Mysterious Bargaining

Dan Usher
Economics Department
Queen's University

January 2003

Abstract

Economists do not understand how bargains are struck. A bargain is the sharing of a pie between two or more people who are collectively entitled to the pie but cannot appropriate it until they agree how large each person's slice is to be. We know that people do strike bargains and that civilized life could not proceed otherwise. We do not know how the required agreement is reached. Theorists have solved the bargaining problem, but only by the imposition of strong, artificial and unrealistic constraints. Trusting that the existence of some complex solution has been demonstrated, applied economists are content to postulate a simple one: that bargainers split the difference in actual disputes. This paper begins with examples of imposed bargaining solutions in politics and corporation finance. There follows a critical examination of the principal bargaining theories - based on notions of fairness or of imposed bargaining procedures - with emphasis on the fragility of their assumptions and on their susceptibility to threats and blackmail. The paper closes with a brief discussion of connections among theories of bargaining, rent-seeking and conflict.

JEL Classification C7

Keywords: bargaining, fairness

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

This paper is not an addition to the world's stock of knowledge or ideas. It is rather a confession of ignorance. It is a reminder that neither economists nor anybody else understands how bargains take place. There are theories of bargaining, but these are more like routes around bargaining, replacements of indeterminate bargaining with some artificially determinate process. The thing itself remains elusive.

The matter is important because the absence of a satisfactory account of how rational and self-interested people strike bargains leaves a substratum of indeterminacy beneath our models of markets and politics. From a distance, we know how markets work in guiding production and allocating output among people. Up close, production and distribution are the outcomes of deals the exact content of which cannot be predicted from a knowledge of the interests of the participants. Politics is even more elusive. Democratic politics is a complex interplay between voting and bargaining. The outcome of voting cannot be predicted from a knowledge of the interests of the voters. There is no denying that people do strike bargains and that bargaining is not outrageously expensive most of the time. There remains a nagging fear that what we do not understand may in time foil us in unexpected ways.

This article may be seen as a protest against the tendency in much of the literature of economics today to assume away the bargaining problem on the strength of bargaining theory that would seem to suggest a determinacy which, in my opinion, is largely illusory. I am not objecting to the theory *per se*. As will be discussed below, determinacy does follow from the assumptions. Nor am I objecting to the practice of reasoning as though bargaining were determinate in many situations, for theory can be insightful even though indeterminacy is assumed away. I am objecting to our tendency to forget how fragile and arbitrary our theories may be. Perhaps, in this age where politicians are expected to apologize on behalf of their constituents for all manner of wrong-doing by themselves or their ancestors, I am suggesting that economists should apologize for their assumptions. We should at least take care not to convey the impression on the strength of bargaining theory that our working assumptions about bargaining are accurate depictions of real markets or real politics. Beyond that, I believe politics and law place too much faith in the efficacy of negotiation. These concerns will be illustrated in a few examples below.

Ubiquitous in actual markets, bargaining is absent altogether from thenomists' principal model of how markets work. The crux of general equilibrium is price-taking. As goods are produced, traded and consumed, there emerges a set of market-determined prices seen as invariant in the special sense that no person's actions bulk large enough in the market as a whole to affect prices by more than an insignificant amount. Each person looks upon prices as fixed independently of how he chooses to behave. Everyone knows that prices change in response to changes in technology and preferences in the community as a whole. What matters is that nobody can budge prices, all by himself. Everyone conducts his affairs as though prices were invariant even though, strictly speaking, this can never be quite so.

With prices invariant, there can be nothing to bargain about. People trade not with one another, but with "the market". People in general equilibrium live astonishingly lonely lives. As producers, traders and consumers, people never threaten one another, never bluff, never form alliances, never even talk to one another. Taking prices as fixed, they produce to maximize income and they consume, given their incomes, to maximize utility. Perfect competition is the perfect absence of competition as the term is commonly understood.

Useful as the model of perfect competition may be as a foundation for the analysis of public policy, universal price-taking is not entirely accurate as a description of how markets work . People do bargain. Economies could not function otherwise. Failure to account for bargaining as rational self-interested behaviour is a major gap in the economist's understanding of the world. Among the principal bargains we would like to explain are

- allocating 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,

- sharing between firms of the returns from a joint venture where each firm's input is essential,

- vote-trading among legislators with different interests, where no faction is large enough to attain all of its objectives without support from other factions,

- wage-setting in negotiation between employer and union,

- negotiating in the shadow of the law when litigation is expensive,

- settling disputes among nations or between the central government and provincial governments in a country where a federal constitution assigns ill-defined powers to both levels of government.

The prototypical bargain is a unanimous and voluntary agreement by a group of people on the allocation of a sum of money among themselves. On the table lies a pile of P dollars, where P is mnemonic for pie or prize depending on the context. A group of N people sitting around the table may allocate the money among themselves in any way they please, but they must agree upon an allocation before any of the money can be touched. The agreement specifies each bargainer's slice of the pie. It establishes a set $\{y_1, y_2, ... ,y_N\}$, such that

$$y_1 + y_2 + ... + y_N \# P \tag{1}$$

where y_i is the income, or slice, of person i and where the inequality becomes an equality in the event that none of the pie is wasted.

It is sometimes convenient to add the assumption that people do not leave the table entirely empty-handed in the event of a failure to agree. Instead, it might be supposed that each person would be supplied with a "no-agreement" income x_i as a consolation prize. The bargaining problem becomes to choose a set $\{y_1, y_2, ... ,y_N\}$ conforming to equation (1) but with the additional constraint that

$$x_i \# y_i \tag{2}$$

signifying that everybody becomes better off with an agreement than without one.

The problem is only interesting when the prize is larger than the sum of no-agreement incomes.

$$x_1 + x_2 + ... + x_N < P \tag{3}$$

With only two bargainers (N = 2), the bargaining problem is illustrated as in figure 1 with the one bargainer's slice on the vertical axis and the other's slice on the vertical axis. For reasons that will soon become evident, it is convenient to refer to the bargainers not as 1 and 2, but as E and O, mnemonic for even and odd, so that their slices become y_E and y_O. All efficient allocations between the bargainers are represented as points on the diagonal line PP, intersecting both axes a distance P from the origin. An agreement providing person E with an income of y_E can provide person O with an income of at most $y_O = P - y_E$. Inefficient allocations, providing less than the total available sum to the two bargainers together, can be represented by points inside of the line PP.

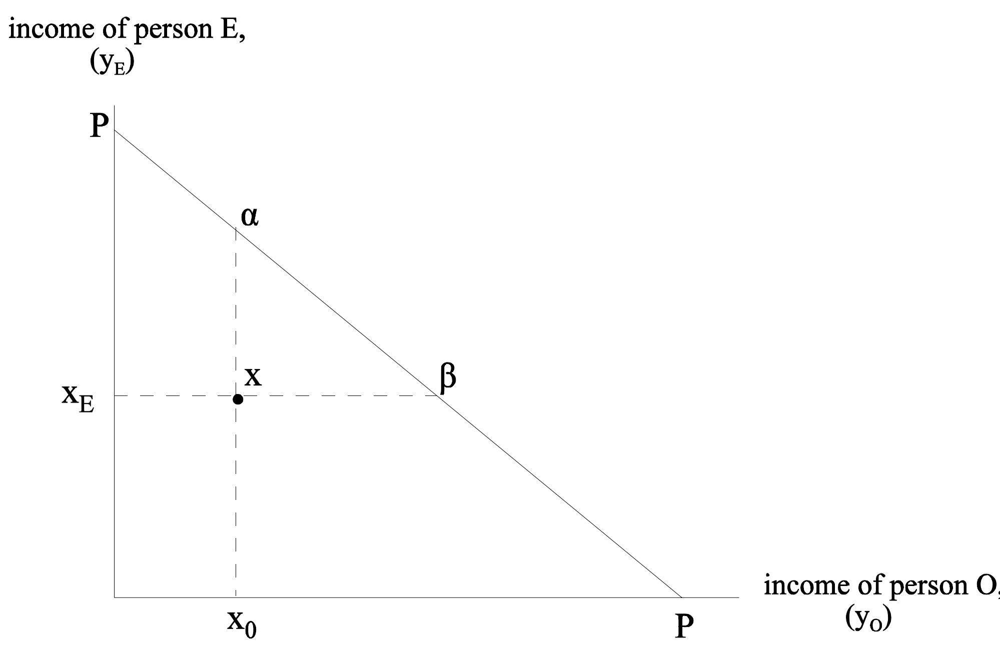
The no-agreement point is x at which person E gets x_E and person O gets x_O. Mutually-advantageous bargains, for which both parties are better off agreeing than failing to agree, are represented by points north-east of x, lying within the triangle αβx. Bargains allocate of the "surplus", S defined as $P - (x_E + x_O)$, the amount over and above what both parties can be sure of acquiring if negotiation breaks down. The parties' shares of the surplus are S_E and S_O where

$$S_E = y_E - x_E \quad \text{and} \quad S_O = y_O - x_O \tag{4}$$

Shares of the surplus need only be distinguished from shares of the pie when there are outside options. Bargains that are efficient as well as mutually-advantageous are represented by points on the outer edge of the triangle, along the line PP between α (directly above the point x) and β (directly to the right of the point x).

In a nutshell, the bargaining problem is how the parties share the surplus by agreeing on a single point on the line between α and β. We know that people do often reach agreements, sometimes quickly and easily, sometimes after considerable waste of resources and time in the process of bargaining. We have theories of bargaining based on additional assumptions about fairness or about how the bargainers are allowed to behave. For the simple case in figure 1, we do not understand how bargains are struck.

Figure 1: A Two-person Bargain



To set the scene for the discussion of bargaining models below, we examine briefly three instances where a solution to the bargaining problem has been assumed. The first is from theoretical politics where determinacy in legislation requires that bargains be struck between legislators in different political parties and between the legislature and the president. The second is from commerce where profit maximization sometimes requires that bargains be struck between suppliers and their customers. The third is from the law where the Supreme Court of Canada mandated that a bargain be struck without specifying what the terms of the bargain must be.

1) Legislature and Executive

In *Partisan Politics, Divided Government and the Economy* (Cambridge University Press, 1995), Alesina and Rosenthal assume that all political outcomes can be represented by points on a left-right continuum and that politics is a contest between two parties, Democrat and Republican, with different ideal points - θ_D and θ_R - on that continuum. For instance, it might be supposed that the choice of the tax rate is the only political issue, that the Republicans favour a low tax rate and that Democrats favour a high tax rate. To ensure that Democrats are to the left of Republicans in accordance with common political terminology, the indicator θ might be interpreted as the ratio of post-tax to pre-tax income, in which case $\theta_D < \theta_R$. All Democrats have the same preferences. All Republicans have the same preferences. The political outcome is a point on the left-right continuum, determined simultaneously by two bargains, one within the legislature and another between the legislature and the president.

The first bargain establishes a preference for the legislature as a whole when some legislators are Democratic and others are Republican. Specifically, Alesina and Rosenthal assume that the preference of the legislature is

$$\theta_L = \theta_R(V) + \theta_D(1 - V) \qquad (5)$$

where V is the proportion of Republicans in the legislature. Within the legislature, and regardless of the preference of the President, a bargain is struck between Republicans whose first preference is the point θ_R and the Democrats whose first preference is the point θ_D to establish a legislative preference at the point θ_L in between θ_R and θ_D in accordance with equation (5).

Under the bargaining rule in equation (5), the legislative preference becomes the Republicans' first preference in the event that the Republicans win 100% of the seats in the legislature (V = 1), the Democrats' first preference in the event that the Democrats win 100% of the seats in the legislature (V = 0), and a number in between the parties' first preference - the closer to each party's first preference the larger its share of the seats in the legislature - in the event that both parties win shares of the seats (0 < V < 1). That in itself seems reasonable, but there is no explanation within the model of the specific functional form in equation (5). It might be assumed instead that θ_R and θ_D are weighted by V^2 and $(1 - V^2)$, transforming the expression $\theta_R(V) + \theta_D(1 - V)$ into the expression $\theta_R(V^2) + \theta_D(1 - V^2)$, and supplying a majority in the legislation with influence more than proportional to its numbers. Alternatively, it might have been assumed that $\theta_L = \theta_R$ if V > ½ and $\theta_L = \theta_D$ if V< ½, indicating that a bare majority in the legislature is sufficient for control of the preference of the legislature. The preference of the legislature, θ_L, might be connected to θ_R and θ_D by any function whatsoever, as long as extra seats in the legislature conveys extra bargaining strength. There is no explanation of why or how the value of θ_L in equation (5) emerges as the necessary outcome of rational, self-interested behaviour. Nor is it explained why bargaining never degenerates into chaos. It is simply assumed that a bargain is struck costlessly and that θ_L in equation (5) is what the bargain turns out to be.

The second bargain is between the legislature and the president, so that the outcome depends critically on whether the President is Republican or Democratic. The President is assumed to be no different in his preferences from any other member of his party. His preference, θ_P, is either θ_D or θ_D depending the party to which he belongs. The national decision, θ_N, becomes

$$\theta_N = \alpha\theta_P + (1- \alpha)\theta_L \qquad (6)$$

where α is the postulated, but entirely unexplained, bargaining strength of the President.

In their treatment of bargaining, the Alesina and Rosenthal model has one principal virtue and one corresponding vice. Its virtue is the explicit recognition that democratic politics is a complex interplay between voting and bargaining. Typically, voting theory allows no place for bargaining because the first preference of the median voter is expected to prevail.[1] Alesina and Rosenthal treat bargaining as an essential and indispensable ingredient of democratic politics, an ingredient without which democratic politics would not work at all. We vote for the president. We vote for legislators. Whoever we elect must strike a bargain among themselves before any policy can be adopted. In practice, voting takes place within political parties as well as among elected politicians. The Alesina and Rosenthal model emphasizes the latter but ignores the former by postulating political parties as pre-existing entities with given preferences on a single left-right axis.

Bargaining has an equally large a role to play in the choice by political parties of platform and programs. Platforms established at any moment of time may be altered in response to threats by dissidents to withdraw their allegiance or in the attempt to attract new supporters. It is often argued that the fundamental difference between first-past-the-post and proportional representation lies in the bargaining they provoke. First-past-the-post voting requires relatively little bargaining within the legislature or between legislature and executive, but a great deal of bargaining in formation of a party likely to win elections. Proportional representation requires less bargaining within political parties because every group of like-minded politicians can

[1]See, for instance, Plott, R.C., "A Notion of Equilibrium and its Possibility under Majority Rule", *American Economic Review*, 1978, 146-160. Bargaining is also circumvented in the citizen-candidate model of voting. See Osborne, M and Slivinski, A, "A Model of Political Competition with Citizen-Candidates", *Quarterly Journal of Economics*, 1996, 65-96.

establish a political party of its own. Proportional representation requires more bargaining within the legislature in forming a governing coalition, in holding it together and in deciding among the alternatives for public policy.

The corresponding vice in the Alesina and Rosenthal model is its gratuitous resolution of the bargaining problem. Bargains are struck in accordance with the parties' bargaining power, a concept never justified or explained. Within the legislature, each party's bargaining power is proportional to its number of seats. Between legislature and President, bargaining power depends on a parameter pulled out of thin air. The problem with this analysis is not that it is wrong or useless, but that it is ungrounded in rational, self-interested behaviour. Bargaining power is postulated, not rationalized or explained.

On the strength of their assumptions about how bargains are struck, Alesina and Rosenthal develop interesting and insightful propositions about voting and the formation of public policy. They explain, for example, how rational voters' choices between Republican and Democratic candidates for the legislature are influenced by whether the President is Republican or Democratic. Their assumptions about bargaining remain unexplained and ungrounded in any persuasive explanation of how rational and self-interested people come to agree. Like any theory, this theory of political behaviour might verified by the accuracy of its predictions, but that line of defense is only available if we are not interested in bargaining *per se* and 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.

2) Make or Buy

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 the General Motors company 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 investments are unverifiable. Hart's model is quite complicated, but a stripped down version is sufficient to focus on the 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 π_F and π_G. If they amalgamate into one large firm, its profit would be π_T. 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' profit, π_F and π_G, in the absence of cooperation, the profit, π_T, of the amalgamated firm and the profit function, π_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, g or π_C (f, 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

$$T = \pi_C(f, g) - [\pi_F + \pi_G] - f - g \quad (7)$$

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

$$S = \pi_C (f, g) - [\pi_F + \pi_G] \quad (8)$$

Two key behavioural assumptions are now 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, π_{FC} where C is mnemonic for cooperation, becomes

$$\pi_{FC} = \pi_F + S/2 - f \quad (9)$$

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

$$\pi_{GC} = \pi_G + S/2 - g \quad (10)$$

Hart's principal proposition is that, without verification, both relation-specific investments are too small and total profit, π_C, is less than it might be. From equation (7), it follows at once that the joint profit of the two companies together is maximized when f and g are chosen so that

$$\delta[\pi_C (f, g)]/\delta f = 1 \text{ and } \delta[\pi_C (f, g)]/\delta f = 1 \quad (11)$$

From equations (9) and (10), it follows that each firm maximizes its own profit within the cooperative arrangement when f and g are chosen so that

$$\delta[\pi_C (f, g)]/\delta f = 2 \text{ and } \delta[\pi_C (f, g)]/\delta f = 2 \quad (12)$$

Comparison of equations (11) and (12) shows that both f and g are lower in the cooperative arrangement than would be warranted to maximize the true profit, T, as long as the function π_C(f, g) is concave in f and g. If f and g were observable and verifiable by both firms, they would be chosen in accordance with equation (11). Any other combination of f and g would be inefficient. Any other combination of f and g would make both firms potentially worse off no matter how the surplus is allocated between them. Thus it easily could 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. On the other hand, Hart's purpose is not to explain bargaining, but to articulate the impediments to cooperation between firms in the absence of a complete set of market-clearing prices. His assumption that bargaining would be efficient but for unverifiable or unshared information directs the readers attention to other phenomena that are his main concern.

3) 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 a separate country with French as the only official language. The governing party in the province favours secession, but, so far, two referenda in the province have failed to produce a majority for secession. There is a general understanding that 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 neither raised nor answered.

The judgment in this case exemplifies the commonly-held view, seen as the epitome of wisdom among Canadian political scientists, 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. It is almost an article of faith that negotiation yields a determinate outcome. This view cannot be entirely wrong. Negotiation does yield a determinate outcome much of the time. The world's work could not be accomplished otherwise. Yet negotiation does sometimes break down into non-agreement or outright violence. That bargains are often struck is an empirical regularity rather than a precise consequence of rational calculation by the parties involved.

*****

Four features of these bargains should be noted: First and most important is the implied analogy between bargaining and production. Just as given inputs of labour and land are assumed to yield a determinate amount of some good, such as wheat, in a classic production function, so too is the given bargaining situation assumed to yield a determinate division of the surplus. When employing the production function in general equilibrium models or in analyzing the incidence of taxation, we do not probe behind the production function or ask what it is about the labour and land that enables them to be combined into wheat. Nor do these models of bargaining in politics, business and law probe into the innards of the bargaining process. It was not asked how or why bargaining works. Instead, for bargaining as for the technology production, it was postulated that the outcome is somehow predetermined by the initial conditions. There is, however, an important difference between the two. Production is impersonal and, presumably, in accordance with the laws of nature governing the material world. We may not know the shape of the production function, but we are confident there really is one. By contrast, bargaining is interpersonal in a domain where physical laws alone are not sufficient to determine the outcome completely.

Second, the technology of bargaining is pulled out of thin air. The very same absence of grounding in rational self-interseted behaviour that was noted in the bargaining procedure of equation (5) in Alesina and Rosenthal's political model is also evident in Hart's model of cooperation among firms. To assume an equal split of the joint surplus in equation (8) is no less arbitrary. That the surplus is shared equally between cooperating firms has little more than symmetry to recommend it.

Third, in imposing assumptions about bargaining, the authors are not seeking to explain bargaining *per se*. Rather, they are seeking to explain something else - voting patterns in one case and the structure of contracts in the other - in circumstances where bargains must be struck. For their purposes, any of a large range of bargaining mechanisms would probably do. That being so, they might as well adopt the simplest mechanism, for their models are complicated enough in other respects. "What would you do?", the critic might be asked. "We are seeking to explain voting or contracting as the case may be. If you don't like our assumptions about of bargaining, go invent more realistic and more useful assumptions, or give up trying to explain these phenomena altogether?" Without a wonderful new theory of voting or contracting up his sleeve (and I have none), the critic is stumped by this reply, except for the lame rejoinder that my ignorance does not make your theory true. The main point of this imaginary dialogue is that the bargaining models in Alesina and Rosenthal and in Hart should be looked upon as unexplained empirical regularities embedded in models of rational, self-interested behaviour. Neither model contains an explanation of why or how people bargain as they are assumed to do.

Fourth, in both models, bargaining is costless. No part of the surplus is used up in the process of agreeing how to allocate the surplus between the contending parties. Convenient as it is in some contexts, this assumption is strictly-speaking false. Bargaining is always costly in lawyers' fees, in the wastage of bargainers' time, in delay reaching agreement and in the breakdown of negotiation leading to the abandonment of potentially profitable enterprises.

The Environment of Bargaining

Models of bargaining are intended to explain how bargains are struck, to justify the presumption in the preceding examples that unique bargains are inherent in the circumstances of the bargainers, and to render determinate what Edgeworth in the quotation at the outset of this paper saw as indeterminate and beyond the scope of economic analysis. The principal models will be discussed in the next section where our primary concern will be to judge whether and to what extent the claims for those models are justified. As a preliminary, it may be helpful to list characteristics of bargains, though not all bargains have all of these characteristics in common.

First and foremost, a bargain is a *voluntary sharing of a pie.* Person E and person O bargain over the allocation of a sum of money, P. A bargain is an agreement between them assigning an amount y_E to E and amount y_O to O. Naturally, each bargainer wants his share to be as large as possible. The bargaining problem is to explain how these shares are determined. One may think of the pie as owned jointly by the bargainers, E and O, but of no use to either person until they agree about how it is to be shared. To say that the sharing is voluntary implies a prior agreement that the pie belongs to E and O but to nobody else. In general, there may be more than two bargainers, and the pie need not be money. One would expect that, the larger the number of stakeholders in the bargain, the more difficult it would be come to determine each bargainer's share. Bargaining could also take place over the allocation of a pile of things, as when children divide up parents' possessions. Neither of these considerations will be discussed in this essay.

Typically, in a fresh bargain, each party would have an *outside option*. As an alternative to the bargain, E could use his resources to earn an amount x_E and O could use his resources to earn an amount $x_{O.}$ Thus, in this context, the bargain would not really be about the sharing of the entire pie, P, but about the sharing of the surplus, $S = P - (x_E + x_O)$, over and above the combined value of the outside options. Shares of the surplus - $S_E = P - x_E$ and $S_O = P - x_O$ - would both have to be positive as long as both parties retain the option of not embarking on the common venture within which the bargain is required. On the other hand, once the common venture is under way, the outside options may or may not be preserved. They may, in effect, become sunk costs that cease to constrain the bargain. A central objective of legally-binding contracts is to forestall this possibility.

Bargains differ in *durability.* A distinction may be drawn between a one-shot bargain that must be struck immediately or not at all, and a durable bargain that remains available tomorrow if it is not struck today. The pie may disintegrate if it is not shared immediately, or it may persist - perhaps forever as in some models of bargaining - to be shared by the bargainers whenever a bargain is struck. The pie may shrink over time until, eventually, there is nothing left, or, almost equivalently, the bargainers' present values of the pie may diminish even though the pie itself does not. The picture of bargaining Figure 1 above takes no account of the temporal features of bargaining. The diagram is instructive, but not the whole story.

Failure to agree today may lead to *postponement* rather than termination. Some ventures must be undertaken at once or not at all. Other ventures need not be dismantled if the parties to the bargain wish to continue bargaining in the hope that a deal can be struck later on. Delay may inflict one or both of two costs onto the bargainers, loss of revenue from the venture and *foregone interest*. Suppose y_E is person E's slice of the pie as it would be if the bargain were struck at once and r_E is his rate of discount (for instance, his rate of return on bank deposits or other ventures). His bargaining cost from a one year delay in reaching an agreement is his interest forgone, approximately $r_E\ y_E$. If parties E and O have different discount rates, r_E and r_O, reflecting different opportunities for investing their money, their interest costs per year of delay become $r_E y_E$ and $r_O y_O$. These costs of disagreement may be quite small because failure to agree rarely persists over significantly long periods of time. Nevertheless, a difference between bargainers in their cost of delay may influence the allocation of the pie in the final bargain.

The other, typically larger, cost of delay is a *loss of revenue from the venture*. Delay in reaching an agreement may stop a stream of earnings until such time as the agreement is reached, may reduce the size of the pie so that the eventual bargain becomes about the allocation of the remainder, or may impose costs per unit of time on both parties leaving the pie (evaluated on the day the agreement is reached) undiminished. Costly delay is exemplified by labour contracts where the alternative to settlement is a strike. The cost of a strike is the sum of the workers' the loss of wages during the strike and the management's corresponding loss of profit. These cost are different in kind, and typically larger than the cost of interest foregone.

A deal with costly delay is characterized by annual costs c_E to E and c_O to O for as long as it takes to reach an agreement, though the surplus, S, may remain permanently available whenever the deal is finally struck. With a delay of T years, the total costs of delay are $c_E T$ to E and $c_O T$ to O, and their net gains from an allocation S_E and S_O are reduced to $S_E - c_E T$ and $S_O - c_O T$. An important difference between a postponable deal and deal with costly delay is that the net returns in a deal with costly delay must eventually go negative. A lengthy strike may leave both workers and their employer significantly worse off than if they had agreed to the other's terms immediately. An agreement after a delay of T years may be worse *ex post* for both parties than no deal at all. Mutual stubbornness can convert a surplus into a burden.

Bargains differ in the opportunity for and the credibility of *threats*. The outcome of bargaining may depend very much on the parties' ability to threaten one another. Two types of threats may be distinguished, the threat of non-agreement and the threat of punishment. One may threaten not to agree to any bargain unless one's share of the pie exceeds some given amount. Alternatively, one may threaten one's partner with some harm over and above non-agreement, where, typically, that harm can only be procured at the cost of harming oneself to some extent too. Of course, non-agreement is itself a kind of punishment, but one or both parties may have punishments over and above that.

Bargains may or may not have an *imposed sequence of speech*. An imposed sequence of speech is a specification of the order of speech, the time that must elapse between utterances and the kind of things bargainers are permitted to say to one another. Sequence plays no role at all in some models of bargaining, but it is central and fundamental in others. Discussion of sequence is best put off until the models themselves are examined.

Bargains differ in the rules for *termination.* When two bargainers fail to agree, the process of bargaining may be terminated by either party unilaterally, by unanimous consent of both bargainers, or not all. When one party can terminate bargaining unilaterally, the prize may accrue automatically to the other party, or the prize may evaporate. Outside options may or may not be preserved. Options that were available to the bargainers as alternatives to participation in a venture may or may not remain available in the event of the bargainers' failure to agree. Failure to agree may leave both parties are left worse off than if there had been no bargaining relation at all.

Consider a potential partnership between an engineer and an accountant in a venture that requires both skills. The partnership may fail to materialize because the would-be partners cannot agree on the sharing of the profit from the venture. If the partnership is not established for want of agreement on the allocation of the profit, the engineer and the accountant might still earn what they would have earned if the partnership had never been considered. On the other hand, substantial resources may have been invested by both parties in the attempt to establish the partnership, or options that had been available before the partnership was contemplated may have disappeared once the attempt to establish a partnership is finally abandoned. E*x ante*, such a partnership requires a sharing of the surplus, $P - (x_E + x_O)$, as defined above. *Ex post*, the bargain may extend over the entire pie, P, if termination occurs at a time when the outside options are no longer available.

Some Models of Bargaining

Economists' willingness to postulate a simple bargaining equilibrium in politics or in commerce may be attributable in part to the persuasive power of established bargaining theories. Equilibria emerging from these theories are rarely as simple as those postulated above, but the simple procedures - such as splitting the difference or supposing legislative outcomes to be proportional to numbers of seats held by political parties - may be justified as convenient first approximations to equilibria proved elsewhere to exist. Four models of bargaining will be discussed. The first is based on a willingness of bargainers to make *concessions*. The second is based on a common recognition of *fairness* among the participants in the bargain. The third and fourth are based on imposed bargaining *procedures*. In the third model, bargainers alternate offers for as long as necessary until some offer is accepted. In the fourth model, bargainers make demands simultaneously in circumstances where the pie is lost if the demands are not compatible. As we shall see, the theories do yield equilibria within the confines of their assumptions. Bargaining becomes determinate just as prices, production and allocation of goods and services are determinate in a competitive economy. The question at hand is whether and how closely bargaining theories correspond to actual bargaining as experienced in practice. Do the models capture the essence of bargaining, or are critical and essential features of bargaining assumed away? Is the existence of a bargaining equilibrium established on the strength of realistic and compelling assumptions, or is the gap between assumptions and implications so narrow that the theories amount to little more than an elaboration of the assertion that bargains are somehow struck?

1) Concessions Proportional to the Cost of Disagreement: Zeuthen, Hicks and Cross

The two principal models of this sort were developed by F. Zeuthen in *Problems of Monopoly and Economic Warfare* (George Routledge and Sons,1930) and by J. R. Hicks in *The Theory of Wages* (Macmillam, 1932). Both models are about bargaining between workers and their employer, but the structure of these models is more general. Parties E and O bargain over the sharing of a surplus, S. If they can agree on shares, y_E and $y_{O,}$ the pie is allocated accordingly. If they fail to agree, they must bear costs proportional to some variable, b. In Zeuthen's model, b is the probability of conflict in the absence of an agreement. In Hick's model, b is the duration of a strike. For any given b, the costs to parties E and O of a failure to agree are A_Eb and A_Ob respectively, where A_E and A_O are parameters reflecting each bargainer's intensity of harm from a failure to agree. The key assumptions in both of these models are that i) in the absence of agreement, b automatically grows until the entire surplus is eaten up in bargaining cost, and ii) each bargainer concedes to the other an amount equal to the harm he would experience if no agreement were reached. In other words, the parties to the bargain agree on shares y_E and y_O such that

$$y_O = bA_E \qquad \text{and} \qquad y_E = bA_O \tag{13}$$

when b grows to the point that

$$bA_E + bAo = S \tag{14}$$

indicating that the entire surplus, as it would be if agreement could be reached without delay, is wiped out. It follows immediately that the parties' shares of the surplus become

$$y_E = SA_O/[A_E + A_O] \quad \text{and} \quad y_O = SA_E/[A_E + A_O] \tag{15}$$

Each person's share of the surplus is proportional to the other person's cost per unit of harm from disagreement.

In both models, employees and their employer bargain over the setting of a wage, w, in a range between an upper limit w^H and a lower limit w^L. The surplus is $w^H - w^L$ of which $w - w^L$ accrues to workers, and $w^H - w$ accrues to the employer. In Hicks' model, w^H is the most the employer can afford without actually going broke, w^L is the least workers would accept in preference to seeking employment elsewhere and w is agreed upon in the light of both parties' costs of a strike that is long enough to wipe out the surplus altogether. Hicks draws what he calls an "employer's concession curve" and a "union's resistance curve" (Hicks, *The Theory of Wages*, page 143) which cross at some duration of the strike.

In the Zeuthen model, the failure of employees and their employer to agree leads to "conflict", the exact meaning of which is not spelled out in detail. It is sufficient for the model that there be some high wage, w^H, for which the employer would be indifferent between accepting that wage and "conflict" and some low wage, w^L, for which the union would be indifferent between accepting that wage and "conflict". Zeuthen goes on to assume that each bargainer's concession to the other equals its expected harm from conflict when antagonism between the bargainers rises to the point where the sum of their expected harms from conflict eats up the entire surplus. This process can be represented by equations (13) above where b is interpreted as the probability of conflict and A_E and A_O can be interpreted as the bargainers' harm from conflict.

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, in these models as well as in other models to be discussed below, 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. Bargains are made determinate within these models, but only by ignoring essential features of the world where bargains are struck.

Nor is there an explanation of the usual sequence of concessions leading to agreement in actual bargains with small concessions by one bargainer followed with small concessions by the other until some price between the original bid-price and the original ask-price is agreed upon. Zeuthen talks about a sequence of concessions, but his model contains no explanation of the magnitude of each concession and fails to allow for the possibility that the final agreement is conditioned upon the history of bidding as well as on the initial values of the bargainers' harms from conflict. Hicks does not even try to account for the sequence of concessions within his formal model. Yet, as Bishop pointed out in "Game-Theoretic Analysis of Bargaining" (*Quarterly Journal of Economics*, 1963), the process of negotiation would be nothing more than play-acting if the ultimate bargain were entirely inherent in the initial characteristics of the participants and environment of the bargain.

Genuine concessions are modeled by Cross in "A Theory of Bargaining" (*American economic Review*, 1965). Both parties' rates of concession are rendered determinate by the principle that delay is costly and that, if you do not concede quickly, then I must. 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 the bargainers' concessions are what they are or 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.

2) Fairness: The Nash Bargaining Solution[2]

As illustrated in figure 1, a sum of money, P, is to be divided between two people, E and O. If no agreement is reached, person E automatically acquires an amount x_E and the person O automatically acquires an amount x_O , where $x_E + x_O < P$ (for there would otherwise be nothing to bargain about). A surplus, S (equal to $P - x_E + x_O$), is left unallocated, and therefore wasted, unless some agreement can be reached. The Nash bargaining solution supplies y_E to person E and y_O to maximize the value of the product

$$[u(y_E) -u(x_E)][v(y_O) - v(x_O)] \qquad (16)$$

where the functions u and v are the Neumann-Morgenstern utility of income functions of persons E and O.

The Nash bargaining solution is called a "solution" rather than a postulate or empirical regularity because it is the solution to a very special mathematical problem. Nash derives the solution as the consequence of eight postulates about the nature of utility and the circumstances of the bargain. This is not the place to review the formal proof, but three important aspects of that proof might be mentioned here:

First, there is little distance between Nash's axioms and his principal theorem allocating

[2]John Nash, "The Bargaining Problem", *Econometrica*, 1950, 155-62, reprinted, together with the papers on bargaining by Cross, Bishop, Nash and Schelling cited here and with a useful selection of other papers on bargaining written prior to 1975, is reprinted in Oran R. Young, *Bargaining: Formal Theories of Negotiation*, University of Illinois Press, 1975.

the disputed prize to maximize the expression in equation (16) above. Nash's eighth axiom[3] is comes very close to the principle of difference-splitting in the examples at the outset of this paper. The axiom is that the bargain must supply both participants with equal utilities whenever the range of options is symmetric. Symmetric in this context means that if it is feasible within the bargain to provide you with a bundle of goods B and to provide me with a bundle of goods C, then it must also be feasible to might be provide you with a bundle of goods C and to provide me with a bundle of goods B. By itself, that postulate does not govern the choice of the maximand when the range of options over which people bargain is not symmetric, but it carries much of the weight of Nash's proof.

The equal division of the joint surplus in Hart's model of bargaining between associated firms would seem to be warranted by Nash's eighth axiom alone, with no contribution from any of Nash's other axioms, for the bargain is about money in a context where income and utility are one and the same. Translated as statements about the terms in the expression in equation (16), Hart's model requires that $u(y_E) = y_E$, $v(y_O) = y_O$ and $x_E = x_O = 0$, reducing the expression in equation (16) to $y_E y_O$ which - since $y_E + y_O = S$ - is automatically maximized when $y_E = y_O$. In Hart's context, Nash's other axioms become redundant.

That is not quite true of Alesina and Rosenthal's political compromise in equations (5) and (6) between the President and Congress and among legislators from different political parties, for Alesina and Rosenthal's notion of bargaining power has no counterpart in Nash's model. Alesina and Rosenthal's analogue of a bargainer's utility in equation (16) is the distance between his first preference and the actual political outcome on the left-right scale. Nash's eight axiom is sufficient to generate Alesina and Rosenthal's compromise in the special case where the President and Congress are assumed to have equal bargaining power and where every legislator's bargaining power is the same. Otherwise, if some parties have more bargaining power than others, the political compromise no longer conforms to Nash's eighth axiom and equation (16) is automatically violated.

Second, the Nash bargaining solution carries the odd implication that the allocation of the pie depends critically on the bargainers' risk aversion. To make this point, there is no harm in supposing i) $x_E = x_O = 0$ so that the Nash bargaining solution boils down to the choice of y_E and y_O to maximize product $u(y_E)v(y_O)$, and ii) the utility functions are of the form $u(y_E) =(y_E)^\alpha$ and $v(y_O) = (y_O)^\beta$ where α and β are characteristics of the parties E and O. Confined to the range between 0 and 1, the parameters α and β can be thought of as measures of risk aversion. For example, the more risk averse person E happens to be, the smaller is his value of the parameter α

[3]Nash's eighth axiom is "If S is symmetric and u_1 and u_2 display this, then c(S) is a point of the form (a,a). That is, a point on the line $u_1 = u_2$." Nash, J., "The Bargaining Problem", op.cit. P.57. In ths axiom, S refers to the set of all possible outcomes, u_1 and u_2 refer to the utilities of the bargainers, 1 and 2, and c(S) is the set of fair outcomes.

and the more favourable the odds he requires for a gamble to be acceptable.[4] Person E accepts a fair gamble if $\alpha = 1$ but not if α is less than 1. It follows at once that, when the product uv is maximized over all y_E and y_O such that $y_E + y_O = P$, the allocation of the pie becomes $y_E = [\alpha/(\alpha + \beta)]P$ and $y_O = [\beta/(\alpha + \beta)]P$. The less risk averse you are, the greater is your share of the pie. In some contexts, this implication may not seem unreasonable. But if the Nash bargaining solution is to be interpreted as a principle of fairness, the implication is very peculiar indeed. If you and I are bargaining over the allocation of a dollar between us, it is hard to see why I *should* be entitled to the larger share for no other reason than that I am the more willing to bear risk. That is not what most people mean by fair.

This peculiar implication of the Nash bargaining solution might be defended by a reinterpretation of the parameters α and β. It might be argued that these parameters reflect marginal utilities of income rather than risk aversion. For any given income, y, the parties' marginal utilities of income are $\alpha(y)^{\alpha-1}$ and $\beta(y)^{\beta-1}$. Their ratio is $(\alpha/\beta)(y)^{\alpha-\beta}$ which is greater

---

[4]A person with an income of \$100,000 is offered a gamble with equal chances of gain or loss. A coin is tossed, the person wins if it comes down heads and he loses if it comes down tails. If he wins, he receives \$50,000. If he loses, he must pay a somewhat smaller amount, \$50,000 less R, where R is his premium for accepting the gamble. The minimal premium, R, for which the gamble would be accepted may be taken as an indicator of a person's degree of risk aversion. The more risk averse one is, the larger the minimum acceptable R must be. Consider a group of people whose utility of income functions are all of the general form $u(y) = y^\alpha$ but with different values of α. Within that group, a person's degree of risk aversion can then be measured by the value of α in his utility function because, the smaller α, the larger R must be.

In choices among risky options, a risk neutral person seeks to maximize expected income. A risk averse person does not maximize expected income, but, since all choice may be interpreted as the maximization of something, utility is *defined* as whatever it is one can be interpreted as maximizing in one's choice among risky options. Everybody - whether be risk neutral, risk loving or risk averse - seeks to maximize the expected value of some utility function, and a person's response to risk is automatically reflected in the shape of his utility function. Consider a person with a utility function $u(y) = y^\alpha$ for some value of α. Suppose his initial income is \$100,000 and he is offered a gamble at fifty-fifty odds paying \$50,000 if he wins and requiring him to pay \$(50,000 - R) if he loses. The smallest R for which he accepts the gamble - the R for which he is indifferent between accepting and not accepting - is identified by the

$$(100{,}000)^\alpha = \tfrac{1}{2}(100{,}000 + 50{,}000)^\alpha + \tfrac{1}{2}(100{,}000 - 50{,}000 + R)^\alpha$$

It is immediately evident that R is an increasing function of α . From inspection of the equation, it is obvious that $R = 0$ when $\alpha = 1$, signifying that the person is risk neutral. A bit of calculation shows that R equals \$10,102 when $\alpha = \frac{1}{2}$, R rises to \$13,683 when $\alpha = 1/4$, and R rises again to \$15,534 when $\alpha = 1/10$.

than 1 as long as $y > 1$ and $\alpha > \beta$ signifying that person 1 is the less risk averse. Assume that to be so. Even then, the interpretation of the parameters α and β as measures of the marginal utility of money is questionable in this context because utility is being defined up to a linear transformation and because *some* linear transformation of the functions would raise person 2's marginal utility of income above that of person 1. Marginal utility of income has little or no significance in this context.

This difficulty may be seen as an instance of a larger problem. Among the postulates of the Nash bargaining solution is the Neumann-Morgenstern cardinalization of utility as a reflection of one's willingness to bear risk. In principle, one's utility of income function is elicited by a long series of questions of the form, "Do you prefer this to that?", where "this" is a sure income and "that" is a gamble. Essential to this cardinalization of utility is the presumption that a person who may be unwilling to accept fair gambles of income is always prepared to accept fair gambles of utility, for that is how utility is defined. Thus, the more risk averse one is, the better the odds he demands before accepting a gamble and the more concave his utility function must be. The larger problem, of which our difficulty with the Nash bargaining solution is an instance, is whether and to what extent utility functions so derived can be employed in other situations. Can they be reasonably employed as ingredients of a social welfare function? Can they reasonably serve as building blocks for a model of fair division? The argument in the preceding paragraphs suggests that they cannot.

Third, the Nash bargaining solution is disturbingly vague about the identity of the parties to the bargain. Consider, for example a bargain over wages between the owners of a firm and its unionized employees. Denote the owners by O and the employees by E, suppose the bargain is over a sum of money P which evaporates unless some agreement is reached and let $x_E = x_O = 0$ so that the Nash bargaining solution boils down to the choice of y_E and y_O to maximize product $u(y_E)v(y_O)$. That outcome may seem plausible and fair as long as there are only two parties involved, but suppose instead that there are two unions so that the total pie has to be divided three ways: y_O for the owners of the firm , y_{E1} for the first union and y_{E2} for the second union. Generalizing the two-party Nash bargaining solution, one might suppose that a fair choice of the shares, y_O, y_{E1} and y_{E2}, would maximize the joint product, $u(y_O)v(y_{E1})w(y_{E2})$ where v and w are now to be interpreted as the utility functions of the two unions. Plausible as it may seem at first glance, this generalization of the two-party solution carries the implication that the fair share of the owners is automatically reduced when one union covering the entire labour force is divided in two. In other words, the Nash bargaining solution carries no implication about the participants of the bargain. Does the unions' share depend on the total number of workers and the allocation of workers among the different trade unions? Does the owner's share depend on the number of shareholders or the proportion of each shareholder's income represented by his ownership of the firm? A model of fairness might be expected to supply answers to these questions, but, so far as I can tell, the Nash bargaining solution does not.

3) Structure Induced Equilibrium: The Staahl-Rubinstein Bargaining Solution[5]

The term "structure-induced equilibrium" was introduced into the literature of voting theory to circumvent the paradox of voting where any majority in a legislature could employ the vote to take what it pleased from the corresponding minority[6]. The term is employed here to indicate how imposed constraints on the process of bargaining might ultimately determine which bargain, within the feasible range of bargains, will be realized. A structure-induced bargaining equilibrium is a division emerging automatically as the outcome of self-interested behaviour by rational bargainers in response to an externally-imposed bargaining procedure where offers, acceptances and rejections must conform to a preordained pattern and with well-specified consequences of a failure to agree. How this comes about is best examined sequentially, from a very simple procedure to the progressively more complex.

Begin with a single offer. You and I are jointly entitled to a sum of $1000 if and only if we agree on each person's share. The bargaining process is radically simple. You are entitled to propose an allocation the money, so much for you and so much for me. I may accept or reject your proposal. If I accept, the money is allocated accordingly. If I reject your proposal, the money vanishes and neither of us gets anything. The outcome of this procedure would depend on how we choose to behave, but if we both behave "rationally" (in a sense to be specified below) then the outcome is as follows: You offer me some very small amount, such as a penny, reserving the remaining $999.99 for yourself. I accept because a penny is better than nothing and because the imposed bargaining procedure denies me the opportunity of making any alternative proposal. . Having to choose between a penny and nothing, I choose to take the penny. Of course, the outcome is reversed if I get to make the initial offer. Either way, the outcome is lopsided but determinate.

Note how the imposed structure of bargaining determines the outcome. The structure is that each of us may speak once and only once and in a prescribed order. Speaking first, you may state a number. You may say "$y", where $y is my share of the pie on the understanding that the remainder is for you. The choice of y is up to you, but, having spoken, you must never speak again. Then, I may say "yes" or "no", nothing more, and I must thereafter remain silent. The imposed procedure governs the ordering and content of speech. It is as though we would be

---

[5]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.

[6]K. A. Shepsle and B. R. Weingast, "Structure -Induced Equilibrium and Legislative Choice", *Public Choice*, 1987, 503-19.

severely punished for talking out of turn. It is as though we are both gagged except for one brief moment when we play our parts in the bargaining process.

The meaning of rationality in this context is that we each do what is best for ourselves at the moment, regardless of what has transpired before. It means that I do not reject the penny out of spite. I am not vindictive. I do not spurn the penny for the pleasure of denying you your ill-gotten $999.99. History, on this assumption, is indeed bunk. I take what I can get now.

A second bargaining procedure is only slightly more complex. Suppose that the $1000 is stacked in three piles of $250, $400 and $350 and that we bargain over these piles one by one. For each pile the bargaining procedure is the same as described above, except that I get to make the offer on the first plie, you get to make the offer on the second and, finally, I get to make the offer on the third. Again, it is obvious what happens. As long as there is no extra communication between us, I get all but a penny of the first pile, you get all but a penny of the second, and then I get all but a penny of the third. Thus, ignoring the pennies, I come away from the table with $600 [250 + 350] and you come away with the remaining $400. Dividing the one big bargain into a series of little ones may permit a genuine sharing of the pie.

This second procedure is only of interest as the introduction to a third. Now we bargain over the entire pie - the entire $1000 - but we make alternating offers in three rounds of bargaining where the pie diminishes after each round until, after the third round, there is nothing left: In the first round with the full $1000 on the table, I get to make the offer - so much for you and so much for me - which you may either accept of reject. If you accept, the $1000 is allocated accordingly. If you reject my offer, the pie is reduced by $250, leaving only $750 on the table. In the second round with only $750 on the table, you get to make the offer which I may either accept of reject. If I accept your offer, the $750 is allocated accordingly. If I reject your offer, the pie is reduced by an additional $400, leaving only $350 on the table. Finally, in the third round with only $350 left on the table, I get to make the offer again which you may either accept or reject. If you accept my offer, the $350 is allocated accordingly. If you reject my offer, then the remaining $350 is swept off the table and the bargaining is over because there is nothing left to bargain about.

In the light of the second bargaining example, it is obvious what must happen in the third. In the very first round of bargaining, I offer you offer you $400 (plus a penny) and you accept, leaving me with the remaining $600. The reasoning is by “reverse induction”. Suppose no agreement had been reached until the third round when there remains only $350 on the table and when I alone am entitled to propose an allocation between us. Clearly, I can offer you as little as a penny, for a refusal on your part would eliminate what is left of the pie leaving us both with nothing. As in the first example, you prefer a penny to nothing. Now go back to the second round when you get to make the offer with $750 remaining on the table. You cannot offer me less than $350 because I can acquire $350 by holding out for the third round, but you need not offer me more. You offer you $350 plus a penny, keeping the remaining $400 less a penny for yourself. Finally, consider the first round of bargaining when I get to make the offer with the full $1000 on the table. I cannot offer you less than $400 because you can acquire that for yourself by waiting until the second round, but I need not offer you more. I offer you \$400, keeping the remaining \$6000 for myself, and you, in your own interest, accept. Together, the imposed structure of bargaining and rationality postulate have yielded a unique outcome.

Of course, there is nothing special about three periods. A pie, P, available to the bargainers at time 0 may disappear bit by bit over a total of T rounds of bargaining, diminishing by an amount p_t after the t^{th} round until there is nothing left. Necessarily,

$$P = \sum_{t=1}^{T} p_t \qquad (17)$$

All that is left of the pie at the beginning of the last round is p_T , all that is left at the beginning of the second to last round is $p_T + p_{T-1}$, and so on. Suppose once again that party E and party O make alternate offers to one another, party O in the first round, party E in the second, party O in the third and so on. Assuming T to be an odd number, party O must make the final offer in the last round, T. As above, if bargaining ever got to the last round, party O would offer party E virtually nothing and would take all of p_T for himself. Entitled to make the offer in the preceding round (which is even), party E would offer no less than p_T to party O because party O can acquire that much by waiting until the next round, but party E need not offer more. Since the value of the pie at the beginning of the second to last round is $p_{T-1} + p_T$, party E offers p_{T-1} to party O and, as his offer is accepted, keeps the remaining p_T for himself. Similarly, entitled to make the offer in the preceding round, T-2, party O offers p_{T-1} to party E, and, since his offer is accepted, keeps the remaining $p_{T-2} + p_T$ for himself. By this logic it follows that the parties divide the pie at the first opportunity in the year 1 and that their slices, y_E and y_O, become

$$y_E = \sum_{t \text{ even}} p_t \qquad \text{and} \qquad y_O = \sum_{t \text{ odd}} p_t \qquad (18)$$

At each round of bargaining, the offer is the least that the other party is prepared to accept. One need not offer more, but cannot offer me less because it would then be in the interest of the other party to wait until the second round of bargaining when he would be entitled to make the offer, and he in turn would offer the least that one would rationally accept. The bargaining equilibrium is that, as shown in equation (18), each party acquires the sum of the reductions in the pie in the rounds when he is entitled to make the offer.

Time plays no role in this equilibrium. Sequence is critical, but not time itself. Nothing is specified so far, and nothing need be specified, about the amount of time between rounds. Rounds may be separated by years, weeks, seconds or nanoseconds. The story is the same because impatience is abstracted away. The parties to the bargain care about their shares of the pie, but are assumed to be indifferent about when a given income is acquired. A share of - say - \$350 is as welcome in the third round as in the first.

The model is easily modified to account for time-preference. Time preference is introduced by substituting impatience for the shrinking of the pie. The pie no longer shrinks from year to year. Instead, the size of the pie remains the same, but the present value of the pie, as assessed in year 1, diminishes with the length of time until the pie is actually acquired. Rounds are separated by years - the first round in the first year, the second in the second year and so on - with each offer and rejection occurring almost simultaneously on January 1. Think of the offer in each year t as made at 9 AM of January 1, to be followed at 10AM on that same day by an acceptance of a rejection, and then by silence until 9AM on January 1 of the following year. Determinate shares are generated by essentially the same process of offer and counter-offer as in the preceding example, but special care must be taken to account for the fact that the present value of a given share of the pie never falls to zero no matter how far ahead the bargain is struck and the pie is actually divided. A unique bargain can still be identified by backward induction with the aid of an additional assumption: that, if no deal has been struck by the year T (which may be very far ahead), the cycle of offer and counter-offer is brought to an end by a compulsory allocation of the pie with an preassigned shares to each party.

As shown in Appendix 1, the deal is once again struck in the very first period. Person O proposes a division of the pie in the year 1, and person E accepts, avoiding the waste from delay if the bargaining were allowed to continue. Person O's proposal is dependent on both persons' discount rates. The outcome depends very little on the preassigned shares in the compulsory allocation at the termination date T as long as everyone is rational and T is very far ahead. In the special case where the termination date, T, is infinitely far ahead, person E's slice of the pie - offered by person O in the very first round of bargaining and immediately accepted by party E - becomes

$$y_E = Pr_O/[r_O + r_E] \tag{19}$$

where r_O and r_E are the parties' discount rates and where party O's slice, y_O, is automatically equal to $P - y_E$.

A crude justification of this formula rests on the idea that neither party must stand to gain by waiting in the hope of getting a better deal later on. Party E's cost of waiting one year is $r_E y_E$, the interest forgone in acquiring an amount of money y_E next year rather than this year. Party O's cost of waiting is $r_O y_O$ which is equal to $r_O(P - y_E)$. Think of the bargain as equating the costs of delay, so that the threat of delay in supplies neither party with an advantage over the other. Appendix 1 contains a formal derivation of the equilibrium bargain, including a specification of when equation (19) is strictly valid and when it is merely an approximation. Note the similarity between Hicks and Zeuthen's bargaining solution in equation (15) and the assignment of bargainers' shares in the Nash bargaining solution when people differ in their willingness to bear risk. Both solutions can be seen as special cases of the general principle that each bargainer's share of the surplus is proportional to the other's cost of disagreement. The forms of the equations are the similar, each party's share depending on the other's characteristics: the cost of a strike, harm from conflict or cost of delay.

As constructed so far, these models of structure induced equilibrium are somewhat deceptive. The sequence of offers and counter-offers looks at first glance like what happens when buyer and seller exchange offers in, for instance, the sale of a house, but there is an important difference. Offers and counter-offers in the sale of a house are genuine. The offers are actually made, one after another, and nobody is quite certain until the process is complete that it will, in fact, be completed. Offers and counter-offers in the Staahl-Rubinstein procedure are theoretical, even bogus. They are offers and counter-offers that could, in principle, be made, but are not actually made because the process terminates before real bargaining, as the world understands the term, begins. These models are not of bargaining, but of the avoidance of bargaining, of how bargaining as commonly understood is circumvented.

A genuine sequence of offers and counter-offers can be generated by assuming that bargainers are less than fully informed about the range of bargains their opponents are prepared to accept. For instance, in bargaining between employer and employees, there may be a lowest wage, w^L , below which employees would seek employment elsewhere, and a highest wage, w^H , above which the employer would abandon his business as unprofitable. If both bargainers knew the other's "reservation" wage and as long as $w^H > w^L$, then the difference, $w^H - w^L$, between the reservation wages is just like the pie in the preceding examples. If employer and employees do not know one another's reservation wages, a sequence of genuine offers and counter-offers may be the way for the parties to feel one another out.

In explaining how the want of information may affect the bargaining procedure, it is convenient to suppose that the employer's reservation price, w^H , is common knowledge, but that employer does not know the employees' reservation wage, w^L . Instead, suppose the employer knows w^L lies between a minimum of w^{L1} and a maximum of w^{L2}, and believes there is an equal chance of w^L lying anywhere between these limits. Suppose the employer's gain from the bargain be $A(w^H - w)$ where w is the agreed-upon wage and A is some constant. As always, the outcome of the bargain depends on the imposed bargaining procedure. With a single offer followed by acceptance or termination of bargaining, the outcome when the employees are entitled to make the offer is w^H at which the employees capture the entire surplus. The outcome is less one-sided when the employer is entitled to make the offer. The employer cannot offer as little as w^{L1} because there would be a virtual certainty of rejection. Instead, the employer chooses w to maximize the expected value of his benefit from the bargain, which is the product of the employer's value of the bargain if accepted, $A(w^H - w)$, and the probability, $(w - w^{L1})/(w^{L2} - w^{L1})$, of acceptance. It follows at once that the employer's optimal offer of w is $(w^H + w^{L1})/2$. The employer makes an offer half way between his reservation wage, w^H and his estimate of the lowest possible reservation wage of the employee's. In the special case where w^H and w^{L2} are the same, there is only a 50% chance of the employer's offer being accepted. Half the surplus is wasted, and the remainder is shared equally between employer and employees.

If allowed to do so, the employer might wish to make a second offer in the event that his first offer is rejected. A second offer might be warranted if the first offer turns out to be below the employees' reservation wage, but the employees would have to be penalized sufficiently to forestall rejection of the first offer for no other reason than to wait for a better offer later on. The owners might have to delay the second offer long enough that the employees' gain from waiting exceeds the increase in the wage between the first and second offers.[7]

There is also a mismatch between the postulated structure of bargaining and the circumstances in which people actually bargain. The imposed sequence of offers and counter-offers acquires a certain plausibility from its resemblance to conversation, which, by its very nature, is a sequence of utterances and replies. The resemblance is superficial. Essential to the Staahl-Rubinstein bargaining process are a prescribed spacing between utterances and a prescribed order of speech, neither of which are intrinsic to actual conversation or to negotiation between firms, between employer and employees or between the buyer and seller of a house. *Ex post*, negotiation may have been an alternative sequence of offers. *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. There is nobody to enforce the prescribed sequence of offers, the required lapse of time between 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.

Implicit in the rigidly-imposed speech sequence in the Staahl-Rubinstein bargaining process is a prohibition of threats, dignified by the high-sounding mathematical name "sub-game perfect equilibrium". Broadly-speaking, sub-game perfect equilibrium means that I cannot promise today to act tomorrow in a way that will not be in my interest at that time. If all anybody can say is "$y", "yes" and "no", then I cannot threaten you with dire consequences if you do thus-and-such tomorrow. Threats are simply not in the vocabulary.

Consider the third of the bargaining examples above where I offered you $400 out of a $1000 pie and keep the remaining $600 for myself. As long as we adhere to the rigid sequence of speech and as long as we both behave rationally, no other outcome is possible. But if either of us can say or do what we please, there is nothing to stop you, at the very beginning of the process, from promising never under any circumstances to accept less than - say - $700. You might make the promise credible through a contract to pay a third party $1000 if you accept a penny less than $700. Do that, and you have beaten me. Since any disbursement of the pie requires our agreement, I now have no option other than to accept the remaining $300 or to allow the clock to run out, leaving myself with nothing at all. Nor is it any consolation to me that you get nothing either. In my own interest, I must agree to a $700-$300 split in your favour.

To be sure, you have no monopoly on threats. If you can make threats backed up by side contracts or by a need to preserve your reputation as a tough and astute bargainer, then so too can I. If we both threaten one another and if our threats are incompatible, adding up to more than the value of the pie to be shared, there can be no agreement and we both end up with nothing.

---

[7]For a thorough analysis of structure induced equilibrium with asymmetric information, see John Kennan and Robert W. Wilson, "Bargaining with Private Information", *Journal of Economic Literature*, March 1993, 45-104.

The risk of this outcome makes us 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.

4) Bargaining as the Confrontation of Possibly Irreconcilable Demands:

Once again two people, even and odd, are bargaining over the allocation of a pie of size P. Think of each party as making a "demand" for a share of the pie. Person O demands D_O. Person E demands D_E. If their demands are compatible - that is, if $D_O + D_E < P$ - then, the pie is allocated accordingly.[8] Otherwise, the pie disintegrates and both parties come away from the bargain empty-handed. Unlike the preceding models of bargaining, this framework is captures the property of real-life bargains that such bargains do sometimes end in failure where the potential surplus is lost altogether because no agreement can be reached about how the surplus is to be shared.

This representation of bargaining differs from the Staahl-Rubinstein representation in its timing. There, bargainers can meet over and over again, alternating offers for the sharing of a steadily diminishing pie. Here, bargainers meet once, making demands simultaneously, sharing the pie if and only if their demands are compatible, and losing the pie altogether if they are not.

Each person may be thought of as choosing his demand to maximize his expected return. Person O chooses D_O to maximize

$$D_O \, \text{prob}^O[D_E < (P - D_O)] \qquad (20)$$

where $\text{prob}^O[D_E < (P - D_O)]$ is the probability *as assessed by person O* that person E's demand will be low enough for the two demands to be compatible. Similarly, person E chooses D_O to maximize

$$D_E \, \text{prob}^E[D_O < (P - D_E)] \qquad (21)$$

where $\text{prob}^E[D_O < (P - D_E)]$ is the probability *as assessed by person E* that person O's demand will be low enough for the two demands to be compatible.[9]

[8] Compatible demands could give rise to any of several sharing rules: The rule might be that person O gets D_O, person E gets D_E , and the remainder of the pie, $P - (D_O + D_E)$, is wasted. Alternatively, the rule might allocate the remainder of the pie equally or in proportion to the bargainers' demands.

[9] This framework was employed by John Nash in "Two-Person Cooperative Games", *Econometrica*, 1953, 128-140, reprinted in Young, above. Nash solved the problem by imposing

In principle, the demands could be sequential or simultaneous. If the demands were sequential, the entire pie would accrue to the party entitled to present his demand first, because the other party would have no choice but to accept what is left. Only if the demands are simultaneous is the problem interesting; each party must presents his demand not knowing what the other will do. However, with demands presented simultaneously, the outcome of bargaining can only be predicted on the strength of presumed probability distributions of each party's expectations of the other's demands. Person O's demand depends on his expectation of what person E will do. Person E's demand depends on his expectation of what person O will do.

Suppose, for example, that each person's expectations about the other's demand can be represented by a rectangular probability distribution: As seen by person O, D_E has an equal chance of lying anywhere between 0 and P. As seen by person E, D_O has an equal chance of lying anywhere between 0 and P. In other words, for any and every D_O and D_E,

$$\text{prob}^O[D_E < (P - D_O)] = (P - D_O)/P \qquad \text{and} \qquad \text{prob}^E[D_O < (P - D_E)] = (P - D_E)/P \qquad (22)$$

With these probability distributions of expectations and for any demand D_O, the expected gain of person O - his demand weighted by his probability of getting it - becomes

$$D_O \, \text{prob}^O[D_E < (P - D_O)] = D_O(P - D_O)/P \qquad (23)$$

which is maximized by setting $D_O = P/2$. Similarly, if person E's probability distribution of person O's demand, D_O, were also rectangular, then person E would choose $D_E = P/2$ as well. The pair of probability distributions in equation (22) is especially fortunate because it leads to an exhaustive division of the pie, with nothing left over, and it allocates the pie equally.

Other expectations would have other consequences. If both parties were certain that the other was to demand 55% of the pie, they would both demand 45% for themselves, and 10% of the pie would be left over once both demands are fulfilled. Were that so, the person O's probability distribution of person E's demands would be

$$\text{prob}^O[D_E < (P - D_O)] = \{1 \text{ if } D_E < .45P\} \text{ or } \{0 \text{ if } D_E > .45P\} \qquad (24)$$

and person E's probability distribution would be defined accordingly. Alternatively, if both parties were certain that the other would demand 45% of the pie, they would both demand 55% for themselves, and neither would get anything because a combined demand for 110% of the pie counts as a failure to agree. In general, if person O's expectations about person E's demands could be expressed by a probability distribution $f^O(D_E)$, then

$$\text{prob}^O\left[D_E < (P - D_O\right] = \int_O^{P-D_O} f(D_E)\, dD_E \qquad (25)$$

---

axioms that do not, so far as I can tell, correspond to reasonable behaviour by parties bargaining with one another.

and $\text{prob}^E[D_O < (P - D_E)]$ would be defined accordingly.

This static formulation of the bargaining problem immediately raises the question of the how each bargainer comes to acquire his expectations of the other's behaviour. Presumably the probabilities, $\text{prob}^O[D_E < (P - D_O)]$ and $\text{prob}^E[D_O < (P - D_E)]$ are what they are by virtue of some evidence of past bargains between these two people or between other people with similar attributes. For example, if all previous bargains had resulted an equal sharing of the pie, then both parties to this bargain, person O and person E, might come to believe that the other will demand half the pie. Alternatively, if participants of all past bargains could be classified as O-types and E-types - for instance, as landlord an tenant, or boss and worker - and if the pattern in such bargains had been an apportionment of the one third of the pie to the O-type person and two-thirds of the pie to the E-type person, then both person O and person E in the current bargain might believe that the other will pose his demand accordingly. Person E may come to believe that $D_O = P/3$ and person O may come to believe that $D_E = 2P/3$. Any such "convention" would be self-sustaining. The emergence, sustainability and eventual alteration of conventions can be accounted for by the introduction of randomness in the size of the pie and the bargainers' perceptions of their opponents.[10]

The problem becomes fuzzy when the size of the pie is unobservable. The surplus in the sale of a house is the difference between the most the buyer is prepared to offer (the buyer's reservation price) and the least the seller is prepared to accept (the seller's reservation price). Typically, neither party can observe the other's reservation price, and, not infrequently, one or both parties is not too sure of its own reservation price. With such information concealed, there can be no history of demands from which each person's probability distribution of the of the other's demand may be inferred.

Nor is there place within this model of bargaining for the repeated offers and for the give and take characteristic of genuine real-world bargaining. Each negotiation is confined to two sentences,"I demand D_O." and "I demand D_E." with no second chance to for either party to propose an acceptable offer. What some see as the essential feature of bargaining - that people negotiate until their differences are resolved - is abstracted away. The contrast between the Staahl-Rubinstein bargaining model is instructive. Two essential aspects of bargaining are that negotiation may be protracted and that, in the end, no bargain may be struck. The Staahl-Rubinstein bargaining model incorporates the first aspect but not the second. The present bargaining model incorporates the second but not the first. Admittedly, the protracted negotiation in the Staahl-Rubinstein model is abstract and artificial, for the bargain is struck at the very first contact between bargainers in anticipation of a long series of contacts that would otherwise take

---

[10]See H.P. Young, "An Evolutionary Model of Bargaining", *Quarterly Journal of Economics*, 1993, 145-168. In this model, each bargainer's probability distribution of the other's response is a reflection of the observed history of the apportionment of the pie in past bargains. See also Tore Ellingsen, "The Evolution of Bargaining Behavior", *Quarterly Journal of Economics*, 1997, 581-601.

place. In this model, there is no such possibility. There can be no additional round of demands in the event that the first round ends in failure to agree.[11]

5) Threats and Blackmail

"Demands" could be interpreted as threats. Person O might as well be saying, "Give me D_O or I will destroy the pie altogether.", and person E might as well be saying, "Give me D_E or I will destroy the pie altogether."At least three kinds of threats can be distinguished: i) the automatic threat that the entire pie is wasted in the event that demands are not compatible and that neither party backs down, ii) the threat by one person to inflict harm on the other over and above the harm from failure to agree, and iii) the threat by one person to inflict harm on himself in the event that he backs down from his original demand and agrees to accept something less, a threat to induce concessions from the other person on the principle that, "One of us has to be reasonable, and it is not going to be me".

The *locus classicus* on threats and blackmail is Schelling's "An Essay on Bargaining" (in *The Strategy of Conflict,* 1960). 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. Presumably, 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. Much may depend on the bargainers' concerns for their reputations. Here, two considerations would seem to point in opposite directions. On the one hand, to induce prospective partners to join with you in other ventures, you want a reputation for being reasonable and accommodating. Nobody wants to become your partner if he expects you to be too demanding whenever conflicts of interest arise. Intransigence today may deter future partners from claiming too much. One the other hand, costly intransigence today may pay off tomorrow as a warning to your partners in future bargains that you are tough. Your partners might be

---

[11]There may be allowed a second round bargaining when no agreement is attained in the first. In "A Theory of Disagreement in Bargaining" (*Econometrica*, 1982, 607-637), V. P. Crawford postulates that both parties are prepared to back down in the second round as long as the cost of conceding is not too high. Each person's cost of backing down is assumed to be a random variable, the value of which is only discovered after the first round is complete. Following an impasse in the first round, person E reduces his demand from D_E to $P - D_O$ in the second round if and only if $P - D_O > C_E$ where C_E is person E's cost of backing down from his original demand. Person O acts accordingly. When one or both persons back down at the second stage, their new demands must be compatible and a bargain must be struck. Only if both persons' costs of concession are prohibitive can the impasse remain. Each person's strategy in choosing his demand in the first stage - D_E and D_O - comes to depend not just on expectations of the other's demand, but on the probability distribution of the costs - C_E and C_O - of backing down in the second stage.

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.

You cannot do both completely. There is an extensive literature on the practice of bargaining, covering such matters as when and how much to concede, and how to help your partner to save face in the event that he concedes to you.[12] All such considerations are assumed away in most bargaining theories by the assumption of "sequential rationality" or "subgame perfection". Seduced by 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.[13]

From Bargaining to Arbitration to Rent-seeking to Conflict to War

The outcome of bargaining may depend not on the cost of calamities to be averted by agreement, but on the cost of bargaining itself. In the theories discussed so far, the cost of failure to agree might take any of several forms: delay in apportioning the pie, wastage of all or part of the pie, a strike, or the breakdown of an otherwise harmonious association. The cost of bargaining itself might include the diversion of the bargainers' time from production to negotiation, employment of expensive professional negotiators, or harm to both parties from actions intended to induce the other to settle on terms favourable to oneself.

The starting point for a theory of bargaining based on the cost of negotiation is the original bargaining problem in figure 1 where person E and person O are jointly entitled to a pie of size P. Suppose for convenience there are no outside options, so that $x_E = x_O = 0$ and $P = y_E + y_O$. The core of this theory is the postulated connection between inputs and outputs, between expenditures on negotiation and the resultant shares of the pie. A "conflict success function" attributes shares of the pie - s_E and s_O where $s_E = y_E /P$ and $s_O = y_O /P$ - to the bargainers' expenditures on negotiation, F_E and F_O. The conflict success function is

$$s_E = s(F_E , F_O) \qquad (26)$$

where $s_O = 1 - s_E$ and where the function s must be increasing in F_E and decreasing in F_O. Person E and person O choose F_E and F_O to maximize their net incomes over and above their expenditures on negotiation. [14] Person E seeks to maximize z_E where

[12] See, for instance, Howard Raiffa, *The Art and Science of Negotiation*, 1982.

[13] For survey of contemporary bargaining theory, see Abhinay Muthoo, *Bargaining Theory with Applications*, Cambridge University Press, 1999.

[14] A conflict success function was implicit Winston Bush's model of anarchy, "Individual Welfare in Anarchy", in Gordon Tullock, ed., *Explorations in the Theory of Anarchy*, Center for

$$z_E = y_E - F_E = s_E P - F_E \tag{27}$$

Person O seeks to maximize z_O where

$$z_O = y_O - F_O = s_O P - F_O \tag{28}$$

Each person chooses his negotiating expenditure, F_E or F_E, to maximize his net income, z_E or z_O, in recognition of the negotiating expenditure of the other. It is customary, though not strictly necessary, to postulate an equilibrium where person E maximizes z_E and person O maximizes z_O as though the other person's choice where fixed independently of one's own behaviour.[15]

Expenditure on negotiation is wasted in the sense that both bargainers could be made better off if they could agree to whatever shares they bargain for without the expense of bargaining. They would acquire the entire pie if they could agree to keep F_E and F_O infinitesimally small and to accept whatever shares would emerge in the original bargaining equilibrium. But the very essence of bargaining as costly negotiation is that no such agreement is possible. The bargainers are caught in a standard prisoners' dilemma. They share the remainder of the pie after some portion has been eaten up by costly negotiation.

The Achilleas' heel of this model of bargaining is the interpretation of the conflict success function, the imposed resolution without which no equilibrium of negotiating costs can be identified. The conflict success function is an odd beast, neither technology nor taste, though a plausible interpretation will emerge in other contexts to be discussed below. Bargainers do negotiate, negotiation is expensive, the expense would not be incurred except to influence the outcome of the bargain, but, to the best of my knowledge, no function specifying the outcome of bargaining as dependent on expenditures for negotiation is derivable from the laws of nature or from rational self-interested behaviour. The conflict success function is every bit as *ad hoc* in

---

the Study of Public Choice, 1972. An explicit function was employed in Gordon Tullock, "Efficient Rent-Seeking" in Buchanan, Tollison and Tullock, eds., *Toward a theory of the Rent-Seeking Society*, Texas A & M University Press, 1980. The conflict success function was Christened and alternative functions were compared in Jack Hirshleifer, "Conflict and Rent-Seeking Success Functions: Ratio vs. Difference Models of Relative Success", *Public Choice*, 1989, 101-12.

[15] In the special case where the conflict success function is

$$s_E = F_E / (F_E + F_O)$$

the optimal values of F_E and F_O are are both equal to P/4, and the net incomes, z_E and z_O, of persons E and O turn out to be P/4 as well. Half the pie is wasted in expenditure to influence the arbitrator and the other half is shared equally as net incomes of the bargainers.

this context as the equal splitting of the difference in the models of voting and corporation finance at the beginning of this paper. The function is pulled out of thin air. The function is better grounded in other interpretations of the model to be discussed presently, but some of the fragility of the conflict success function as an explanation of the outcome of negotiation remains.

The conflict success function acquires some plausibility in the context of arbitration. Suppose, instead of negotiating directly, person E and person O agree to submit their dispute to an arbitrator, so that what we have been calling negotiating expenditure becomes expenditure to influence the arbitrator's decision. A God-like arbitrator would be imperious to the pleas of the parties to the dispute. Actual arbitrators are likely to be impressionable and open to solicitation or propaganda. The more impressionable the arbitrator, the greater the elasticities of s_E to F_E and of s_O to F_E , and the more of the pie will be wasted in expenditure to influence the arbitrator.

The contest success function may bear either of two interpretations in this context. The term s_E may be interpreted as the share of the pie accruing to person E as we have so far supposed, or it may be interpreted as the probability that the entire pie accrues to person E, in which case s_O would be the probability that the entire pie accrues to person O. On the latter interpretation, z_E and z_E become the *expected* net incomes of person E and person O, but the rest of the model remains the same.

It is a small step from arbitration to rent seeking. Instead of the bargainers choosing the arbitrator, it in now the arbitrator who chooses the bargainers. The paradigmatic rent seeking contest is over the allocation of shares of an import quota. As an alternative to a tariff, the government restricts the quantity of imports, raising the domestic price and generating "rents" for importers fortunate enough to be assigned shares of the quota. Rent-seeking in this context is expenditure by importers and would-be importers to acquire or to enlarge their shares. Expenditures F_E and F_O become rent-seeking expenditures, the proportions s_E and s_O become proportions or probabilities assigned by the rent-setter - in accordance with the conflict success function - to person E and person O, and the efficacy of rent-seeking is indicated by equations (26) to (28) above. Rent-seeking may be lobbying, advertising or political pressure, all of which are expensive. Rent-seeking merges into corruption. Rent-setters may be bribed or compensated in more insidious ways that are not exactly against the law, as when regulators may be offered jobs by the firms they are empowered to regulate.

A rent-seeking contest may be closed or open. A contest is closed when the government selects the contestants. The model in equations (26) to (28) is constructed on the assumption that the government selects only person E and person O, denying access to anybody else. In principle any number of rent-seekers may be selected. Alternatively, a rent-seeking contest may be open to everybody. With unrestricted entry into the contest, a modest fixed cost of entry and an unlimited supply of essentially identical entrants, the only possible equilibrium is where the entire pie is wasted.

Normally, rent-seeking expenditure is pure waste because the prize is what it is regardless of who finally acquires it or how it is divided among contestants. That is not invariably so. Rent-seeking becomes socially-advantageous when there is a spin-off or externality that is beneficial to the community at large. The principal example of socially-advantageous rent-seeking is the patent race. Monopoly is *per se* socially-disadvantageous, but it is awarded nevertheless as a patent on a newly-invented product to induce invention. From the point of view of the inventor, inventing is a form of rent-seeking, but it is no less desirable on that account. The analogy between inventing and rent-seeking is closest when pure research has identified the possibility of developing a patentable product, but nobody is quite sure how best to create it.

An additional assumption converts rent-seeking to "conflict".[16] Bargaining became rent-seeking when the prize was supplied by the government and the shares were made dependent of expenditure by the bargainers. Rent-seeking becomes conflict, as the term is commonly used in the literature of economics, when the prize becomes endogenous. A two-party model of conflict model contains i) an assignment of given resources (or income-earning capacity) R_E and R_O to person E and person O, ii) an apportionment by each party of his resources between production of goods, G, and fighting, F, and iii) an allocation of the combined production of the two parties together as a prize to be apportioned between the parties in accordance with the conflict success function in equation (26) above. Formally, equations (26), (27) and (28) remain unchanged and are supplemented with equations (29), (30) and (31) below

$$R_E = G_E + F_E \qquad (29)$$

$$R_O = G_O + F_O \qquad (30)$$

and $$P = G_E + G_O \qquad (31)$$

where F_E and F_O should now be reinterpreted as expenditures on fighting rather than on rent-seeking and where the conflict success function may be thought of as a representation of the technology of war.

As defined in economics, conflict occupies an intermediate position between commercial rivalry and outright war. As a pattern of commercial rivalry - for instance, between two would-be monopolists - the model of conflict is perhaps too all-encompassing, for the entire combined income of the contestants is at stake with no safe harbour, no secure source of income, remaining to the loser in the contest. On the other hand, as a representation of outright war, the model is too constricting for it fails to differentiate between arming and fighting. At a minimum, one would want a model of warfare to draw a distinction between costly preparation for war and the cost of war itself. Having armed, the potential combatants may choose to live with the status quo, or one of them may trigger a real war in which a significant part of their combined resources, over and above that devoted to military preparation, is destroyed in battle. The parties may choose to

[16]The transformation occurred in Jack Hirshleifer, "The Technology of Conflict as an Economic Activity", *American Economic Review*, 1991, 130-4.

live with the status quo, may precipitate a battle, or may bargain their way to a transfer of resources that both prefer to the contingencies of war. Bargaining to avert war is like bargaining in any other context; sometimes catastrophe is averted, sometimes not. The prospect of war leads back to bargaining which is every bit as mysterious, albeit more lethal, in that context as in any other.[17]

****

The emergence of equilibrium - of a determinate agreement predictable from the initial conditions and the preferences of the parties to the bargain - in the formal models of bargaining we have examined, gives rise to the presumption that there may be an equilibrium in practice too. The presumption is that shares of the prize over which people bargain are allocated by a mechanism comparable to the allocation of the shares of the national income in a competitive market. The presumption is that the old scepticism about bargaining, expressed in the quotation from Edgeworth at the outset of this paper can finally be abandoned. The examples in the first section of this paper reveal no hesitation about postulating a neat and costless division of the pie in politics, business and law.

Our brief examination of bargaining theory suggests that this is much too sanguine a view, and that the old scepticism was entirely justified. Though superficially similar to real bargaining, the bargaining models are really about something else altogether. The Hicks and Zeuthen models were ungrounded in utility. The Nash bargaining solution was *ex cathedra*, assuming the greater part of what was supposedly proved and substituting imposed arbitration for voluntary agreement. The Staahl-Rubinstein model imposed a rigid structure of bargaining with no counterpart in actual negotiation. None of the models could withstand the introduction of bargainers' threats. With no counterpart in actual negotiation, sub-game perfection may appeal more for the sound and connotations of the words than as a reasonable postulate about self-interested behaviour.

We need not deny that simple imposed bargaining solutions, such as those of Alesina and Rosenthal and of Hart, can be useful representations of empirical regularities in models designed explain something other than bargaining itself. Nor need we deny that bargains are often reached quickly and painlessly. What matters is that economists do not know how bargains are struck, why any particular agreement is reached, or why bargaining sometimes fails. The inexplicable does occur frequently, but we would be foolish to depend upon it. Millions of dollars may be divided easily. Valuable associations may suddenly dissolve over trifles. Threats, commitments and blackmail may be expunged from bargaining models but they refuse to depart from actual negotiation.

---

[17]On arming as preparation for bargaining, see Michelle Garfinkel, "Arming as a Strategic Investment in a Cooperative Equilibrium", *American Economic Review*, 1990, 50-68. On arming as preparation for war, see Michael Intrilligator and Dagobert Brito, "Conflict, War and Redistribution", *American Political Science Review*, 1985, 63-84.

The great virtue of the Alesina and Rosenthal model was its emphasis on the interplay of voting and bargaining as indispensable ingredients of democratic government. Democracy is not just a matter of voting. It is a complex mixture of voting and bargaining. Alesina and Rosenthal focused upon the content of bargains, but never questioned the premise that some bargain would be struck. It is precisely that premise which is called into question by the review of bargaining theory in this paper. A natural inference from the Alesina and Rosenthal model would be that, if bargaining breaks down, then democracy probably breaks down as well. We know from experience that people do usually bargain successfully, but it is disconcerting to realize that the process cannot be (or at least has not been) rationalized.

The policy implication of this critique is to reemphasize the virtues of a market with well-established property rights. General equilibrium with universal price-taking behaviour supplies the determinacy that bargaining lacks. Supported by the law of contracts, a competitive market may transact the world's business without much fuss, leaving only small, incidental corners where bargains must be struck and minimizing the harm from failure to agree. The world, of politics, especially international politics, is not so fortunate. Bargaining is less easily and less frequently circumvented. Perhaps a recognition of the indeterminacy of bargaining may forestall unwarranted optimism, rendering bargaining less dangerous in practice. For bargaining really is mysterious.

Appendix 1: Sequential Bargaining with Imposed Arbitration after T Periods

For the reader who is uncomfortable with the "interminability" assumption of the Staahl-Rubinstein bargaining procedure, the Staahl-Rubinstein bargaining solution can be derived as a limiting case of a bargaining procedure that terminates after a finite number of periods. The interminability assumption is that bargaining would continue forever if no agreement is reached. Suppose instead that bargaining would be terminated by compulsory arbitration at the end of T periods. Together with the rest of the Stahl-Rubinstein framework, this assumption yields a determinate outcome that boils down to the Staahl-Rubinstein solution in the special case where T approaches infinity. The proof is tedious but straightforward and intuitive.

Mr. E and Mr. 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 Mr. O in the year 1, the next by Mr. E in the year 2, and so on until an offer has been accepted or until the year T when the process is stopped by compulsory arbitration. Each offer specifies a division of the pie, so much for Mr. O and so much for Mr. E. Both parties are super-rational in a sense to become evident below.

Since bargaining begins in the year 1, Mr. O is entitled to make an first offer. On January 1 of the year 1, Mr. O makes an offer to Mr. E. If Mr E accepts the offer, then the division of the pie is arranged accordingly and the bargaining is over. If Mr E does not accept the offer, there is no further communication between the parties until January 1 of the year 2 when their roles are reversed. On January 1 of the year 2, Mr. E makes an offer to Mr O. Once again, if the offer is accepted, the division of the pie is arranged accordingly and the bargaining is over; and if the offer is not accepted, there is no communication between the parties until January 1 of the year 3.

If no bargain has been struck by the year T, compulsory arbitration supplies a share s to Mr. E leaving as share (1 - s) for Mr. O. In other words, an amount \$sP would be allotted to Mr. E and the remainder, \$(1-s)P, would be allotted to Mr. O. Suppose, for convenience that T is even. If neither party discounted the future - if the parties were indifferent between a dollar acquired today and a dollar acquired t years from today - then neither party would care when the bargain is struck, but, whenever the bargain is struck, it would have to be in accordance with the imposed allocation - \$sP for Mr. E and \$(1-s)P for Mr. O - because neither party would ever accept less in a year (T - t) than he can acquire by waiting until the year T. On the other hand, if the parties are impatient or if interest is forgone in waiting, both parties acquire an incentive to strike a bargain soon and the specifics of the bargain would depend on the parties' discount rates. Let δ_E and δ_O be the parties' discount factors, where δ_E / $1/(1+ r_E)$ and δ_O / $1/(1+ r_O)$, and where r_O and r_E are the parties' rates of interest.

The central concept in the derivation of the bargain is the minimal share of the pie that a party is prepared to accept in any year τ (where τ runs from 1 to T) in the event that, for whatever reason, no bargain had been struck already. Though, as we will show, the bargain will be struck in the year 1, the determination of that bargain will depend on what would have happened otherwise. Define $s(E, \tau)$ to be the lowest share that Mr. E would be willing to accept (if that share were offered by Mr. O) on January 1 of the year τ, and define s(O, τ) to be the lowest share that Mr..O would be willing to accept (if that share were offered by Mr. E) on January 1 of the year τ. Obviously, s(E, T) = s and s(O,T) = (1-s) because neither party would be prepared to accept less than he would be awarded by the arbitrator in the year T. Suppose for convenience that T is an odd number.

Since delay is costly to both parties, it is reasonable to suppose that the deal might be struck immediately. For this to happen, Mr. O's offer to Mr. E on January 1 of the year 1 must be accepted. Our problem is to deduce what that offer would have to be. Our problem is to deduce s(E, 1), the lowest share Mr. E is prepared to accept in an offer by Mr. O in the year 1 when the bargaining begins.

The problem can be solved by reverse induction. Suppose the parties have failed to reach an agreement until the year T - t, where t is an odd number, so that (since T is assumed to be even) T - t is an odd year and Mr O's is entitled to make an offer to Mr. E in that year. The offer must be s(E, T - t) which is, by definition, the least Mr. E would be prepared to accept in the year T - t in preference to waiting a year until he is entitled to make his own offer to Mr. O. Never mind for the moment how s(E, T - t) is determined or why it is what it is. Just assume that there is some value of s(E, T - t) which is known to both parties. It follows immediately that, if s(E, T - t) is offered to and accepted by Mr. E, then Mr. O's share of the pie in the year T - t would be 1 - s(E, T - t).

Now consider the preceding year, T - t - 1, when Mr. E is entitled to make an offer to Mr. O. Since Mr. O can acquire a share 1 - s(E, T - t) in the following year, he would not accept less than the present value of 1 - s(E, T - t) in the year T - t - 1, and need not be offered more. Specifically, Mr E would offer Mr O a share of $\delta_O[1 - s(E, T - t)]$ and Mr O would accept, leaving $1 - \delta_O[1 - s(E, T - t)]$ for Mr. E.

Finally, coming forward one more year, we see that the least Mr. O can offer Mr. E in the year T - t - 2 is the present value of $1 - \delta_O[1 - s(E, T - t)]$ one year ahead. Mr. E is offered $\delta_E\{1 - \delta_O[1 - s(E, T - t)]\}$, and he accepts. Thus, if Mr. E is prepared to accept a share of s(E, T - t) in the year T - t, he must be prepared to accept a share of

$$s(E, T - t - 2) = \delta_E\{1 - \delta_O[1 - s(E, T - t)]\} = (1 - \delta_O)\delta_E + \delta_O \delta_E\, s(E, T - t)$$

in the year T - t -2. By the very same reasoning

$$s(E, T - t - 4) = \delta_E\{1 - \delta_O[1 - s(E, T - t - 2)]\} = \{1 - \delta_O)\delta_E + \delta_O \delta_E\, s(E, T - t - 2)$$

$$= (1 - \delta_O)\delta_E + \delta_O \delta_E\, s(E, T - t - 2)$$

$$= (1 - \delta_O)\delta_E + \delta_O \delta_E \{(1 - \delta_O)\delta_E + \delta_O \delta_E\, s(E, T - t)\}$$

$$= (1 - \delta_O)\delta_E [1 + \delta_O \delta_E] + (\delta_O \delta_E)^2 s(E, T - t)$$

Repeating the procedure again and again, we see that

$$s(E, 1) = (1 - \delta_O)\delta_E\ [1 + \delta_O\ \delta_E + (\delta_O\ \delta_E\ )^2\ + ... + (\delta_O\ \delta_E\ )^{(T - t)/2}]\ + (\delta_O\ \delta_E)^{(T-t)/2} s(E, T - t)$$

Finally, recalling that s(E, T) = s which is Mr. E's assigned share in the imposed in the year T, we see that

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

which can be computed from values of the parameters δ_E, δ_O, T and s. For the special case where the compulsory arbitration takes places in the very distant future - that is, where T approaches infinity so that $(\delta_O\ \delta_E)^{T/2}s$ approaches 0 -

$$s(E, 1) = (1 - \delta_O)\delta_E\ /[1 - \delta_O\ \delta_E] = r_O\ /[r_O + r_E\ +\ r_O r_E]$$

where r_O and r_E are interest rates, so that

$$s(O, 1) = (r_E + r_O r_E)/[r_O + r_E\ + r_O r_E]$$

The pie may be thought of as carved into three shares, the "basic" shares of Mr. E and Mr. O - $r_O\ /[r_O + r_E\ +\ r_O r_E]$ and $r_E\ /[r_O + r_E\ +\ r_O r_E]$ - plus a premium to Mr. O - $(r_O r_E)/[r_O + r_E\ + r_O r_E]$ - as a consequence of his entitlement to make the first offer.

Equation (19) in the text becomes strictly valid when the time that must elapse between offers approaches 0 so that r_O and r_E become very small, the product $(r_O r_E)$ becomes infinitesimal by comparison with either r_O or r_E, and the premium for the first mover shrinks to 0 as well.

On these assumptions, the bargain is struck immediately with each party's share of the pie set in proportion to the other party's interest rate. Mr. E gets $\$r_O/[r_O + r_E]P$ and Mr. O gets $\$r_E/[r_O + r_E]P$.

Appendix 2: Reductio ad Absurdum when an annual cost of delay replaces the preference for present income over future income.

Suppose once again that E and O bargain over the allocation of a prize worth \$P, and that, if there were still no agreement after T periods have elapsed, compulsory arbitration would supply a share s to E and a share (1-s) to O. Continue to assume that T is even. Now, however, there is no discounting. Instead, E and O bear costs, C_E and C_O, per period for as long as the agreement is delayed. Suppose, for example, that no bargain is struck until period t and that the bargain supplies a share s(E, t) to E and the remaining share 1 - s(E, t) to O. The net gain to E - his slice of the pie *less* his cost of waiting - would be

$$y_E(t) = s(E, t)P - (t-1)C_E$$

and the net gain to O would

$$y_O(t) = [1- s(E, t)]P - (t-1)C_O.$$

Once again, it is in the common interest of E and O to strike a bargain at once. Our problem is to determine what s(E, 1) in that bargain must be. As in the preceding appendix, the procedure for discovering s(E, 1) is to reason backwards, from the year T toward the present.

If no agreement were reached by period T, the arbitrated shares of E and O would become s and (1-s) and their slices would be sP and (1-s)P. Now move forward one period from T to T-1. If no agreement has been reached by the year T-1 (where T-1 is odd because, by assumption, T is even), it would be O's turn to make an offer to E. Since E can assure himself of sP in the year T and must bear a cost C_E if he does not accept O's offer in the year T -1, E would be willing to accept any offer equal to or greater than sP - $C_{E\,,}$ and there is not reason why O should offer more. O would offer E a slice sP - C_E in the year T-1, raising O's own slice from (1-s)P to
(1-s)P + C_E and making O better off in two respects: because he gets (1-s)P a year earlier than if no agreement were reached until the year T and because he acquires C_E as well.

Move forward another period from T-1 to T-2 when it becomes E's turn to make the offer. Since O can acquire a slice (1-s)P + C_E by waiting until period T-1 and since the wait imposes a cost on O of C_O, the most E need offer O in period T-2 becomes (1-s)P + [C_E - C_O], leaving himself with slice P - [(1-s)P + C_E - C_O] which is equal to sP - [C_E - C_O].

The pattern should now be evident. Every two periods adds [C_E - C_O] to O's slice and subtracts [C_O - C_E] from E's slice. Since there are T/2 pairs of periods between 1 and T, the slices in period 1 must be

$$s(E, 1)P = sP - (T/2)[C_E - C_O]$$

and $$s(O, 1)P = (1-s)P + (T/2)[C_E - C_O]$$

A peculiar feature of this computation of equilibrium slices by backward induction is that the slices need not both be positive. For any difference, C_E - C_O, in the parties' costs of waiting, there must be some T large enough that one of the shares turns negative. The equilibrium bargain in the year 1 would then supply one of the two bargainers with the entire prize *plus* some premium reflecting the difference in the bargainers' costs of waiting, leaving the other party worse off than if there had been no prize to bargain over at all. Odder still, as T approaches infinity, one party must pay the other *an infinite amount of money* in the first year of the bargain regardless of the amount of money, P, at stake in the original bargain! If C_O exceeds C_E by as little as one penny, then O must pay E his entire fortune, or so the reasoning implies. This follows despite the fact that both parties' losses would be astronomical if they really did wait T periods before striking a deal.

The critical assumption in this demonstration is that neither party can escape from the agreement without the consent of the other. For example, if C_O exceeded C_E, then s(O, 1) might easily be negative, in which case O would happily walk away from the bargain in period 1 leaving the entire prize to E. If the bargain cannot be dissolved without mutual consent, O could easily claim the entire prize by promising not to release E from the bargain for a long time unless E conceded some very large amount. The essential difference between the discounting of future income in Appendix 1 and the fixed cost per period to each party in the event of a failure to agree lies in the bargainers' present values of their slices of the pie. With discounting, a long wait is at least conceivable because the present value of a slice acquired any number of years ahead remains positive, ensuring that nobody wants to walk away from the bargain leaving the entire prize to the other. With fixed costs of waiting per period, both bargainers might be prepared to walk away and the bargain deduced by backward induction seems, to say the least, implausible.

What Bargaining theory is Really About

Bargaining procedures can be looked upon from two points of view: They can be interpreted as what people *really do* when sitting across a table with a pie to be divided between them, or they can be interpreted as frameworks *imposed externally* upon bargainers by law or by prior agreement. Typically, the externally imposed frameworks yield precise solutions, for they would not otherwise be imposed. The question is whether such solutions track what people actually do in a pure bargain, as described in figure 1 above, without such imposition. Let it be agreed in advance to resolve disputes in a Nash bargaining solution, or within as Staahl-Rubinstein framework, or by means of a double auction, or in a variety of other ways, and a unique bargain emerges when both bargainers act rationally within the prescribed rules. The bargain may or may not be efficient depending on the circumstances, but it exists in the sense that some unique outcome emerges from the prescribed bargaining procedure.

A bargaining framework may be imposed by law or by prior agreement. Consider persons A and O who have established themselves as joint-owners of a newly-established firm. They get along well enough today, but they know there is some risk of conflict tomorrow. They realize that there may come a time when it is best for one of them to leave the firm. It would then become necessary for the leaver to be compensated. Compensation cannot be set at the time the firm is established because nobody knows if the partnership will ever be dissolved, when it will be dissolved or what the firm will be worth at the time. The amount of compensation has to be decided if and when the time come for one of the partners to go. Who goes, who stays and what compensation is paid is a standard bargaining problem. In these circumstances, a bargaining procedure may be written into the contract by which the firm is established. A possible procedure is this: If either party wishes to leave the firm, he can he can specify any amount compensation from the remaining partner, who must then either pay that amount or accept that amount himself as compensation for leaving the firm. Suppose person E wishes to leave the firm. He specifies an amount \$x. Then person O must either become sole owner of the firm on payment of \$x to person E, or person O can take the \$x himself, abandoning the firm to person E who, as sole owner, can either sell the firm or continue to run it. A similar procedure may be designed to accommodate the case where one partner wishes to buy out the other.

Bargaining works smoothly in this example because the bargaining procedure was set by prior agreement between the parties who might, equally-well have established some other procedure instead. Bargaining works because the bargainers had a common interest in setting up the bargaining procedure, despite the eventual conflict of interest when the bargain actually takes place. I think it fair to say that most of what goes by the name of bargaining theory is really about properties of bargaining procedures for future bargainers who have common interests today because they do not know their circumstances tomorrow when the bargain actually occurs.

It may, on the other hand, be argued in support of this or that bargaining theory that the theory represents how people actually behave when they are bargaining regardless of whether or not there is a prior agreement to accept the rules. Partners E and O may adopt the above procedure automatically even in the absence of any contractual requirement to do so. People's innate sense of fairness may lead them to agree on the Nash bargaining solution. Or they may agree to establish a double auction once the existence of a surplus is recognized.. On the other hand, it is by no means clear which among the competing bargaining theories would be appropriate. Referring again to figure 1 above, when parties A and E bargain over the surplus ($P - x_E - x_O$), do they split the difference as warranted by the Nash bargaining solution, or do they weight shares by the parties' impatience as warranted by then Staahl-Rubinstein bargaining solution, or do they make simultaneous demands as in a double auction. Worse still, the parties know that their choice of a bargaining procedure boils down in the end to a choice among outcomes, for the expected outcome of each bargaining procedure can be predicted in advance. They might agree on a procedure for themselves before they know the circumstances of the bargain that must be struck, but once they find themselves sitting across the table with a pie to be divided, procedure and outcome are very much the same.

Bargaining theories are like the clear and frank answers that politicians sometimes provide to questions just different enough from the questions actually asked to avoid embarrassment or to conceal critical information. The embarrassment is that we really do not know how bargains are actually struck.

Bargaining models can be classified as behavioural, ethical, and procedural. The behavioural models of Hicks and Cross postulate what bargainers do with no guarantee that the bargainers' behaviour is individually rational. Ethical models are about fairness. Procedural models impose a framework within which individually rational behaviour yields an outcome that may or not be efficient. Procedural models constrain offers. Some constrain offers sequentially. Others require offers to be simultaneous (the double auction)

Behavioural models supply wrong answers to the right question. Ethical and Procedural models supply right answers to wrong questions. Not quire, because bargaining really is sometimes constrained.