

# Mysterious Bargaining

Dan Usher  
Economics Department  
Queen's University  
Revised, April, 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

It is fifty years since Thomas Schelling published “An Essay on Bargaining” (*American Economic Review*, June 1956), and, though the paper has attained the status of a still much cited classic, the principal idea in the paper tends to be overlooked. At its core is the proposition that bargaining is fundamentally indeterminate. Nothing in the economist’s models of rational behaviour predicts which bargain - from a set of mutually-advantageous bargains - will ultimately be struck, especially as it is rational for either party to concede to the other’s demands when those demands are backed up by an irreversible promise to accept no less favourable deal. Since that article appeared, there have emerged elaborate bargaining theories, based for the most part on a postulated shared sense of fairness or on an imposed bargaining process. A bargaining solution is implicit in models of rent seeking, conflict and transaction cost. But the fundamental bargaining problem remains as elusive as ever.

As these words are being written, elected politicians in Iraq are attempting to form a government. No party has a clear majority, but some majority coalition must emerge if government is to function at all. By the time the reader sees these words, the matter may have been resolved one way or another, in agreement or in chaos. Whether the politicians will prove to have been sufficiently reasonable, fairminded and trusting in one another for democratic government to proceed is a question our bargaining solutions - to be discussed presently - simply cannot resolve. There is nothing unique in this regard about Iraq. There is a comparable, if less

lethal, uncertainty in the United States today, where the Senate and the House of Representatives, having passed very different versions of an immigration bill, must reconcile their differences if a any bill is to emerge as law.

All democratic government is an intricate fusion of voting and bargaining. Political parties are inevitably coalitions of interests that must be accommodated in party platforms and in public policy once a party is elected. Differences between legislature and executive must somehow be resolved. The great principle of countervailing power among legislature, executive and judiciary is enshrined in the Constitution of the United States and is implicit in the unwritten constitutions of all democratic countries. Checks and balances are built into the constitution to keep the dictator at bay. In all this, there is an unspoken premise that the branches of government will, one way or another, succeed in reconciling their differences and working together.

Bargaining is pervasive in the economy too. Among the great virtues of the principal model of the competitive economy is its freedom from bargaining. Given an initial allocation of resources among people, self-interested behaviour in response to market- prices determines what is produced and how goods and services are allocated. The model is at once a great intellectual edifice and a basis for the formation of public policy. Exiled though it may be from the core model of the economy, bargaining is nevertheless indispensable for i) the sharing of the profit of a firm among partners with different skills and different outside options, where each partner's contribution is unique and no partner's contribution can be replicated exactly by services that may be purchased at invariant market-determined prices, ii) the sharing between firms of the returns from a joint venture where each firm's input is essential and iii) wage-setting in negotiation between employer and union. Something of the determinacy of the competitive model places limits upon outcomes in these situations, but negotiation cannot be dispensed with altogether. No modern economy could run if bargaining - the knack of striking deals not entirely predetermined by the initial conditions - were absent altogether.

Our legal system would become very much more expensive, if not completely unworkable, but for the prospect of resolving differences privately in the light of what the parties to a dispute expect the outcome of litigation to be. In international relations, bargaining may be the only alternative to war.

Recognition that people do bargain and that civilized life would otherwise be impossible motivates to a search for models within which the outcome of bargaining is determinate. The search is for a well-specified social technology - analogous to the physical technology in standard

economic models - that transforms competing interests into a unique agreed-upon outcome, just as inputs of factors of production are transformed into outputs of goods in an ordinary production function. The price mechanism supplies order without orders and without cooperation. The great theorem in economics, rendered commonplace by repetition, is that with nothing more than an initial ascription of property rights, people acting in their own interest in response to market-determined prices give rise collectively to a unique, predictable outcome that is best for everybody in the sense that no planner, however wise, however benevolent, however powerful, can rearrange the economy to make everybody better off. The ideal in bargaining theory is to render the outcome of negotiation equally determinate.

The objective of bargaining theory is to discover a mechanism enabling self-interested bargainers to resolve their differences, a mechanism that is at once a plausible representation of the circumstances of negotiation and a device enabling an outsider to predict what the outcome of any particular bargain will be. This essay is a review of bargaining theories to determine how well and to what extent that objective has been attained. It is argued that the objective has at best been obtained 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.

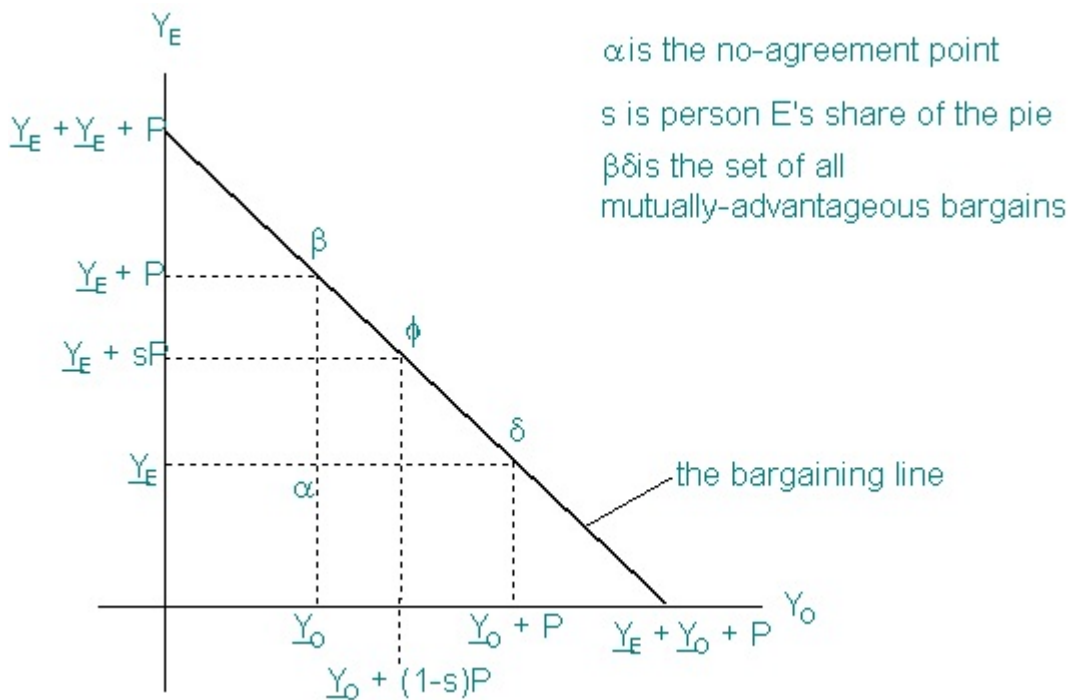
From here on, this article proceeds as follows. We begin with the classic bargaining problem where two parties are jointly entitled to a pie if and only if they can agree between themselves how large each person's slice is to be. There follows a series of examples from the literature of economics, politics, international relations and law where simple bargaining solutions are postulated. It is not my purpose here to fault the literature to be discussed, especially as much of it is directed to drawing out the implications of bargaining solutions or to explaining why simple models may fail to capture complex aspects of real live bargains, rather than to justifying bargaining solutions per se. Nevertheless, regardless of the intentions of the authors, such literature may easily convey the impression that bargaining is more determinate and more predictable than is actually the case. The principal bargaining models - models of fairness and of imposed process - are then examined with particular emphasis upon the correspondence between the pictures of bargaining in the models and what most people imagine actual bargaining to be, and upon the extent to which the models supply confidence that bargains will actually be struck. Also to be discussed are "transaction cost" as an aspect of bargaining not always incorporated in formal bargaining models, the representation of bargaining as "conflict" and the critical role of bargaining in "checks and balances" as a necessary condition for the viability of democratic government.

## 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.

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.

**Figure 1: The Paradigmatic Bargain**



The money at stake is  $\$P$  (mnemonic for pie). 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 horizontal 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 bargains 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. Post-bargaining incomes for which both bargainers are better off after the bargain than they were before are represented by points on the segment of that line between  $\beta$  and  $\delta$ . The chosen bargain is represented by the point  $\varphi$ . An increase in  $s$  moves  $\varphi$  toward  $\beta$  and a decrease in  $s$  moves  $\varphi$  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, a specification of the bargainers' untouchable property rights apart from the object of the bargain, and an understanding that the bargain is final, so that no bargainer can return to the table demanding a portion of the other's agreed-upon share. Otherwise, bargaining would be impossible because, in so far as bargaining leads to sharing, a person could appropriate another's property completely by 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) 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.

iii) 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.

iv) 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

$$s_A + s_B + s_C = 1 \tag{1}$$

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 feasible bargains is expanded from a line to a plane, the set of all mutually-advantageous bargains (generalizing the line segment from  $\beta$  to  $\delta$  in Figure 1) becomes a triangle in a three dimensional 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.

v) 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 where coalitions tend to be formed on the basis of characteristics of bargainers, such as race, religion or geography, but where any member of the coalition can be replaced if he becomes too greedy.

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 reserve the entire national income for itself. Members of the majority coalition may be bound together perhaps by some identifiable characteristic of people - such as race, ethnicity 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.

vi) Asymmetric Information: Bargainers may disagree about the location of the no-agreement point and the size of the pie, or they may have private information that they cannot or are unwilling to share. Consider an astigmatic extension of figure 1 in which person E and person O have different versions of the figure and where, depending on the circumstances, each party may or may not be able to see the other party's version. One can easily imagine a situation in which what seems appropriate to person E seems inappropriate to person O, and vice versa.

vii) 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 uses up resources in the attempt to augment his share. Actual bargaining, may be different. The bargaining process may be time-consuming. 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. Bargainers may devote labour time and money to persuading one another to moderate their demands. Mere passage of time may be more or less costly to a bargainer depending on his rate of discount. One would 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.



viii) Commitment: The paradigmatic bargain in Figure 1 is atemporal; negotiation and the splitting of the pie take place in an instant. Actual contracts may not be like that at all. Person E may be required under the terms of the contract to do this today, while person O is required to do that tomorrow. An investor supplies an innovator with cash in return for a share of the profit. The passage of time may have no significant impact on bargaining if everybody is guaranteed to keep his word. Problems arise when there is some risk that one or both bargainers may fail to do so. Mutually-advantageous deals may break down because participants cannot credibly commit themselves to do as they promise.

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) Externalities: By implication, since nothing was said about the matter, the paradigmatic bargain affected nobody but the bargainers themselves. The pie over which people bargained was at the same time an addition to their combined incomes and to the entire national income. Not all bargains are like that. All bargains are advantageous to the bargainers, but some bargains are *neutral* in the sense of having no impact on anybody but the bargainers themselves, some bargains are *beneficial* to other people, and some bargains are *detrimental* to other people. Beneficial bargains are exemplified by a deal among scientists with different skills to invent a new product which they will patent. They must bargain over the sharing of the revenue from their patent, but others gain too because the value of a new product to society exceeds the revenue from the patent. Detrimental bargains are exemplified by cartels. Bargains between worker and employer tend to be neutral in that incomes of other people are more or less the same regardless of whether wages are high and profits low, or vice versa.

### **Bargaining Assumptions About Industrial Organization, Politics, War and Constitutions**

As preface to the discussion of formal bargaining models in the next section, this section is a demonstration of how a postulated bargaining solution is invoked in several contexts:

dealings between firms, formation of public policy within the legislature and by agreement between legislature and executive, negotiation as an alternative to war, and constitutional prescriptions. In each context, bargaining can be looked upon from two distinct points of view. On the one hand, determinant bargaining may be postulated as an ingredient of models of business, politics and war where the object of the analysis is to explain something other than bargaining itself, where a simple assumption about bargaining, such as a fifty-fifty split, is good enough for the purpose and where the working assumption that bargains will be struck is not seen as assurance that this is really so. On the other hand, the incorporation of bargaining into models of business, politics and war could be seen as ground for confidence that bargains will somehow be struck, that two parties caught in the simple bargaining paradigm in Figure 1 will focus on one of the many mutually-advantageous bargains and will strike a deal accordingly.

Negotiation may come to appear more determinate, cooperation may seem easier and difficulties may be postulated out of sight by the presumption that disputes will, one way or another, be resolved. It is often said that disagreements among reasonable people can be settled amicably by getting together “at the table” to work out a mutually acceptable solution. Models incorporating the assumption that bargaining is determinate may seduce us into believing that is really so and, perhaps, into placing too much reliance on bargaining to get the world’s work done. Confidence in the determinacy of bargaining is demonstrated by the final example in this section, a decision of the Canadian Supreme Court not specifying rights, but constitutionally mandating the Federal and provincial governments to strike a 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 on economies of management and on impediments to cooperation when neither firm can verify the other’s relation-specific investments. 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

$$T = \pi_C(f, g) - [\pi_F + \pi_G] - f - g \quad (2)$$

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

$$S = \pi_C(f, g) - [\pi_F + \pi_G] \quad (3)$$

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

$$\pi_{FC} = \pi_F + S/2 - f \quad (4)$$

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

$$\pi_{GC} = \pi_G + S/2 - g \quad (5)$$

Hart's principal proposition is that, without verification, both relation-specific

investments are too small and total profit,  $\pi_C$ , 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 when each firm's relation-specific investment is concealed from the other.

Our concern here is with the role of the bargaining assumption in this argument. Hart never explains why it is reasonable to suppose that bargainers agree to split the verifiable surplus equally. Except for an off-handed reference to the Nash bargaining model (to be discussed below), he treats his bargaining assumption as self-evidently valid.

## 2) Legislature and Executive

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

$$x_Q = \alpha x_L + (1-\alpha)x_R \tag{6}$$

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

$$x = \beta x_Q + (1 - \beta)x_P \quad (7)$$

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 lie 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 different connection between voting and bargaining is studied in David Baron and John Frerejohn in “Bargaining in Legislatures” by (*American Political Science Review*, 1989, 1181-1206). Alesina and Rosenthal restricted the objectives of politicians and voters to the choice of points on a left-right continuum. Baron and Frerejohn require legislators to allocate the entire national income among the constituencies they represent, and, in doing so, propose a 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. Their solution is to impose rigid constraints on the right to propose bills and amendments.

The task of the legislature is to allocate a total national income of  $\$Y$  among  $n$  constituencies, each represented by a legislator whose sole objective is to acquire the largest possible income for his constituency. The legislature supplies  $\$y_i$  to each of  $n$  constituencies, where of course  $\sum y_i$  must equal  $Y$ . Allocations of the national income are incorporated into bills that are passed or defeated by majority rule voting. Since every such bill can be defeated in majority rule voting by some other bill with a different allocation, there can be no equilibrium outcome without the imposition of strong procedural rules. The paper is largely about the consequences of different rules. In the extreme case where a bill is to be voted up or down on the understanding that the entire national income would be wasted if the bill were voted down, the randomly-chosen proposer would assign almost the entire national income to his constituency with only a pittance left over for the rest of the country. That outcome is avoided by complex rules leading not to full equality of income, but to considerably more equality than would arise in the extreme case.

The reader may well ask what, if anything, this has to do with bargaining. It turns out that an important variant of bargaining theory (the Staahl-Rubinstein bargaining solution to be discussed below) is based upon an imposed process not unlike that postulated here.

### 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

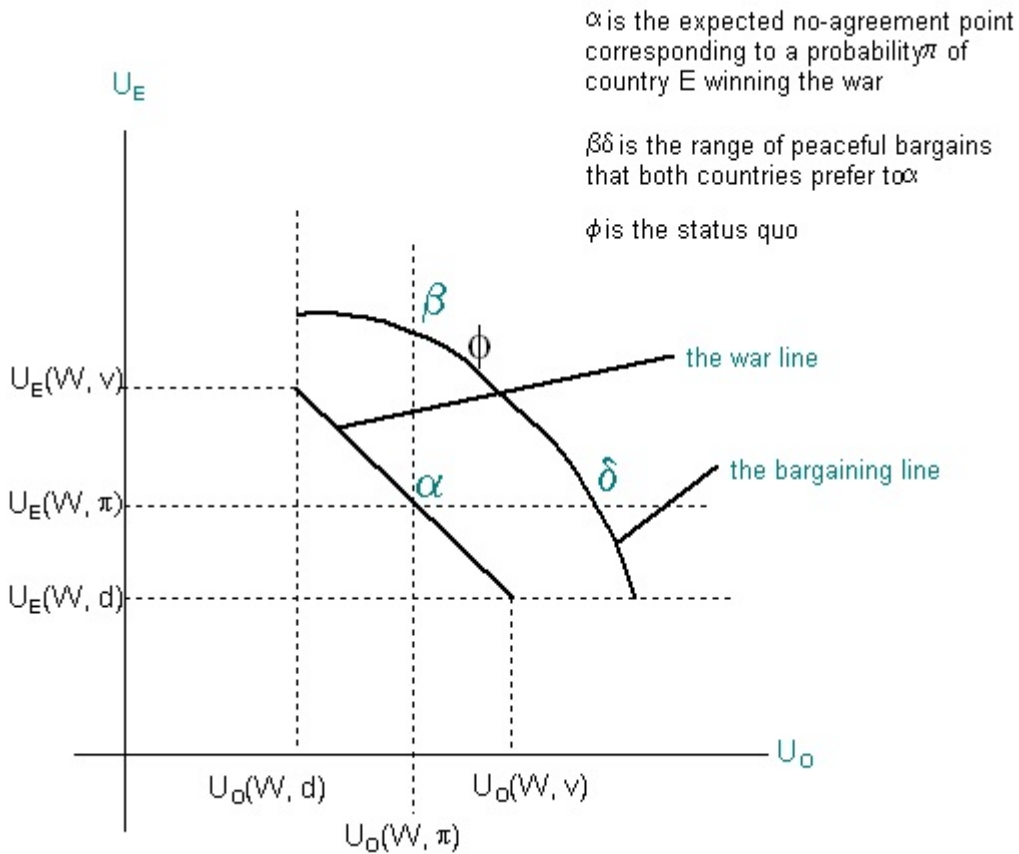
bargaining, for “...under broad conditions, the fact that fighting is costly and risky implies that there should exist negotiated agreements that rationally-led states in dispute would prefer to war” (p. 409). The article is a well-reasoned account of the circumstances where the statement may turn out to be false, where war may be rational even though there can be found a bargain that both states prefer to war. War may be rational when states are unable to locate or agree on such a bargain because “(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. The paper is a careful analysis of these propositions. So far as I can tell, it is correct. My concern here is with a major premise of the analysis, that war is averted by bargaining in the absence of these impediments.

Fearon’s purpose is to explain war, modeled as a variant of the no-agreement point in the paradigmatic bargain in Figure 1. Since both countries could expect to be better off by bargaining than by fighting, since war only occurs once bargaining fails, and since the failure of bargaining is attributed to circumstances postulated away in the paradigmatic model, it is not unreasonable to begin with the assumption that some bargain would be struck if the assumptions of the model were strictly true. The logic of Fearon’s paper is that, *even if* people were sure to strike a bargain in some situations, they would be unlikely to do so in others.

Fearon’s model of bargaining as an alternative to war is illustrated on Figure 2 which is a modification and extension of the paradigmatic bargain in figure 1. The bargainers are countries rather than people. The disputing countries are England (E) and the Ottoman Empire (O). The countries’ objectives, as shown on the vertical and horizontal axes, are utilities,  $U_E$  and  $U_O$ , rather than incomes. The dispute is about something, such as territory, that can be apportioned between them, with a share  $s$  to country E and a share  $(1-s)$  to country O. The dispute may be resolved by bargaining or by war.

Suppose, first, that the dispute is resolved by a war. 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  $U_E(W, d)$ . Obviously, as it is better to win than to lose,  $U_E(W, v) > U_E(W, d)$  and  $U_O(W, v) > U_O(W, d)$ .

**Figure 2: Utilities in Bargaining and in War**



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

$$U_E(W, \pi) = \pi U_E(W, v) + (1-\pi)U_E(W, d) \quad (8)$$

so that  $U_E(W, \pi) = U_E(W, v)$  when  $\pi = 1$  and  $U_E(W, \pi) = U_E(W, d)$  when  $\pi = 0$ .

Essentially the same relation holds for country O.

$$U_O(W, \pi) = (1-\pi) U_O(W, v) + \pi U_O(W, d) \quad (9)$$

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 the point  $\alpha$  on the war line. The point  $\alpha$  will play the role of the no-agreement point in Figure 1, once the opportunities for bargaining are introduced.

Since war is costly while bargaining is not, any outcome on the war line can be bettered for both countries by a deal to avoid war. Let the countries' utilities by bargaining be  $U_E(B, s)$  and  $U_O(B, s)$  where B is mnemonic for bargaining, where s can be thought of as country E's share of the territory or whatever else is in dispute, so that  $U_E$  increases steadily with s and  $U_O$  decreases accordingly. Outcomes of all possible bargains are illustrated on the "bargaining line" in Figure 2. The bargaining line is 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 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.

The points  $\beta$  and  $\delta$  on the bargaining line play exactly the same role as their counterparts in figure 1. All points on the bargaining line between  $\beta$  and  $\delta$  leave both countries better off than they could expect to be in the event of war. At the point  $\beta$ , the entire surplus accrues to country E, leaving country O just as well off but E distinctly better off than they could expect to be 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 than they could expect to be if there had been a war. It is convenient to think of s as varying from 1 to 0 between the points  $\beta$  and  $\delta$  along the bargaining line; if the bargaining line were a straight line, the variable s could ne interpreted as country E's share of the surplus. As in figure 1, the bargaining problem is to choose a point between  $\beta$  and  $\delta$ .

Fearon's claims about bargaining and war can be condensed into three propositions: 1) There is almost always a bargain that leaves both countries better off than if they had gone to war. 2) In so far as compromise to avert war can be represented by the simple paradigmatic bargain in Figure 1, there would be no war, for rational countries would bargain their way to a peaceful resolution of conflict. 3) Wars occur when circumstances of countries in conflict differ significantly from the paradigmatic bargain. I take issue not with the first or third propositions, but with the second exclusively. Nothing need be said here about the second proposition because it is the principal subject of this entire paper: that there is no equilibrium bargain - to avert war or for any other purpose - comparable to the equilibrium outcome in a competitive economy. Countries may well go to war because they cannot agree to any one of a number of mutually-advantageous bargains. It may, however, be of interest to list some of the war-provoking modifications of the paradigmatic bargaining problem, if only to emphasize the significance of its

implicit assumptions.

An important difference between bargaining to avoid war in Figure 2 and the paradigmatic bargain in Figure 1 lies in the relation between the no-agreement point and the status quo. In the paradigmatic bargain, they are one and the same; bargainers start from the status quo, better themselves if they can strike a deal, and fall back to the status quo if they cannot. In bargaining to avoid war, the status quo is peace and the no-agreement point is war which is usually, though not invariably, worse for both countries than peace. The no-agreement point is the failure to strike a deal in Figure 1 and the resort to war in Figure 2. In both figures, the no-agreement point,  $\alpha$ , is below the bargaining line and is worse for both countries than any point on the bargaining line between  $\beta$  and  $\delta$ .

On the assumption that the status quo is peaceful - if only the calm before the storm - the status quo may be represented by a point,  $\phi$ , on the bargaining line, where  $\phi$  lies between  $\beta$  and  $\delta$  if the status quo is preferable to war for both countries, to the right of  $\delta$  if war would be advantageous in expectation to country E, and the left of  $\beta$  if war would be advantageous in expectation to country O.

Distance between status quo and the expected outcome of war can be a strong force for peace. As long as the status quo lies within the range from  $\beta$  to  $\delta$ , it serves as a focal point for peaceful cooperation. Each country says to the other, "I'll fight rather than accept any point on the bargaining line other than the status quo itself." Then, neither country is 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 - to the point  $\delta$  for country E, and to the point  $\beta$  for country O - at which peace is no better than the expected outcome of war. This mechanism breaks down when  $\phi$  is to the left of  $\beta$  or to the right of  $\delta$  or when countries disagree about the location of the status quo. 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.

Causes of war can be illustrated by the replacement of a single no-agreement point  $\alpha$  with a pair of no-agreement points,  $\alpha^E$  and  $\alpha^O$ , corresponding to the countries' estimates,  $\pi^E$  and  $\pi^O$ , of their chances of winning the war. When both countries are relatively optimistic about their chances of winning a war, that is when  $\pi^E > \pi^O$ , country E's no-agreement point  $\alpha^E$  must lie to the left of country O's no-agreement point  $\alpha^O$  along the war line. Corresponding to no-agreement points,  $\alpha^E$  and  $\alpha^O$ , are 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 country O's bargaining range  $\{\beta^O, \delta^O\}$ , but the ranges may or may not overlap. If the ranges  $\{\beta^E, \delta^E\}$  and  $\{\beta^O, \delta^O\}$  overlap, there is still room for a bargain to avert war. Otherwise there is none.

Why might  $\pi^E$  differ from  $\pi^O$ ? Three possible explanations, corresponding to three strong reasons why war might not be averted, are a first strike advantage, change over time in relative strengths of the countries, and the impact of the bargains themselves upon the countries' prospects in war.

The simplest of these explanations is the first strike advantage. Think of  $\pi^E$  as country E's chance of winning the war if country E strikes first, and of  $\pi^O$  as country E's chance of winning if country O strikes first instead. Clearly,  $\pi^E > \pi^O$  (so that  $\alpha^E$  is to the left of  $\alpha^O$ ) as long as striking first conveys any advantage at all. If  $\alpha^E$  and  $\alpha^O$  are close enough that there is an overlap between the corresponding ranges,  $\{\beta^E, \delta^E\}$  and  $\{\beta^O, \delta^O\}$ , on the bargaining line, then war might be averted by judicious bargaining with an agreement in the range between  $\beta^O$  and  $\delta^E$ . But if  $\alpha^E$  and  $\alpha^O$  are far enough apart that  $\delta^E$  is to the left of  $\beta^O$ , then war can never be averted by bargaining and the country that strikes first wins.<sup>1</sup>

A second explanation of the resort to war is that the position of  $\alpha$  may change over time. Think of  $\pi^E$  and  $\pi^O$  as pertaining to the times when each country is relatively strong. 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 lower and when the terms of a bargain to avert war would be unfavourable.

A third explanation is that  $\pi$  as it will become tomorrow depends critically on the bargain that is struck today. For example, when the bargain is about disputed territory, each country could

---

<sup>1</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.

become stronger tomorrow depending on how much territory it acquires today. Let  $s$  be the share of the disputed territory acquired by country E today. Country E's chance of winning a war tomorrow may be  $\pi(s)$  where  $\pi$  is an increasing function of  $s$ . A "reasonable" country might be nickle-and-dimed into defeat. Bargains that might otherwise be struck today become unacceptable to one country or the other because of their influence on the prospect of war tomorrow.

Behind these causes of war lies a problem of commitment. Almost regardless of the 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 countries. 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. The problem of commitment is compounded when, as is argued in this paper, bargaining may break down from a failure of the countries to agree about which among a set of mutually-advantageous bargains - some favouring one country, and some favouring the other - will finally be struck .

Some of the characteristics of bargaining to avert war between countries have their counterparts in bargaining among firms and among branches of government. In politics, there may be a status quo on the left-right continuum that can only be abandoned gradually or political outcomes may be influenced by who has the most to lose from the disarray in society that might be expected if no deal among politicians can be struck. In business, the break-up of an on-going relation among firms may be worse for everybody than if no such relation had ever been established.

The main consideration for the purposes of this article is that, in none of these three models, is bargaining really explained. In each case a simple bargaining solution is postulated as part of an explanation of something else: of forms of industrial organization, of durable democratic politics, and of the descent from negotiation to war. The combined impact of these models should be to heighten rather than to diminish concern about the fundamental bargaining problem in Figure 1, for they highlight the need for a bargaining solution without supplying any real ground for the belief that such a solution exists. Psychologically, they may well have the opposite effect. Models postulating a bargaining equilibrium may somehow generate confidence that an equilibrium exists. Influence of models such as these is unlikely to have permeated the Canadian Supreme Court, but a prevailing attitude that reasonable men can strike a deal is well-illustrated in a recent case.

#### 4) A Constitutional Duty to Negotiate

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.

## 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>2</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.

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 not be represented except by gains in utility. Utility may supplant 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 and not comparable from one person to the next. When bargaining is over a sum of money,  $P$ , a bargain is deemed to be fair when  $\Delta Y_E$  is equated to  $\Delta Y_O$  where  $Y_E$  and  $Y_O$  are the incomes of bargainers  $E$  and  $O$ , where  $\Delta Y_E = sP$ ,  $\Delta Y_O = (1-s)P$  and  $s$  is person  $E$ ’s share of the pie. A fair bargain would be one for which  $\Delta Y_E = \Delta Y_O$ , or, equivalently,  $s = 1/2$ . 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

---

<sup>2</sup>For an excellent survey of contemporary bargaining theory, see Abhinav 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.

O. We would like to define a fair allocation of the pie as one for which  $\Delta u^E = \Delta u^O$

$$\text{where } \Delta u^E = [u^E(\underline{Y}_E + sP) - u^E(\underline{Y}_E)] \quad (10)$$

$$\Delta u^O = [u^O(\underline{Y}_O + (1 - s)P) - u^O(\underline{Y}_O)] \quad (11)$$

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>3</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 of shares cannot. The Nash bargaining solution is to choose  $s$  to maximize the product

$$\Delta u^E \Delta u^O = [u^E(\underline{Y}_E + sP) - u^E(\underline{Y}_E)] [u^O(\underline{Y}_O + (1 - s)P) - u^O(\underline{Y}_O)] \quad (12)$$

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 or 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.

---

<sup>3</sup>John Nash, "The Bargaining Problem", *Econometrica*, 1950, 155-62.

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>4</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.

Regardless of whether the notion of fairness pertains to money or to utility, there remains some question about whether bargainers would be inclined to accept a fair allocation as binding upon themselves. For bargains over the allocation of money, there are reasons why an agreement to split the pie equally might be expected. A fifty-fifty split is what Schelling, in the article cited above, referred to as a focal point. 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 which bargainer is to play the role of person E and which 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>5</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

---

<sup>4</sup>Suppose i) that persons E and O are bargaining over the allocation of \$P between them, 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  $[(sP)^{1/2}][(1-s)P]$  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.

<sup>5</sup>Such a mechanism is analyzed by Robert Axelrod in *The Evolution of Cooperation*, Basic Books, 1984.



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 fairly, where a fair resolution of disputes is understood to be an even splitting 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 property to be allocated in the service of the common good. A person may hold both views simultaneously, but in different contexts.

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>6</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.

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. Both Hicks and Zeuthen discuss bargaining as a sequence of concessions, but neither provides a satisfactory explanation of the timing and the magnitude of concessions and neither allows 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 and that, if you do not concede quickly, then I must. At least three kinds of harm might be identified. a) refusal to agree on a deal so that each party loses what

---

<sup>6</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.

would otherwise be his share 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.<sup>7</sup>

### 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.

---

<sup>7</sup>The postulated sequence of concessions could be replaced by a pattern of simultaneous demands. Call person E's demand  $P_E^*$ , and call person O's demand  $P_O^*$ . It might be assumed that both parties get what they demand as long as the sum of the demands is less than or equal to the pie, but that, otherwise, the entire pie is wasted. Person E gets  $P_E^*$  and person O gets  $P_O^*$  if and only if  $P_E^* + P_O^* < P$ . Any residual could be shared equally or wasted. The outcome would then depends on each person's expectation of how the other will behave. This framework was employed by John Nash in "Two-person Cooperative Games", *Econometrica*, 1953, 128-140. Dynamic extensions of this framework, allowing bargainers to learn from experience how large their demands can safely be, were developed in H.P. Young, "An Evolutionary Theory of Bargaining", *Quarterly Journal of Economics*, 1993, 145-168 and in Tore Ellingsen, "The Evolution of Bargaining Behaviour", *Quarterly Journal of Economics*, 1997, 581-601.

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 if 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 if 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

$$P_E = \sum_{t \text{ even}} p_t \quad \text{and} \quad P_O = \sum_{t \text{ odd}} p_t \quad \text{and} \quad P_E + P_O = P \quad (13)$$

where  $p_t$  is the size of the slice of the pie that disappears on the  $t^{\text{th}}$  year and where, so defined, the values of  $P_E$  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,

$$P_E = r_O / (r_E + r_O) \quad \text{and} \quad P_O = r_E / (r_E + r_O) \quad (14)$$

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>8</sup> A simple, and hopefully intuitive, proof of the theorem is contained in the appendix to this paper.

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 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

---

<sup>8</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.

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. For example, if the interest rates of bargainers E and O were 9% and 1%, and if bargaining would persist forever unless a deal were struck, then, as shown in equation (14), the bargainers would immediately strike a deal assigning 10% of the pie to person E and the remaining 90% of the pie to person O. By contrast, if bargainers do not discount future income but if half of the pie vanishes six years ahead and the other half vanishes in the seventh year, then, as shown in equation (13), the bargainers would immediately agree to a fifty-fifty split. A simple extension of the proof of the Staahl-Rubinstein theorem in the appendix allows for the computation of equilibrium shares in intermediate cases where bargainers discount future income *and* the pie vanishes over time. Specifically, if half of the pie vanishes six years ahead and the remaining half vanishes in the seventh year, if person E discounts at a rate of 9% and if person O discounts at a rate of 1%, the equilibrium bargain assigns 40.5% of the pie to person E and 59.5% of the pie to person O. Person E's share of 40.5% in this deal is much closer to his share of 50% when the pie vanishes over two years but the bargainers do not discount future income, than to his share of 10% when the pie remains whole but the bargainers discount future income. 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.

The Staahl-Rubinstein solution is entirely dependent upon the imposed bargain procedure. 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 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.

This questionable feature of the Staahl-Rubinstein bargaining solutions is highlighted by Baron and Frerejohn's appeal to the solution in "Bargaining in Legislatures", included in our list of applications above. In their model, the entire national income is allocated by majority rule voting in the legislature, the sole concern of each of the  $n$  legislators is with the portion of the national income acquired by his constituents, and the imposed sequence of offers in the Staahl-Rubinstein bargaining model is replaced by a procedure where entitlement to propose a bill or an amendment each time period is determined by lot in a process that is repeated again and again until some bill is accepted by a majority of the legislature. Legislators vote self-interestedly in anticipation of what else may be proposed if the present bill or amendment is rejected. There is no direct communication, one on one, between legislators. Legislators do not talk to one another or split the difference when they disagree. The skeptical reader may well ask whether it might not be more nearly correct to say that the Staahl-Rubinstein model is a representation of Parliamentary procedure than to say that the Baron and Frerejohn model is a representation of bargaining in the legislature.

Clean and interesting as it may be, the Staahl-Rubinstein solution 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 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 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 Staahl-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 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" referred to at the beginning of this paper. 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, laying 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>9</sup> Actual real-life 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 performance, and of changing institutional arrangements. In short, they comprise all those costs not directly incurred in the physical process of production."<sup>10</sup>

- "executing contracts effectively, attenuating opportunism"<sup>11</sup>

---

<sup>9</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.

<sup>10</sup>Steven Cheung, "Economic Organization and Transaction Costs" *The New Palgrave: A Dictionary of Economics*, volume 2, 56.

<sup>11</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

- “the cost of establishing and maintaining property rights”<sup>12</sup>
- “the cost of effecting a transfer of rights”<sup>13</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)

- “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 term). 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)

---

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>12</sup>Douglas W. Allen (“What are Transaction Costs?”, *Research in Law and Economics*, 1991, 1-18)

<sup>13</sup>Richard Posner’s *Economic Analysis of Law* (fifth edition, 1998) , 39.

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 are expected to design laws, and

judges are expected to interpret laws, 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>14</sup> Good law is seen as whatever minimizes transaction cost, regardless of how that cost is apportioned among plaintiff, defendant and the state. Transaction cost becomes the difference between the national income as it might be and the national income as it is with impediments to efficiency collected under the heading of transaction cost. 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.

---

<sup>14</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.

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>15</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.

#### 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 be assigned a share of the quota, 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

---

<sup>15</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.

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>16</sup> Formally, conflict differs from rent seeking in that the combined income of the competitors becomes the prize over which they compete.

The main ingredient of both models is the “conflict success function”.<sup>17</sup> With only two rent seekers (or two contestants), person E and person O, the conflict success function becomes

$$s = C(F_E, F_O) \quad (15)$$

where  $F_E$  and  $F_O$  (with F mnemonic for fighting) are the expenditures of the two rent seekers and where  $s$  can be interpreted in either of two equivalent ways: 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, it must be the case that  $s$  is 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  $F_E$

---

<sup>16</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.” (page 130).

<sup>17</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  $F_O$  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.

$$N_E = sP - F_E = C(F_E, F_O)P - F_E \quad (16)$$

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

$$N_O = (1-s)P - F_O = [1 - C(F_E, F_O)]P - F_O \quad (17)$$

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  $Y_E$  and  $Y_O$  of which they devote  $F_E$  and  $F_O$  to fighting. Their combined net income,  $(Y_E + Y_O - F_E - F_O)$ , replaces the fixed prize,  $P$ , as the object over which they compete. Person  $E$  chooses  $F_E$  to maximize expected net income

$$N_E = C(F_E, F_O)(Y_E + Y_O - F_E - F_O) \quad (18)$$

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

$$N_O = [1 - C(F_E, F_O)](Y_E + Y_O - F_E - F_O) \quad (19)$$

As between rent seeking and conflict, it is not always obvious which of the two is more representative of civilian disputes where bargaining is thought to be involved. Though rent seeking models were designed to represent conflict over an externally-supplied prize such as entitlement to an import quota, the prize might equally-well be thought of as a surplus in labour disputes or other ordinary commercial transactions. 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 potential 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 the surplus itself 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

distribution of income in the nation as a whole.

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 distinct 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.

The central question in assessing the conflict success function as an explanation of bargainers' shares is whether there really is any such underlying technology of bargaining. Here the Fearon model may be useful in sorting things out. The conflict success function might best be thought of as a model of war that may be averted by bargaining, rather than of bargaining itself. In war, as in ordinary production, there may be a determinate relation between inputs and output, where the output of war is a probability of winning. Bargaining is fundamentally different. Bargaining requires agreement and voluntary cooperation between independent agents in a way that production, rent seeking and war do not.

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 other's actions are what they are regardless of what he does. The assumption is entirely reasonable as applied to a competitive market where each person is an infinite 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 blocked by the imposition of a Nash equilibrium and has no counterpart in models of rent seeking and conflict.

### **Impediments to Bargaining as the Savior of Democracy**

From our account of bargaining so far, one might reasonably infer that impediments to bargaining are unambiguously harmful: blocking mutually-advantageous deals between firms, impeding political accommodation and even precipitating war when, for one reason or another, bargains to avert war are thwarted. Good laws are virtually defined as minimizing transaction cost. Costless bargaining makes possible a higher output than would be obtainable otherwise. There is, however, another side to the story. Not all bargains are socially-advantageous, and impediments to bargaining may be desirable when the bargains themselves are not.

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 exploit and expropriate the corresponding minority, depriving 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> Recognizing its fate at the hands of a predatory majority, the minority may rebel, abolishing democracy in the process. Or, recognizing 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

---

<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 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.

dominated by a predatory majority in a democratic state.<sup>19</sup> Recognition of a certain validity to 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. The rule of 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 associated with the organization and 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

---

<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).

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>20</sup>

What has this to do with bargaining? The connection is that the anti-democratic argument and its implications for the design of government in a democracy can be restated as propositions about bargaining. The exploitation of minorities in majority-rule voting is a predatory bargain, and the separation of powers is intended to block such bargains or to render them much less advantageous to the bargainers.

Imagine a small society, like Ancient Athens, where every citizen is a member of the legislature. The formation of a predatory majority requires i) that income, office and privilege can be reserved for the members of a majority coalition, ii) that potential members of a majority coalition succeed in finding one another, and iii) that members of the coalition strike a deal with one another about the allocation among themselves of whatever benefits the coalition supplies. When there are many issues to be resolved - the death penalty, punishments for the different crimes, school funding, public health and, and above all, potential transfers of income from one ethnically, geographically or religiously defined group to another - every legislator could choose to vote sincerely on every issue, but, in all probability, a majority coalition of legislators can make themselves and their constituents better off by log rolling, by agreeing on a common platform and voting according, even though each member of the coalition would prefer to vote otherwise on some issues. Despite the constraints, each member of the majority coalition can expect to become better off by participation than by voting in accordance with his first preference on each issue separately, very much better off than he could expect to be if a coalition had formed from which he is excluded.

Everything now depends on whether a bargain can be struck, and, if struck, can be

---

<sup>20</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)

expected to hold. If any of our bargaining theorems are right - if each legislators is governed by a sense of fairness in his dealings with other members of his voting coalition, or if an imposed or agreed-upon bargaining process yields a determinate outcome - then, though we cannot say in advance which coalition will form, we can say with some confidence that some coalition will form, that legislators outside the coalition will fare badly, that legislators within the coalition will fare well and that democracy will not last very long. On the other hand, if bargaining is difficult, costly and perhaps unfruitful, coalition formation may be blocked except where outcomes are advantageous to the great majority of legislators and their constituents. Ordinary cost-benefit analysis can be looked upon as the outcome of a deal among all legislators. The Ministry of Transport may be empowered to pick and choose from a virtually unlimited menu of new roads and road improvements in accordance with rules that make each legislator and his constituents as well off as possible in expectation and in the long run, even though nobody gets exactly what he wants today. Only when relatively small majority coalitions can be blocked by impediments to bargaining from adopting their favourite bills and squashing the rest, will democracy prove durable and satisfactory.

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>21</sup> (*The Federalist Papers* , #51)

The proposed remedy for faction is the separation of powers, but what exactly is this power of which political theorists speak? It is certainly not military. The legislature does not literally go to war with the executive. Nor do the two houses in a bicameral legislature go to war with one another. I suggest that at least part of the meaning of “power” in this context is “veto power”, the right of the President to veto legislation, or the right of one branch of a bicameral legislature to block bills passed in the other, requiring negotiation and compromise on a single bill that can be passed in both houses. In so far as this is the meaning of power, it would seem to stem primarily from the difficulty in bargaining not just in the simple paradigmatic case in figure

---

<sup>21</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.

1, but when many actors have a collective right to the pie.

Countervailing power is in some respects like an increase in the size of the coalition required to trigger an allocation of the pie. The larger the required coalition the more difficult it becomes to strike a bargain and the smaller the loot that can be appropriated from people remaining outside.<sup>22</sup> The larger the required coalition, the more likely it becomes that policies for the general benefit of society will be adopted and that policies benefitting one group at the expense of the rest will not.

A different interpretation of power may also be relevant. Power may be the capacity to command, where the organized few can overcome the unorganized many. A soldier acting alone may be no match for a few determined civilians, but an entire country of civilians may be no match for a disciplined brigade, acting in unison as commanded by its leader. Countervailing power exerted by one branch of government over another may in part be the power to interrupt the chain of command. The policeman has no hesitation in obeying his superior's orders to arrest me when I am seen to have committed a crime and when his actions are sanctioned not just by his immediate superiors but by the law as he understands it and by the entire apparatus of the state. He may hesitate to arrest me for no other reason that he is commanded to do so when he knows I have broken no law and, especially, when the legislature and judiciary has decreed that the arrest would be illegal. Soldiers may disobey their commanding officers if the army is seen to be usurping the authority of other branches of government. Commands may be unreservedly binding if and only if consistent with a consensus among all three branches of government, a consensus that only negotiation can create.

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

---

<sup>22</sup>This view of the interaction of bargaining and voting is in sharp contrast to Alesina and Rosenthal's model of democratic politics, in which bargaining is typically benign. The principal difference lies, in my opinion, in assumptions about the content of political disputes. All political preferences in their model are squeezed onto a one-dimensional left-right scale, automatically restricting the content of bargains - among legislators and between the legislature and the executive - to compromises about the location of a unique political outcome on that scale and simultaneously ruling out the opportunity for one group to exploit another. On that assumption, the outcome of bargaining is a point on the scale not enormously different from the first preference of the median voter.

inefficiency in divided government is the friction created by 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 a sort of 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**

Our search for a plausible bargaining equilibrium has proved unsuccessful despite a stacking of the deck in bargaining's favour. In the discussion surrounding figure 1, we have postulated a well-specified framework for the bargaining problem. We have postulated a unanimous recognition of what exactly is at stake in the bargain, who is entitled to negotiate for a share of the pie and who is altogether excluded. Bargainers cannot 'raise the stakes' to enforce compliance with their demands. No third party can intervene in negotiation, threatening to block a deal unless he is compensated; one is either recognized from the beginning as a bargainer, or one is not. Even within these restrictions, bargaining remained indeterminate in the sense that the assumptions required to render bargaining determinate - assumptions such as that one party has the sole right to make a take-it-or-leave-it offer or that there must be an equal split of the pie - are simply too far from the circumstances in which actual bargaining takes place. Without these restrictions, bargaining becomes even more elusive.

The holy grail of bargaining theory is a model comparable to the general equilibrium model of the economy, a model capturing essential features of bargaining as we know it while at the same time depicting the outcome of bargaining as the joint consequence of self-interested actions by rational bargainers, each maximizing something in his own interest exclusively within the constraints at hand. In a competitive economy, price-taking behaviour by consumers or firms is rational in precisely that sense. Consumers maximize utility. Firms maximize profit. That the world never conforms exactly to the model is true enough but beside the point. Circumstances may be imagined where rational self-interested actions by all parties give rise to a well-defined outcome. Bargaining is different. Empirical regularities can be observed, and one cannot deny that bargains do get struck successfully, but there is no counterpart to the maximization subject to



constraints in the competitive market. Bargaining theory can be seen as attempts to supply a such a counterpart. Models of fairness and process do so, but only, in my opinion, by postulating away what I see as intrinsic and essential aspects of the bargaining process. With a common conception of fairness, there is really nothing to bargain about. Externally-imposed processes remove the need for what we commonly think of as negotiation. The notion of transaction cost leaves open the question of how cost is allocated among bargainers. The conflict success function is more applicable to fighting than to bargaining.

This is not to deny that bargaining is both ubiquitous and indispensable. People bargain all the time, and civilized society could not exist otherwise. Actual markets could not function without bargaining, notwithstanding its exile from the world of perfect competition. Not even in theory can democratic government be maintained without a substantial capacity for deal-making among politicians. Thus, it can hardly be deemed unreasonable for authors articles on relations among firms, democratic government or war to postulate bargaining solutions.

But an imposed solution is not a rational explanation, and there remains a nagging suspicion that empirical regularities cannot be relied upon in any particular case. What is inexplicable in theory may be untrustworthy in practice too. Bargains over large sums may be struck easily. Bargains over trifles may fail at considerable cost to all concerned. Bargaining within the economy is easier than bargaining in politics because, in the economy, the failure of a deal between A and B does not usually preclude a similar deal between A and C or between B and D, while, in politics, there may be no additional participants waiting in the wings, as illustrated in the example of Iraq where the alternative to agreement may be chaos. That is why the hardening of the line between left and right in contemporary democratic politics is so concerning.

We know that bargains are often struck. Models of bargaining can be insightful. Statistical regularities may be observed. Yet our capacity to bargain is neither rationally explained nor entirely trustworthy in practice. With reference to figure 1 above, we simply do not know which, if any, among all mutually-advantageous bargains from  $\beta$  (the most advantageous for person O) to  $\delta$  (the most advantageous for person E) will in the end be chosen. Bargaining remains 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.

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

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

Staaht, 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

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: 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

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

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 statements are made from time to time, and mutually advantageous bargains do fail to be struck because the bargainers are stubborn.

