T - 4 and two years ahead is the year T - 2. Specifically,

$$\begin{aligned} s(E, T - 4) &= (1 - \delta_O)\delta_E + \delta_O \delta_E s(E, T - 2) \\ &= (1 - \delta_O)\delta_E + \delta_O \delta_E \{(1 - \delta_O)\delta_E + \delta_O \delta_E S\} \end{aligned}$$

$$= (1 - \delta_O)\delta_E [1 + \delta_O\delta_E] + (\delta_O\delta_E)^2 S \quad (A4)$$

Similarly,  $s(E, T - 6) = (1 - \delta_O)\delta_E + \delta_O\delta_E s(E, T - 4)$

$$= (1 - \delta_O)\delta_E + \delta_O\delta_E \{(1 - \delta_O)\delta_E(1 + \delta_O\delta_E) + (\delta_O\delta_E)^2 S\}$$

$$= (1 - \delta_O)\delta_E [1 + \delta_O\delta_E + (\delta_O\delta_E)^2] + (\delta_O\delta_E)^3 S \quad (A5)$$

and  $s(E, T - 8) = (1 - \delta_O)\delta_E [1 + \delta_O\delta_E + (\delta_O\delta_E)^2 + (\delta_O\delta_E)^3] + (\delta_O\delta_E)^4 S \quad (A6)$

Moving toward the present, every extra two years adds an additional term to the series in square brackets and adds a power to  $\delta_O\delta_E$  in the final expression so that, for any even number n,

$$s(E, T - 2n) = (1 - \delta_O)\delta_E [1 + \delta_O\delta_E + (\delta_O\delta_E)^2 + \dots + (\delta_O\delta_E)^{n-1}] + (\delta_O\delta_E)^n S$$

$$= (1 - \delta_O)\delta_E \left[ \frac{1 - (\delta_O\delta_E)^n}{1 - \delta_O\delta_E} \right] + (\delta_O\delta_E)^n S$$

$$= \{1 - (\delta_O\delta_E)^n\} \left[ \frac{(1 - \delta_O)\delta_E}{1 - \delta_O\delta_E} \right] + (\delta_O\delta_E)^n S$$

$$= \{1 - (\delta_O\delta_E)^n\} \left[ \frac{r_O}{r_E + r_O + r_E r_O} \right] + (\delta_O\delta_E)^n S \quad (A7)$$

The first step in the derivation of equation (A7) is to replace the series in square brackets with a simpler equivalent. The second step is just to alternate terms in the first expression. The third step is to replace a function of  $\delta_O$  and  $\delta_E$  with an equivalent function of  $r_O$  and  $r_E$ , specifically,

$$(1 - \delta_O)\delta_E / (1 - \delta_E \delta_O) =$$

$$= \{1/(1 + r_E)\} \{1 - 1/(1 + r_O)\} / \{1 - 1/[(1 + r_E)(1 + r_O)]\}$$

$$= r_O / [r_E + r_O + r_E r_O] \quad (A8)$$

The meaning of equation (A7) is that, for any n less than T/2, person E's share of the pie as it would be if a deal were struck in the year T - 2n is a weighted average - the weights being  $\{1 - (\delta_O\delta_E)^n\}$  and  $(\delta_O\delta_E)^n$  - of the expression  $r_O / [r_E + r_O + r_E r_O]$  and of person E's share, S, as it would be if no deal were struck by the year T when the pie is allocated arbitrarily.

To determine  $s(E, 1)$ , we need only choose n so that  $T - 2n = 1$ . The required n is equal to  $(T - 1)/2$ , and person E's share of the pie in the bargain struck in the year 1 becomes



$$s(E, 1) = \{1 - (\delta_O \delta_E)^{(T-1)/2}\} [r_O / \{r_E + r_O + r_E r_O\}] + (\delta_O \delta_E)^{(T-1)/2} S \quad (A9)$$

which is equation (9) in the text. It follows immediately from equation (A9) that as T approaches infinity - meaning that the size of the pie never diminishes no matter how long it takes to strike a bargain - then weight  $(\delta_O \delta_E)^{(T-1)/2}$  approaches 0 and person E's share of the pie reduces to

$$s(E, 1) = r_O / \{r_E + r_O + r_E r_O\} \quad (A10)$$

which is the Staahl-Rubinstein bargaining solution (or would be if the "years" were defined to be very short so that the product  $r_E r_O$  becomes insignificant by comparison with  $r_E$  or  $r_O$  alone). If the bargainers discount rates,  $r_E$  and  $r_O$ , are 9% and 1% respectively, then person E's share of the pie in accordance with the Staahl-Rubinstein bargaining solution becomes 10%  $[(.01)/(.09 + .01)]$ , leaving 90% for person O.

It is equally evident from equation (A9) that the Staahl-Rubinstein bargaining solution is a poor approximation to person E's share of the pie if the pie disintegrates in the course of time or if an allocation is eventually imposed in the event that no bargain is struck. Suppose, for example, that, if no deal is struck in the meantime, one half the pie vanishes on the last day of the year 6 and the other half vanishes on the last day of the year 7.

Person O is entitled to make the offer in the year 7. As there would be nothing left in the year 8 when it is person E's turn to make an offer, person O can offer person E nothing (but a penny) and can keep all that remains of the pie (half the pie) for himself.

Move forward to the year 6 when it is person E's turn to make an offer to person O. Since person O can acquire half the pie by waiting one year and since person O's rate of interest is 1%, person E must offer person O a share equal to  $\delta_O/2 = [1/(1.01)]/2 = .49$ . But since none of the pie disintegrates until the very end of the year 6, person E's share acquired during the year 6, while the pie is still whole, is .51.

Finally, in the year 5, person O cannot offer person E anything less than  $\delta_E(.51) = [1/1.09](.51) = .47$  because that is what person E could acquire by waiting until it is his turn to make an offer in the year 6. Thus  $s(E,T) = .47$  when  $T = 5$ .

Person E's share of the pie,  $s(E, 1)$ , in the bargain struck in the year 1 can now be

computed from equation (A10) where

$$T = 5, S = .47, (\delta_O \delta_E)^{(T-1)/2} = \{[1/1.01][1/1.09]\}^2 = .825$$

and  $\{r_O / [r_E + r_O + r_E r_O]\} = .01/[.01 + .09 + (.01)(.09)] = .099$

Person E's share in the bargain becomes

$$(.175)(.099) + (.825)(.47) = .405$$

Person E's share of the pie would be 10% in accordance with the Stahl-Rubinstein solution of equation (14), would be 50% if neither party discounted future income, would be 40.5% when both discounting and disappearance are accounted for in accordance with equation (A9) and would be 50% if a common sense of fairness led bargainers to split the pie equally.

A similar process of backward induction would identify each bargainers' share for any arbitrarily-chosen pattern of annual disappearances of the pie. Note however that the derivation crashes completely unless there is assumed to be a "sub-game perfect equilibrium" where no bargainer can bind himself to any future action that would not be in his interest at that time. Nobody can declare, "I will accept nothing less than 3/4 of the pie even if my refusal means that no deal is struck and the entire pie is wasted." The assumption is very strong. Such threats are uttered from time to time, and mutually advantageous opportunities may be lost because bargainers are stubborn.

## Appendix B: Why Bargaining Fails: Examples from Models of Business, Law and the Choice between War and Peace

Inspection of the paradigmatic bargain in figure 1 suggests a simple reason why bargaining might fail. Parties might fail to strike a deal because they are stubborn, each demanding more of the pie than the other is prepared to concede. If neither gives way, they must both make do with the assignment of incomes at the no-agreement point where they are both worse off than they might otherwise be. The three articles discussed in this appendix take a different tack. They start from the premise that bargainers would not be stubborn, that, one way or another, parties would agree upon some allocation of the pie in the circumstances postulated in the paradigmatic bargain. They then go on to show how bargaining might fail in circumstances that differ in certain respects from what was postulated in the paradigmatic bargain.

### 1) Bargains Among Firms

In *Firms, Contracts and Financial Structure* (Clarendon Press, 1995), Oliver Hart explains the pattern of ownership as a trade off between economies of scale and the loss of incentive when one cannot reap the full benefit from one's activities. Patterns of ownership are exemplified by the relation between General Motors and the Fisher Body company that makes frames for General Motors' cars. The question is whether these two companies i) remain entirely separate, buying or selling from one another or from other companies on the open market, ii) amalgamate into one large company or iii) establish a close working relation with one another. The choice among these options depends upon economies of management and impediments to cooperation when neither firm can verify the other's relation-specific investment. Hart shows that, though the *potential* combined profit under cooperation (iii) may exceed the combined profit under amalgamation (ii), amalgamation may nevertheless be the better option when relation-specific expenditures are unverifiable. Hart's model is quite complicated, but a stripped down version is sufficient to focus on his assumption about bargaining which is our immediate concern.

Consider two firms, F and G. If they remain entirely separate from one another, their profits would be  $\pi_F$  and  $\pi_G$ . If they amalgamate into one large firm, its profit would be  $\pi_A$ . If they remain as separate entities but cooperate, their combined profit,  $\pi_C(f, g)$ , would be dependent upon their relation-specific investments, f by firm F and g by firm G. The critical assumption about the relation-specific investments is that neither firm's investment is *verifiable* by the other. Each firm is assumed to know both firms' profits,  $\pi_F$  and  $\pi_G$ , in the absence of cooperation, the

profit,  $\pi_A$ , of the amalgamated firm and the profit function,  $\pi_C(f, g)$ , of the two firms together in the event that they cooperate. Knowing its own relation-specific investment,  $f$  or  $g$  as the case may be, each firm is in a position to infer the relation-specific of the other firm, but it cannot demonstrate this knowledge objectively to a third party because outsiders cannot be expected to know  $f$  or  $g$ . That being so, an agreed-upon rule for apportioning combined profit between the firms cannot be made to depend upon their relation-specific investments. A distinction is therefore drawn between the *true* surplus,  $T$ , from cooperation where

$$T = \pi_C(f, g) - [\pi_F + \pi_G] - f - g \quad (B1)$$

and the *verifiable* surplus,  $S$ , from cooperation where

$$S = \pi_C(f, g) - [\pi_F + \pi_G] \quad (B2)$$

Two key behavioural assumptions are introduced: the *informational* assumption that only the verifiable surplus can serve as a basis for assigning each firm's share of the benefit of cooperation and the *bargaining* assumption that the two firms split the verifiable surplus,  $S$ , evenly, half to firm F and half to firm G. On these assumption, firm F's profit in the event of cooperation,  $\pi_{FC}$  where C is mnemonic for cooperation, becomes

$$\pi_{FC} = \pi_F + S/2 - f \quad (B3)$$

and firm G's profit in the event of cooperation,  $\pi_{GC}$ , becomes

$$\pi_{GC} = \pi_G + S/2 - g \quad (B4)$$

Hart's principal proposition is that, without verification, both relation-specific investments are too small and the combined profit is less than it might be because both firms invest to maximize their own surpluses rather than the surplus of the two firms together. It could easily happen that - as between amalgamation and cooperation - the total profit of the two firms together would be higher under cooperation if relation-specific investments could be verified, but the total profit is actually higher under amalgamation because each firm's relation-specific investment is concealed from the other.

Bargaining failure in this context is manifest in the form of industrial organization. Even

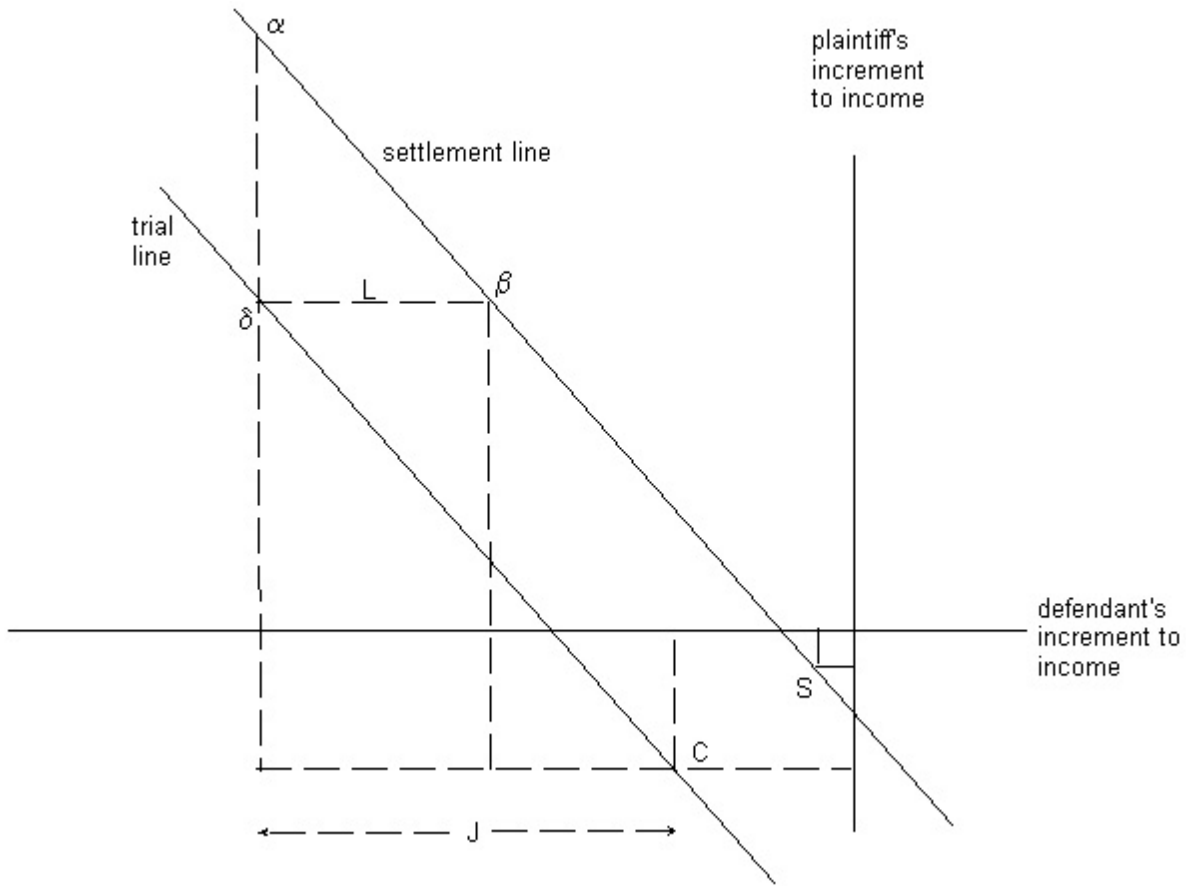
on the assumption that bargainers split the pie evenly where each party's actions can be observed, the "wrong" form of industrial organization may be chosen when the firms' actions are not only observable but verifiable by a third party in a position to punish firms for malfeasance.

## 2) Settlement to Avoid a Trial

Private disputes can be settled by going to court or by a settlement out of court in the light of how the court might be expected to rule. Since the combined cost of settlement, to plaintiff and defendant together, exceeds the combined cost of trial, one might expect that "bargaining in the shadow of the law" would displace trials altogether. Priest and Klein (1984) ask why that is not always so. They begin with two observations: that 1) about 95% of personal injury cases are settled without going to trial and that 2) among cases that do go to trial, the plaintiff's success rate is about 50%. Their explanation is that there must be a major discrepancy between the plaintiff's and the defendant's expectations of the court's decision and coupled with a more or less equal chance that either party is mistaken. I do not take issue with this explanation, but would add a second explanation which flows naturally from the analysis of bargaining in this paper: that the plaintiff and defendant fail to strike a deal because one or both are stubborn, that plaintiff and defendant may fail to settle upon any one of the many mutually-advantageous deals that might be struck.

The dispute between plaintiff and defendant can be represented by a variant of Figure 1 in the text. Represented on the axes are not incomes, but increments to income as a consequent of settlement or trial. The plaintiff's increment to income is shown on the vertical axis. The defendant's increment to income is shown on the horizontal axis. Since the defendant is expected to compensate the plaintiff, his increment to income must be negative. Increments to income (which may be positive or negative) take account of i) each person's legal cost in the event of a trial, ii) the court's judgment mandating a transfer of income, from defendant to plaintiff and iii) each person's the cost of settlement. The analysis will be undertaken in three stages. In the first, the court's judgment is known in advance to both plaintiff and defendant. In the second, there is a universally-recognized probability of a judgment for the plaintiff, but the amount of that judgment is known to both parties. Only in the third stage might plaintiff and defendant have different expectations about what the probability or the magnitude of the judgment might be.

**Figure B1: Settlement versus Trial**



When the court's judgment,  $J$ , is known in advance, the choice between settlement and trial is as shown in figure B1. The legal costs of a trial are  $C^p$  to the plaintiff and  $C^d$  to the defendant, represented together on the figure by the point  $C$ . The costs of arranging a settlement are  $S^p$  to the plaintiff and  $S^d$  to the defendant, represented together on the figure by the point  $S$ . Settlement costs are assumed to be very much less than legal costs of a trial, i.e.  $S^p < C^p$  and  $S^d < C^d$ . The diagonal line - assumed to be at 45 degrees off both axes - through the point  $C$  shows all possible increments to income of plaintiff and defendant when both pay the legal costs required for a trial and there is a transfer of income between them. Starting from the point  $C$ , a transfer from defendant to plaintiff is represented by a movement north-west from  $C$ , and a transfer from plaintiff to defendant would be represented by a movement to the south-east of  $C$ . Similarly, the diagonal line - also assumed to be at 45 degrees to both axes - through the point  $S$  shows all possible increments to income of plaintiff and defendant when both pay the required cost of settling the case without a trial and there is a transfer of income between them. Starting

from the point S, a transfer from defendant to plaintiff is represented by a movement north-west from S, and a transfer from plaintiff by a movement to the south-east of S. The horizontal distance, L, between the trial line and the settlement line can be thought of the surplus, the sum of both parties' gain from settlement rather than trial.

If plaintiff and defendant paid the legal cost of trial but the court excused the defendant from any payment to the plaintiff, the parties would find themselves at the point C. If, instead, the judgment of the court required the defendant to pay the plaintiff an amount J, the parties would find themselves at the point  $\delta$ , at which the total *loss* to the defendant is  $J + C^d$  and the net *gain* to the plaintiff is  $J - C^p$ .

From here on the story is essentially the same as in figure 1 of the text. The trial brings plaintiff and defendant to the point  $\delta$  which is below and to the left of the settlement line, so that there must be room for a mutually-advantageous deal, moving in the first instance to the point S on the settlement line and then, via the appropriate transfer, to some point between  $\alpha$  and  $\beta$  -  $\alpha$  directly above and  $\beta$  to the right of the point  $\delta$  - making both parties better off. A bit of geometry reveals that the surplus, L, is equal to  $(C^p + C^d) - (S^p + S^d)$  which must be positive if settlement is ever to be advantageous. Hence, if the court would mandate a transfer of J from defendant to plaintiff, then, starting from the point S, a mutually-advantageous transfer, T, to circumvent the court, T, must conform to equation (2) above for some x between 0 and 1, where, the larger x, the better off the plaintiff and the worse off the defendant must be.

If this model were a true and complete representation of the confrontation of plaintiff and defendant, there need be no civil trials at all, for all disputes could be negotiated away. But the model suggests reasons why this may not turn out to be so. A possible reason is that one or both parties may be stubborn. Suppose that it costs nothing to settle disputes privately, that plaintiff and defendant would both face legal costs of \$200 if the case went to trial, and that the defendant would be required to pay \$1,000 to the plaintiff if the case when to court. Thus, any transfer from defendant to plaintiff of between \$800 and \$1200 would leave both parties better off (or in the limit no worse off) than if they went to court. But how do they decide on one particular transfer between these limits? What happens if the plaintiff absolutely refuses to accept anything less than \$1150, and the defendant absolutely refuses to provide anything more than \$850? There may be an impasse leading to a trial despite the fact that both parties could be better off without one.

Priest and Klein's evidence that about 95% of cases are settled out of court can be interpreted as meaning that most people are reasonable, and the 5% going to trial might suggest

that a few people are not. The evidence that the plaintiff is successful in about 50% of all cases is not explained by the model at all. If the facts corresponded exactly to the model as developed so far, then plaintiffs would win all the time and not just half the time.

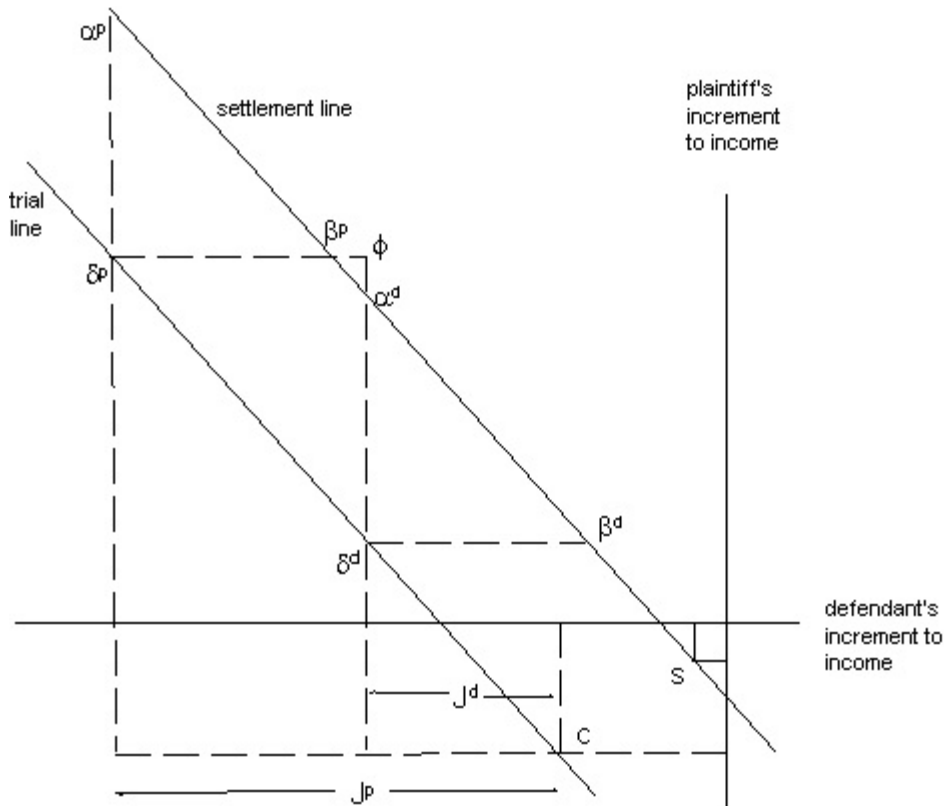
Nothing of importance changes when the assumption of a sure judgment  $J$  is replaced with a universally-recognized probability  $\pi$  of a judgment  $\$ J$  coupled with a probability  $(1 - \pi)$  that the defendant is deemed by the court not to be liable at all. As long as plaintiff and defendant are risk neutral, they are in the same position as they would be with a judgment of  $\pi J$  and no risk at all. Figure 3 and its implications would be unchanged except that  $J$  is replaced by  $\pi J$ . Once again, there is a range  $\alpha\beta$  of mutually-advantageous settlements from which one must be chosen. Settlement would be expected as long as plaintiff and defendant are not too stubborn.

What does make a difference, forming the core of Priest and Klein's analysis, is the possibility that expectations differ. Suppose for convenience that all expectations are held with the force of certainty, that the plaintiff expects the award to be  $J^p$  and that the defendant expects the award to be  $J^d$  where  $J^d$  is considerably less than  $J^p$ . The plaintiff expects the court to award him a large amount, while the defendant expects the award to be much smaller. There is of course no strong reason why expectations should not go the other way, with the plaintiff expecting the expecting the relatively smaller reward and the defendant expecting the relatively larger one, but, if that were so, both parties would be especially averse to trial and especially willing to settle, reinforcing the advantage of settlement when expectations about  $J$  are the same.

The pattern of expectations and its influence on the parties' willingness to settle out of court is illustrated in Figure B2. Anticipating a judgment of  $J^p$ , the plaintiff expects the trial to generate increments to the incomes of plaintiff and defendant as represented by the point  $\delta^p$  and, as a consequence of that expectation, would accept no point on the settlement line below and to the right of  $\beta^p$ . He would accept no transfer less than  $J^p - [C^p - S^p]$ , his anticipated award less the extra cost of trial rather than settlement. Similarly, anticipating a smaller judgment of  $J^d$ , the defendant expects the trial to generate increments to the incomes of plaintiff and defendant as represented by the point  $\delta^d$  and, as a consequence of that expectation, would accept no point on the settlement line above and to the left of  $\alpha^d$ . He would accept no transfer greater than  $J^d - [C^d - S^d]$ .



**Figure B2: Plaintiff and Defendant are Both Relatively Optimistic about the Magnitude of the Judgment**



But - as the figure is drawn with  $J^p$  very much larger than  $J^d$  - there is no point on the settlement line that is at once above  $\beta^p$  (and therefore acceptable to the plaintiff) and below  $\alpha^d$  (and therefore acceptable to the defendant). There is in short no deal, and the parties go to trial despite the fact a settlement would be mutually-advantageous if only they knew what  $J$  would turn out to be. Priest and Klein make essentially the same argument, except that they allow the opinions to differ about the probability of an award to the plaintiff rather than about its magnitude.

When the plaintiff expects to receive a higher award than the defendant expects to provide, there is, in effect, an anticipated surplus to the trial. Never mind that one or both of them must turn out to be wrong. It is the anticipated sums that matter. In this way of looking at the problem, the choice between trial and settlement becomes a choice between two surpluses: the cost surplus  $(C^p + C^d) - (S^p + S^d)$  favouring settlement and the expected judgment surplus  $J^p - J^d$  favouring trial. Together, plaintiff and defendant choose the route with the larger surplus which, as figure 4 is drawn, is the trial route.

There is another way of looking at this. If the plaintiff expects to receive an award of  $J^p$  while the defendant expects to pay an award of  $J^d$ , then the net expected outcome of the trial is the  $J^p - C^p$  for the plaintiff and  $- [J^d + C^d]$  for the defendant, represented on figure 4 by the point  $\phi$  which - given the magnitudes of  $J^p$  and  $J^d$  - turns out to be above and the right of the settlement line, indicating that settlement supplied no anticipated net surplus for plaintiff and defendant to share. Once the true award is known, a true surplus must be known as well, and it will be recognized that both parties could have been better off by settlement than by trial. That information is necessarily hidden as long as expectations differ.

It boils down to this: Plaintiff and defendant fail to resolve their dispute without resorting to a trial a) if plaintiff and defendant cannot agree upon one among many mutually-advantageous deals, some more favourable to the plaintiff and others more favourable to the defendant, and b) if one or both parties is so overly optimistic about his chance of success at trial that there are no perceived mutually-advantageous settlements, for the net expected net gain from the trial - the award expected by the plaintiff less the award that the defendant expects to pay - would exceed the difference between the sum of their costs of trial less the sum of their costs of settlement. These considerations are at bottom compatible; both may be in force at the same time. Priest and Klein stress the second consideration, but not the first.

Several extensions and qualifications may be of interest;

As when anticipations were the same, it makes little difference to the story whether expectations are about probabilities, or magnitudes of the award, or both. If plaintiff expects an award  $J^p$  with a probability  $\pi^p$  while defendant expects an award  $J^d$  with a probability  $\pi^d$ , then  $J^p$  and  $J^d$  in figure 4 must be replaced with  $\pi^p J^p$ , and  $J^d$  in figure 4 must be replaced with  $\pi^d J^d$ , but everything else remains the same.

Priest and Klein's argument that plaintiff and defendant are about equally likely to prevail when disputes come to trial reemerges in this variant of their model. Plaintiff and defendant may have different expectations about the magnitude or probability of an award, but, in the end, courts must resolve disputes. For any given magnitude of the award, there must be some true probability  $\pi$  of which  $\pi^p$  and  $\pi^d$  are only estimates. If plaintiff and defendant are equally likely to be overly optimistic about their prospects, then they are equally likely to misjudge the court's decision and equally likely to lose at trial. Nothing in the model requires  $J$  or  $\pi$  to remain the same from case to case. Some cases are large, others are small. The story remains the same.

There is an implicit assumption here of risk neutrality. Assuming instead that one or both parties are risk averse enhances the advantage of settlement over trial.

Legal costs,  $C^p$  and  $C^d$ , have been assumed invariant. It might instead have been assumed that the magnitude or probability of an award is influenced by the amounts parties are willing to spend to influence these variables. In other words, a rent-seeking model of legal cost might have been appended to the basic model of settlement and trial. The defendant may be in a position to deter suits altogether by forcing up the legal cost of the plaintiff to the point where the expected return to the plaintiff turns negative. Class action suits may then be the plaintiffs' only recourse.

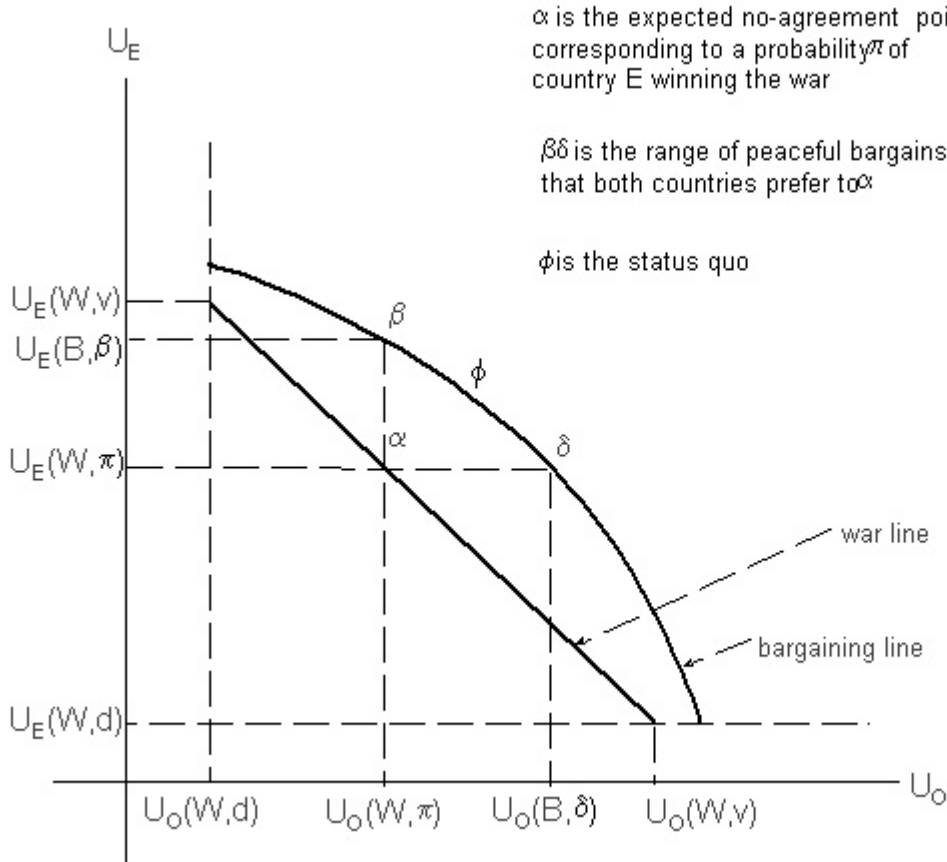
It has been assumed so far that  $J^p$  differs from  $J^d$  because one or both parties is mistaken about the true value of  $J$ . That is a reasonable assumption for traffic accidents which is primarily what Priest and Klein had in mind. But a difference between  $J^p$  and  $J^d$  may be genuine. One party may really have more to gain from a trial than the other has to lose. In compensation for injury, a dollar extra to the plaintiff is a dollar extra from the defendant as well, but, in suits over product liability, the full cost of a trial to the producer may be many times the award to the plaintiff because such an award may set a precedent for many disputes to come. A defendant might prefer settlement to trial if trial sets a precedent but the settlement does not.

### 3) Bargaining as an Alternative to War

In "Rational Explanations for War" (*International Organization*, 49,3, Summer 1995, pp.379-414), James D. Fearon raises the intriguing question of why war is not always averted by bargaining, for "...under broad conditions, the fact that fighting is costly and risky implies that there should exist negotiated agreements that rationally-led states in dispute would prefer to war" (p. 409). The article is a well-reasoned account of the circumstances where the statement may turn out to be false, where war cannot be averted by negotiation and compromise. The claim is that war cannot be averted when "(1) the combination of private information about resolve or capability and incentives to misrepresent these, and (2) states' inability, in specific circumstances, to commit to uphold a deal." (P.409) The first of these impediments to bargaining is exemplified by a situation where leaders in both states believe they have an 80% chance of winning. The second is exemplified by a situation where both states' chance of winning a war are substantially enhanced by striking first and where there is no international policeman to punish states that break a promise not to do so. Fearon's development of these propositions will be reviewed here because they are of interest in their own right but primarily to emphasize a major premise of the analysis, that, but for these impediments, some bargain from a set of all mutually-advantageous bargains would always

be struck.

The paradigmatic bargain in Figure 1 can be reconstructed to represent the choice between negotiated settlement and war. Bargainers E and O become countries, perhaps England and the Ottoman Empire. Incomes of people become supposedly well-defined utilities of nations. The no-agreement point becomes the utilities of nations in the event of war. The bargaining range becomes a continuum of peaceful outcomes, from the most advantageous for one country to the most advantageous for the other. The two principal modifications are the transformation of the no-agreement point to represent expected utilities rather than given incomes, and the distinction between the no-agreement point and the status quo. Resort to war differs significantly from a mere absence of agreement, and harmony between nations may be fostered by a status quo. These modifications will be examined in a reformulation of the bargaining problem in Figure B3.



The intrinsic uncertainty about the outcome of war converts the no-agreement point into an expected outcome. War, in this context, is a gamble in which each country's expected utility is a weighted average of its probabilities of victory and defeat. If country E wins, its utility becomes  $U_E(W, v)$ , where  $W$  is mnemonic for war and  $v$  is mnemonic for victory, and country O's utility becomes  $U_O(W, d)$ , where  $d$  is mnemonic for defeat. If country O wins, its utility becomes  $U_O(W, v)$  and country E's utility becomes  $U_E(W, d)$ . As it is always better to win than to lose,  $U_E(W, v) > U_E(W, d)$  and  $U_O(W, v) > U_O(W, d)$  as illustrated in the figure.

If country E's probability of winning is  $\pi$ , then its expected utility in the event of war becomes  $U_E(W, \pi)$  where

$$U_E(W, \pi) = \pi U_E(W, v) + (1-\pi)U_E(W, d) \quad (B5)$$

and where  $U_E(W, \pi) = U_E(W, v)$  when  $\pi = 1$  and  $U_E(W, \pi) = U_E(W, d)$  when  $\pi = 0$ . Essentially the same relation holds for country O.

$$U_O(W, \pi) = (1-\pi) U_O(W, v) + \pi U_O(W, d) \quad (B6)$$

For all values of  $\pi$  from 1 to 0, the countries' expected utilities can be represented by points on the "war line" in Figure 2, a downward sloping straight line from the point  $\{U_E(W, v), U_O(W, d)\}$  when  $\pi = 1$  to the point  $\{U_E(W, d), U_O(W, v)\}$  when  $\pi = 0$ . For any given value of  $\pi$ , the utilities of the two countries -  $U_E(W, \pi)$  and  $U_O(W, \pi)$  - are represented by a point  $\alpha$  on the war line. The point  $\alpha$  plays the role of the no-agreement point in Figure 1, once the opportunities for bargaining are introduced.

Since war is harmful while bargaining is not, any outcome on the war line can be bettered for both countries by a deal to avoid war, but, as in the paradigmatic bargain, such deals are not unique. All peaceful accommodations are shown on the "bargaining line" in Figure 2, a downward sloping from a point (not labeled) where country O is no better off than if it had lost the war but country E is better off than if it had won, to a point where country E is no better off than if it had lost the war but country O is better off than if it had won. For any point such as  $\alpha$  on the war line, there must be a range of points on the bargaining line at which utilities of both countries are greater than at the point  $\alpha$ .

Not all peaceful allocations are feasible in the sense of being preferable to both nations

than the resort to war. When the nations' expected utilities in the event of war are represented by the point  $\alpha$ , the range of feasible bargains becomes the portion of the bargaining line from  $\beta$  to  $\delta$ . At the point  $\beta$ , the entire surplus accrues to country E, leaving country O just as well off but E distinctly better off than if there had been a war. At the point  $\delta$ , the entire surplus accrues to country O, leaving country E no better off but country O distinctly better off than if there had been a war. As in Figure 1, the bargaining problem is to choose a point between  $\beta$  and  $\delta$ .

The other special feature of bargaining between nations is that there may be a "status quo" represented in Figure 2 by the point  $\phi$  on the bargaining line. In negotiation between businessmen over the sharing of profit from a joint venture, the no-agreement point and the status quo were one and the same. Failure to agree caused a reversion of incomes to whatever they would have been if the joint venture had never been considered. In dealings between nations, failure to agree peacefully can leave one or both nations worse off than before. In one case, negotiation takes both parties from their original position to something better. In the other, negotiation failure takes one or both parties from their original position to something worse.

As long as the status quo lies within the range from  $\beta$  to  $\delta$ , it may serve as a focal point for peaceful cooperation. War may be averted through mutual recognition of its privileged position on the bargaining line.<sup>23</sup> Each country says to the other, "I'll fight rather than accept any point on the bargaining line other than the status quo itself." Neither country may be inclined to push for change, and there is no war. No country would accept a bargain different from the status quo for fear that it would be pushed, bit by bit, to the wrong edge of the bargaining range - country E toward  $\delta$  and country O toward  $\beta$  - at which peace is no better than the expected outcome of war.

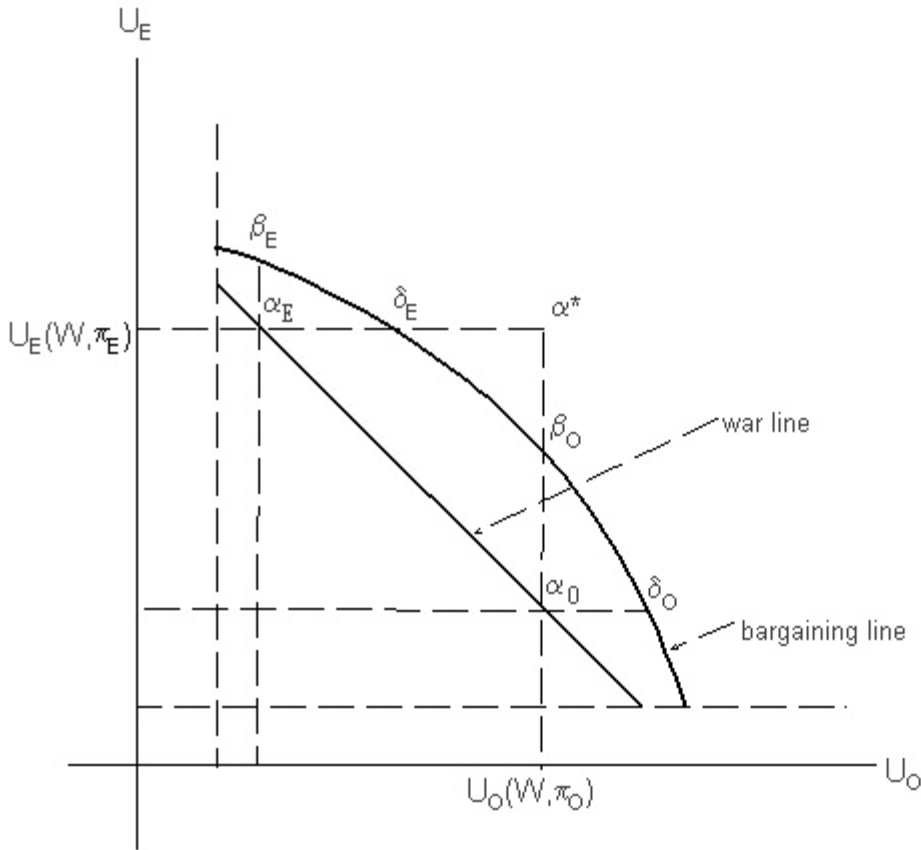
When potential conflict is over territory, the status quo would be a more or less well-recognized border that both nations are willing to respect. But respect for a status quo is far from inevitable. In a dispute about territory, country E may claim ownership because it captured the territory in the glorious battle of 1362, while country O may claim ownership because it captured the territory in the glorious battle of 1478. Or disputes between nations may be about matters such as trade relations or global warming for which there is no prior rule that all nations have been accustomed to recognize. Or a status quo once recognized may not be recognized forever. It is an unfortunate characteristic of any status quo that, as illustrated in Figure 2, both countries may be better off ceding some territory to the other - country E moving along the bargaining line toward the point  $\delta$ , or country O moving along the bargaining line toward the point  $\beta$  - if the inevitable

---

<sup>23</sup>Focal points are discussed in Schelling, "An Essay on Bargaining" (1956).

alternative to concession is war. One country may simply announce that the old status quo is no longer acceptable.

**Figure B4: Some Causes of War**



There is another possibility. As illustrated in Figure B4, countries may resort to war because they differ in their assessments of the prospect of victory. It was assumed in the construction of figure B3 that there was no disagreement about the probabilities of victory or about the countries' expected utilities, as represented by the point  $\alpha$  on the war line. That may not be so. Instead, each nation may be relatively confident of its prospect of victory. Country E may see its probability of winning as  $\pi_E$ , and country O may see its probability of winning as  $\pi_O$  where,  $\pi_E > \pi_O$ . Different expectations of victory imply different expected utilities of war represented by distinct no-agreement points  $\alpha_E$  and  $\alpha_O$ , where excess confidence in victory implies that  $\alpha_E$  is to the left of  $\alpha_O$  along the war line.

Corresponding to the no-agreement points,  $\alpha_E$  and  $\alpha_O$ , are a distinct ranges,  $\{\beta_E, \delta_E\}$  and  $\{\beta_O, \delta_O\}$ , on the bargaining line. Country E's bargaining range  $\{\beta_E, \delta_E\}$  must begin and end to the

left of the beginning and the end country O's bargaining range  $\{\beta_O, \delta_O\}$ , but the ranges may or may not overlap. If the ranges overlap, there is still room for a bargain to avert war, somewhere between  $\beta_O$  and  $\delta_E$ . Otherwise, as illustrated in Figure 3, there is none. As Figure 3 is constructed, the countries' predicted utilities in the event of war -  $U_E(W, \pi_E)$  and  $U_O(W, \pi_O)$  as indicated by the point  $\alpha^*$  above the bargaining line - are high enough that war really would be preferable to negotiation if both countries' predictions of their utilities were correct. One or both predictions must be mistaken, but that mistake leads to war.

Another cause of war can be illustrated by a reinterpretation of Figure B4. Think of  $\alpha_E$  and  $\alpha_O$  not as perceived no-agreement points, but as actual no agreement points depending on which country strikes first. Once again, there would be a mutually-advantageous bargain if  $\alpha_E$  and  $\alpha_O$  were close enough that  $\beta_O$  is to the left of  $\delta_E$ , leaving a range within which both countries are better off than they would expect to be in the event of war, regardless of who strikes first. Such a bargain is possible but very fragile. Fearing that the other country might strike first regardless, each country acquires a strong incentive to do so itself.<sup>24</sup>

A third interpretation of the no-agreement points in figure B4 is that countries' prospects at war may change over time. Think of  $\pi_E$  and  $\pi_O$  as pertaining to different times rather than different expectations or different actions. Let  $\pi_E$  be country E's probability of winning a war today, while  $\pi_O$  is country E's probability of winning a war tomorrow. If  $\pi_E$  is significantly larger than  $\pi_O$ , then country E has an incentive to declare war today, when its probability of winning is high, rather than to wait until tomorrow when its probability of winning is low and when the terms of a bargain to avert war will have become unfavourable. Bargains at different times may be interdependent, as when a country's ability to wage war tomorrow depends on how much territory it will have acquired in bargaining today. A reasonable and accommodating country might be nickle-and-dimed into defeat.

Behind these causes of war lies a problem of commitment. Almost regardless of the

---

<sup>24</sup>Both countries' first strike advantage may be reduced when both countries are armed. By arming itself, country E increases its chances of winning a war, pushing  $\alpha^E$  and  $\alpha^O$  to the left along the war line. Correspondingly, country O would push  $\alpha^E$  and  $\alpha^O$  to the right. When both countries arm themselves, it is at least possible that  $\alpha^E$  and  $\alpha^O$  are pushed together, opening an opportunity for bargaining that might not otherwise exist. On arming see, Garfinkel, Michele, "Arming as Strategic Investment in a Cooperative Equilibrium", *American Economic Review*, 1990, 50-68 and Intrilligator, Michael and Brito, Dagobert, "Can Arms Races lead to the outbreak of War?" *Journal of Conflict Resolution*, 1984, 63-84.



technologies of bargaining and of war, some bargain preferable to war can almost always be devised as long as both sides can be confident that promises will be kept. It is precisely that confidence which may be lacking in relations between nations. Promises not to strike first or not to reopen negotiation tomorrow when relative strengths have changed are not credible unless resort to war is disadvantageous to both countries.

In explaining the descent from negotiation to war, Fearon starts with the working premise that one way or another, some bargain from the set of all mutually-advantageous bargains would be struck if circumstances conformed to the paradigmatic bargain or to its modification in Figure 2. Somehow, countries would settle upon one point on the continuum from  $\beta$  which is the best for country O to  $\delta$  which is the best of country E. Fearon's argument is that even if some bargain would be struck in these conditions, the circumstances of nations are different enough from the paradigmatic bargain that no alternative to war can be agreed upon. The mere statement of this premise suggest an additional cause of war. Countries may resort to war because they cannot agree on which bargain to adopt. Countries may be just plain stubborn, each demanding a better deal than the other is prepared to allow.

A common thread runs through all of the examples in this appendix: the sharing of profit, the choice between settlement and trial and the descent into war. All identify special circumstances of the parties in conflict, on the working assumption that some accommodation will be found when disputes conform to the paradigmatic bargain. Either a specific bargain (such as a fifty-fifty split) is postulated or it is explained how actual disputes differ enough from the paradigmatic bargain that no common ground remains. It is not my purpose to fault this procedure. Bargainers do often split the difference. Special circumstances can be interesting and important. Rather it is my purpose to identify the common premise as preface to a critique of bargaining models *per se*, in an attempt to show that confidence in the determinacy of bargaining can be unwarranted. My prey is the common premise that, but for these special circumstances, bargains are always and inevitably struck. It is the widespread belief that bargaining models justify the common premise, and that bargaining is determinate in much the same way as outcomes are determinate in the general equilibrium of a competitive economy.