Arbitration and Information*

by
Avinash Dixit
Princeton University

First draft February 13, 2003
This version November 26, 2003

Abstract

Specialized institutions of arbitration can acquire and use an informational advantage over general courts of civil law. This paper models the idea to generate more precise implications, showing the connections between the nature of the additional information, the resulting expansion in the set of feasible contracts, and the improvement in the outcomes.

Address of author:
Avinash Dixit, Department of Economics, Princeton University, Princeton, NJ 08544–1021, USA.
Phone: 609-258-4013. Fax: 609-258-6419.
E-mail: `dixitak@princeton.edu`
Web: `http://www.princeton.edu/~dixitak/home`

1 Introduction

Oliver Williamson's discriminating alignment hypothesis offers a link between economic activities and the rules and institutions that govern them: "Transactions, which differ in their attributes, are aligned with governance structures, which differ in their cost and competence, so as to effect a discriminating – mainly a transaction cost-economizing – result"

*I thank Lisa Bernstein, Alessandra Casella, David Epstein, Xavier Gabaix, Karla Hoff, Oliver Williamson, and an editorial reviewer for their perceptive comments, the Russell Sage Foundation for their magnificent hospitality during a year spent there as a Visiting Scholar, and the National Science Foundation for research support. Substantial parts of this article reproduce material from Chapter 2 of Dixit (2004), and I thank Princeton University Press for permission to do so.

(Williamson, 1996, p. 12). How does this apply to arbitration, an institution that governs many commercial transactions? What are the cost or competence advantage of arbitration, and what kinds of transactions will be better governed by arbitration than by alternative methods? Williamson says that arbitration has "the capacity to evaluate disputes in a more knowledgeable way than the courts. ... Many agreements which, were it not for arbitration, would be regarded as excessively hazardous can, in this way, be reached and implemented." (ibid, p. 131-2). Bernstein (2001, p. 1741) makes this more precise: "[B]y providing for the appointment of industry-expert arbitrators, who can make many factual determinations more accurately and less expensively than a judge or jury can, the rules greatly expand the 'contractible' aspects of an exchange. The use of stream-lined procedures together with the appointment of expert adjudicators transforms considerations that in the public legal system would have been only observable to the parties . . . into considerations that are also verifiable . . . thereby encouraging transactors to enter into more complete contracts." In this paper I model this formally, thereby elucidating the precise kind of extra verifiability arbitration must possess if it is to achieve outcomes better than those available using formal contracts enforceable in state courts.

The history of arbitration, and the scholarly literature that describes and analyzes it, are both far too lengthy to summarize here. Landes and Posner (1979) provide a good overview and analysis of arbitration and other methods of private adjudication, and Bernstein (1992, 2001) studies arbitration in two specific industries. Bennett (2002) discusses the more practical aspects of arbitration. I merely extract a couple of brief points pertinent to my present purpose.

Arbitration tribunals run by trade associations are staffed by industry experts who are either active or retired members of the same trade. They have specialized expertise that enables them to understand technical complexities of the transaction, and have knowledge of accepted customs of the trade that may constitute implicit terms for a contract even when these are not explicitly written down. By contrast, the government's courts have to cover the entire range of civil law and do not possess expertise or detailed knowledge of any particular trade. Courts have access to expert witnesses, but many of these are hired by the parties to the dispute. Even when experts are brought in as unbiased consultants, the courts must interpret their findings in their own minds and constrained by their lack of specialized knowledge. The issue is not one of the amount of information available. The process of discovery, whereby one party can gain access to pertinent records in the other's files or e-mails, can bring before a court as much information as either party to the dispute finds it desirable to present. Rather, the issue is one of interpreting that information and using it correctly to adjudicate the dispute in that forum. Arbitrators do not generally permit such discovery.

More generally, Landes and Posner (1979, p. 249) argue that "the rules applied in commercial arbitration ... usually come from the courts and other sources of public law. ... Arbitration is generally limited to disputes where the rules are perfectly clear and the only issue is their application to the facts." Thus the advantage of arbitration resides in its superior ability to determine and interpret the facts. Arbitrators can use this advantage by adopting procedures that are more flexible than those of state law. Therefore the process can be speedier and less costly than state law, but that is not always the case.

Institutions specializing in arbitration but not confined to particular industries also exist, for example the American Arbitration Association, and the International Court of Arbitration. These do not have specific industry expertise, although they run some specific industry tribunals. Arbitration in such fora can be quite costly and the outcome somewhat unpredictable. However, they are often used in international trade because at least one party does not have enough knowledge of the other country's laws, or fears that the other country's courts may be corrupt or biased in favor of the home party. In this paper I am not concerned with this aspect. David (1985), Dezalay and Garth (1996), Casella (1996), and Mattli (2001) describe and discuss the institutions of arbitration in international trade.

Arbitration may render better and faster decisions, but how are these to be enforced? Private institutions lack the coercive power of the government's courts. Many systems of state law recognize and defer to arbitration, and stand ready to enforce the arbitrators' decisions. Even otherwise, institutions of arbitration can record and publicize any defiance of their verdicts, with the result that in a repeated interaction in a group, the cheater will suffer ostracism or boycott of future trade. In other words, enforcement of arbitration can be either formal or relation-based.

In this paper I assume that the government's courts enforce arbitrators' decisions if needed; so this is a model of "arbitration under the shadow of the law." I do not consider the possibility that the arbitrator may be dishonest and collude with one of the parties; this is assumed to be handled by the arbitrator's reputational concerns that are kept in the background.

The contract whose enforcement I consider is of the principal-agent type. This can be interpreted in the context of a firm's outsourcing decision. The principal buys from an agent a component that is an input to one of the principal's final products. The principal firm's ultimate objective is profit, but the contribution of the component to the profit generated by the final product, and indeed the contribution of the final product to the profit of the firm as a whole, are so thoroughly concealed in the accounts of the firm that no one else can disentangle them. Therefore a profit-based contract is infeasible. Matters like the quantities and time of delivery of the component are recorded and easily verifiable; therefore a contract that specifies the firm's payment to the component supplier as a function of these matters can be written and enforced by the government's civil courts. But a specialized arbitrator may be able to verify more subtle aspects of quality and fit of the component; then a contract that conditions payment on such aspects may become feasible under arbitration. The model can apply to some aspects of franchise contracts or managerial incentive contracts. However, I do not consider contracts between management and unions, where the actions of several workers or managers may be an issue.

Baker (2002) argues convincingly that in most agency situations the agent has a large number of available actions, and that the optimal design and power of incentives depends on the collinearity between two vectors of the marginal effects of actions: one on the objective of the principal (call this $\mathbf{y}$), and the other on the publicly verifiable performance measure, call it $\mathbf{x}$, that is the basis of the incentives. I extend this idea to the context where more verifiable measures are available to an arbitrator than to a court of law. In the main text, I develop this theme using simple linear and quadratic functions as in Baker (2002). I find that the extra information available to the arbitrator is useful only to the extent that the vector of marginal effects of actions on this new indicator is collinear with the direction in which it remains desirable to shift the agent's action beyond what is achievable in the governmental court of law; mathematically this is the component of $\mathbf{y}$ that is orthogonal to $\mathbf{x}$. This gives precision to the nature of informational advantage arbitration should have if it is to implement agreements that would otherwise be infeasible.

In an appendix I consider more general functional forms. This sheds a different light on the benefit of arbitration – the verifiability of more information can expand the set of instruments that are feasible in the incentive contract, and therefore can increase the dimensions of control of the agent's actions. Arbitration can achieve better outcomes to the extent that vectors of the marginal effects of the agent's actions on the newly verifiable events and on the measures that could be verified in a court of law are linearly independent. Somewhat paradoxically, if the arbitrator's additional indicator is stochastically independent of the one usable in the court, then the marginal effects of actions have greater linear dependence and therefore arbitration yields a smaller potential benefit.

2 A Simple Model

The interaction needing enforcement is between a principal and an agent. The agent takes an action $\mathbf{a}$, an ℓ-dimensional vector. The dimensions of this can be various aspects of the quantity and quality of effort, exercise of care, and so on. The agent's cost of taking the action is a quadratic function:

$$C(\mathbf{a}) = \tfrac{1}{2}\, \mathbf{a}'\mathbf{a}\,. \tag{1}$$

The outcome (revenue) for the principal is a binary variable y, which can equal 0 or 1 (e.g. failure or success of a project). The probability of success is a linear function:

$$\text{Prob}(\, y = 1 \,|\, \mathbf{a}\,) \equiv Y(\mathbf{a}) = \mathbf{y}'\mathbf{a}\,. \tag{2}$$

(It is assumed that the probability stays within the interval (0,1) for all relevant levels of actions.) Both parties are risk-neutral, but it is assumed that their joint involvement is essential – it is not feasible to sell the operation to the agent for a fixed fee, which would have given him exactly the right incentives. The agent's reservation utility is u_0 and the principal's reservation profit is 0.

3 First Best

The expected joint surplus from the relationship is

$$\begin{aligned} S &= Y(\mathbf{a}) - C(\mathbf{a}) \\ &= \mathbf{y}'\mathbf{a} - \tfrac{1}{2}\,\mathbf{a}'\mathbf{a}\,. \end{aligned} \tag{3}$$

The first-order condition defining the first-best action $\mathbf{a}^{FB}$ to maximize S is $\mathbf{y} - \mathbf{a} = 0$, so

$$\mathbf{a}^{FB} = \mathbf{y}\,, \tag{4}$$

and the resulting total surplus is

$$S^{FB} = \tfrac{1}{2}\,\mathbf{y}'\mathbf{y} \tag{5}$$

The first-best can be achieved if y is verifiable, by offering the agent a contract consisting of a salary s plus a bonus equal to 1 to be paid if $y = 1$. Then the agent's expected utility is

$$U = s + 1 \times \mathbf{y}'\mathbf{a} - \tfrac{1}{2}\,\mathbf{a}'\mathbf{a}\,.$$

Then the first-order condition for the agent's choice of $\mathbf{a}$ is the same as the joint surplus maximization condition (4). The salary s can be determined to meet the agent's participation constraint $U \geq u_0$.

The principal's outcome is often not verifiable. Therefore I now consider various alternative situations where incentive payments to the agent must be based on other verifiable indicators. We will find some other situations where the first-best is attainable; otherwise it serves as a hypothetical ideal standard.

4 Performance Measures and Their Observability

My focus is on a situation where only the principal can observe y. Therefore any contract conditioned on y is infeasible, even on a relational basis. There are two other performance measures x and w that can serve as a basis for contracts. Both are binary variables, taking values 0 or 1. I will assume that x is publicly verifiable, and w can be verified only by an arbitrator with specialized skill. So explicit or formal contracts based on x can be enforced in a governmental court of law. Contracts based on w as well as x can be adjudicated by the arbitrator, and I will assume in this paper that the government's courts stand ready to enforce the arbitrator's judgment if necessary.

This structure is special in two respects – there is only one performance measure of each kind (publicly verifiable and verifiable only by the specialized arbitrator), and each is binary. Both are easy to relax at the cost only of algebra, therefore I have opted for the simplicity afforded by my assumptions.

In the text of the paper I will follow Baker, Gibbons, and Murphy (2002) and assume that x and w are stochastically independent. In the appendix I will consider the more general case, and examine the implications of dependence for the possible set of contracts.

Each of x and w can take on two values; therefore four distinct realizations of the state of the world are possible. We need to specify probabilities for these. The ones most relevant are those on which payments to the agent can be made contingent. The public courts can verify x but not w. The arbitrator can verify both, and the most significant new item of information for him is if his indicator w shows success while the public x shows failure. In the text of the paper, for analytical simplicity, I will make both these probabilities linear functions of the agent's action $\mathbf{a}$; thus

$$\text{Prob}(\, x = 1 \,|\, \mathbf{a}\,) \;\equiv\; X(\mathbf{a}) = \mathbf{x}'\mathbf{a}\,, \tag{6}$$

$$\text{Prob}(\, x = 0,\; w = 1 \,|\, \mathbf{a}\,) \;\equiv\; W(\mathbf{a}) = \mathbf{w}'\mathbf{a}\,, \tag{7}$$

where $\mathbf{x}$ and $\mathbf{w}$ are constant vectors that denote the marginal effects of actions on the respective performance measures. In the appendix I consider more general functions. Also, to simplify later notation and without loss of further generality, I choose units so that the vector $\mathbf{x}$ has unit length, that is,

$$\mathbf{x}'\mathbf{x} = 1\,. \tag{8}$$

5 Court-Enforced Contract

The state legal system or governmental court can verify only x. Then the most general available form of the contract will stipulate payment to the agent in the form of an unconditional salary s, and a bonus ξ to be paid if $x = 1$. Then the agent's utility will be

$$U = s + \xi\, \mathbf{x}'\mathbf{a} - \tfrac{1}{2}\, \mathbf{a}'\mathbf{a}\,.$$

The agent's utility-maximizing choice of action $\mathbf{a}$ is characterized by the first-order condition

$$\xi\, \mathbf{x} = \mathbf{a}\,.$$

The principal chooses s and ξ to induce effort $\mathbf{a}$ so as to maximize his expected profit

$$\Pi = \mathbf{y}'\mathbf{a} - [\, s + \xi\, \mathbf{x}'\mathbf{a}\,]\,,$$

subject to the agent's participation constraint $U \geq u_0$. To meet this, the principal sets

$$s = u_0 - \xi\, \mathbf{x}'\mathbf{a} + \tfrac{1}{2}\, \mathbf{a}'\mathbf{a}\,.$$

Substituting for s, the principal's objective function becomes

$$\begin{aligned} \Pi &= \mathbf{y}'\mathbf{a} - \tfrac{1}{2}\, \mathbf{a}'\mathbf{a} - u_0 \\ &= (\mathbf{y}'\mathbf{x})\, \xi - \tfrac{1}{2}\, (\mathbf{x}'\mathbf{x})\, \xi^2 - u_0 \\ &= (\mathbf{y}'\mathbf{x})\, \xi - \tfrac{1}{2}\, \xi^2 - u_0\,, \end{aligned}$$

where the second line uses the agent's first-order condition and the second uses the normalization (8). The choice of ξ to maximize this, namely the bonus coefficient in the optimal contract with enforcement by governmental courts of law, is therefore

$$\xi^{CRT} = \mathbf{y}'\mathbf{x} \tag{9}$$

Then the agent's action is

$$\mathbf{a}^{CRT} = (\mathbf{y}'\mathbf{x})\, \mathbf{x}\,. \tag{10}$$

The resulting joint surplus is

$$S^{CRT} = \tfrac{1}{2}\, (\mathbf{y}'\mathbf{x})^2\,. \tag{11}$$

This is of course smaller than the first-best surplus in (5), formally this is because of the Cauchy-Schwartz Inequality.

The idea that the size of the bonus (the power of incentive) depends on the correlation between the vectors of the marginal products of efforts in increasing the principal's objective on the one hand, and the verifiable indicator on the other, is known from Baker (2002). An interpretation by analogy with information or regression theory may be more familiar. The principal induces the agent to take a vector of actions that is just the projection of the vector $\mathbf{y}$ (the coefficients of the principal's benefits of action) on $\mathbf{x}$ (the effects of actions on the verifiable performance measure). The shortfall of action below the first-best, namely

$$\mathbf{a}^{FB} - \mathbf{a}^{CRT} = \mathbf{y} - (\mathbf{y}'\mathbf{x})\ \mathbf{x} \tag{12}$$

is orthogonal to $\mathbf{x}$:

$$\begin{aligned}(\mathbf{a}^{FB} - \mathbf{a}^{CRT})'\mathbf{x} &= \mathbf{y}'\mathbf{x} - (\mathbf{y}'\mathbf{x})\ \mathbf{x}'\mathbf{x} \\ &= 0\,,\end{aligned}$$

using the normalization of the length of $\mathbf{x}$. Thus the information contained in the indicator is used to the full extent possible. This will be important when we study the properties of the optimal arbitration contract. Of course this simple interpretation of information in terms of projections and orthogonality is specific to the linear-quadratic structure; in the appendix I will consider more general functional forms.

6 Arbitration

Now introduce an arbitrator who can verify the outcome of w. Since x remains publicly verifiable, the two players can condition their contract on the realizations of both x and w. Now are four states of the world are distinguishable, therefore in addition to the salary s, there can be three distinct bonus payments. To allow direct comparison with the contracts of the previous section enforced in courts where only x was verifiable, I make one bonus contingent only on x. A second bonus is paid in the event of special interest in the context of arbitration, namely where the arbitrator's measure indicates success while the public one indicates failure. The bonuses are denoted by

$$\begin{aligned}&\xi \qquad \text{if } \ x = 1 \\ &\omega \qquad \text{if } \ x = 0 \text{ and } w = 1\,.\end{aligned}$$

Baker, Gibbons, and Murphy (2002) use a similar two-bonus contract for expository simplicity. In principle one can consider a separate bonus for a third contingency:

$$\beta \qquad \text{if } \ x = 1 \text{ and } w = 1\,.$$

I do this in the appendix. It turns out that the third bonus can be expressed as a linear combination of the first two if the two performance measures x and w are stochastically independent conditional on $\mathbf{a}$. I am assuming this in the text, so the two-bonus restriction on contracts involves no additional loss of generality here.

Recall that I am assume the official legal system is available to enforce the arbitrator's decision, backed by the power to levy sufficiently large penalties. Therefore such contracts are binding. We can find the optimal contract following the same steps as in the section on court-enforced contracts above.

The agent's expected utility is

$$U = s + \xi\ \mathbf{x}'\mathbf{a} + \omega\ \mathbf{w}'\mathbf{a} - \tfrac{1}{2}\,\mathbf{a}'\mathbf{a}\,.$$

The agent's choice of action is characterized by the first-order condition

$$\xi\ \mathbf{x} + \omega\ \mathbf{w} = \mathbf{a}\,.$$

The principal chooses s, ξ, and ω to induce effort $\mathbf{a}$ so as to maximize his expected profit

$$\Pi = \mathbf{y}'\mathbf{a} - [\,s + \xi\ \mathbf{x}'\mathbf{a} + \omega\ \mathbf{w}'\mathbf{a}\,]\,,$$

subject to the agent's participation constraint $U \geq u_0$. To meet this, the principal sets

$$s = u_0 - \xi\ \mathbf{x}'\mathbf{a} - \omega\ \mathbf{w}'\mathbf{a} + \tfrac{1}{2}\,\mathbf{a}'\mathbf{a}\,.$$

Substituting for s, the principal's objective function becomes

$$\Pi = \mathbf{y}'\mathbf{a} - \tfrac{1}{2}\,\mathbf{a}'\mathbf{a} - u_0\,.$$

Substituting for $\mathbf{a}$ from the agent's first-order condition and collecting terms into suitable vectors and matrices, we have

$$\Pi = (\,\mathbf{y}'\mathbf{x} \quad \mathbf{y}'\mathbf{w}\,) \begin{pmatrix} \xi \\ \omega \end{pmatrix} - \tfrac{1}{2}\,(\,\xi \quad \omega\,) \begin{pmatrix} \mathbf{x}'\mathbf{x} & \mathbf{x}'\mathbf{w} \\ \mathbf{w}'\mathbf{x} & \mathbf{w}'\mathbf{w} \end{pmatrix} \begin{pmatrix} \xi \\ \omega \end{pmatrix} - u_0\,.$$

The first-order condition to maximize this is

$$\begin{pmatrix} \mathbf{y}'\mathbf{x} \\ \mathbf{y}'\mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{x}'\mathbf{x} & \mathbf{x}'\mathbf{w} \\ \mathbf{w}'\mathbf{x} & \mathbf{w}'\mathbf{w} \end{pmatrix} \begin{pmatrix} \xi \\ \omega \end{pmatrix} .$$

This yields the following solutions for the bonus coefficients under arbitration:[1]

$$\begin{aligned} \xi^{ARB} &= \frac{(\mathbf{w}'\mathbf{w})\,(\mathbf{x}'\mathbf{y}) - (\mathbf{w}'\mathbf{x})\,(\mathbf{w}'\mathbf{y})}{(\mathbf{x}'\mathbf{x})\,(\mathbf{w}'\mathbf{w}) - (\mathbf{w}'\mathbf{x})^2} \\ \omega^{ARB} &= \frac{(\mathbf{x}'\mathbf{x})\,(\mathbf{w}'\mathbf{y}) - (\mathbf{w}'\mathbf{x})\,(\mathbf{x}'\mathbf{y})}{(\mathbf{x}'\mathbf{x})\,(\mathbf{w}'\mathbf{w}) - (\mathbf{w}'\mathbf{x})^2} \end{aligned} \tag{13}$$

This is again interpretable as regressing $\mathbf{y}$ on $\mathbf{x}$ and $\mathbf{w}$ jointly, and the resulting action $\mathbf{a}^{ARB}$ is just the projection of $\mathbf{y}$ on the plane spanned by $\mathbf{x}$ and $\mathbf{w}$.[2] The algebraic expression for the action is not especially insightful so I will omit it.

The solution is valid if the denominator is non-zero. This will be so unless the vectors $\mathbf{x}$ and $\mathbf{w}$ are perfectly collinear. If the marginal effects of actions $\mathbf{a}$ on the probabilities of the two indicators x and w are mutually proportional, then the w indicator will not carry any extra useful information. The two equations in (13) will collapse to one. Thus, if $\mathbf{w} = \lambda\,\mathbf{x}$ for a scalar λ, then the two equations will reduce to one:

$$(\mathbf{x}'\mathbf{x})\,(\xi^{ARB} + \lambda\,\omega^{ARB}) = (\mathbf{x}'\mathbf{y})\,,$$

which determines only the combination $\xi^{ARB} + \lambda\,\omega^{ARB}$. The agent's action also depends only on this combination,

$$\mathbf{a}^{ARB} = (\xi^{ARB} + \lambda\,\omega^{ARB})\;\mathbf{x}\,.$$

Therefore the outcome can be achieved using the ξ bonus alone, without arbitration, relying only on the official civil courts and their ability to verify x.

Now suppose the vectors of marginal effects of action on the two indicators are linearly independent. Then the denominator in (13) is positive by the Cauchy-Schwartz inequality. Substituting these bonuses into the expression for the agent' action and then into the principal's payoff, the expression for the total surplus from arbitration becomes

$$S^{ARB} = \tfrac{1}{2}\,\frac{(\mathbf{w}'\mathbf{w})\;(\mathbf{x}'\mathbf{y})^2 - 2\,(\mathbf{x}'\mathbf{w})\,(\mathbf{x}'\mathbf{y})\,(\mathbf{w}'\mathbf{y}) + (\mathbf{x}'\mathbf{x})\;(\mathbf{w}'\mathbf{y})^2}{(\mathbf{x}'\mathbf{x})\,(\mathbf{w}'\mathbf{w}) - (\mathbf{w}'\mathbf{x})^2}\,. \tag{14}$$

[1] I will use the normalization $\mathbf{x}'\mathbf{x} = 1$ later; for now I keep the more general form to bring out the symmetric structure of the solutions.

[2] This also tells us how the procedure can be generalized to allow multiple performance measures of each kind.

How does this compare to the surplus, found in (11) achievable using the state civil law alone? The difference provides an upper bound to the extra cost the parties would be willing to pay for having access to the arbitration forum. After some tedious algebra, we find

$$S^{ARB} - S^{CRT} = \frac{1}{2} \, \frac{[\,(\mathbf{x}'\mathbf{w})\ (\mathbf{x}'\mathbf{y}) - (\mathbf{x}'\mathbf{x})\ (\mathbf{w}'\mathbf{y})\,]^2}{(\mathbf{x}'\mathbf{x})\ [(\mathbf{x}'\mathbf{x})\,(\mathbf{w}'\mathbf{w}) - (\mathbf{w}'\mathbf{x})^2\,]}\,. \tag{15}$$

By the Cauchy-Schwarz inequality, the expression in the square bracket in the denominator of the right hand side is always non-negative. The other terms on the right hand side are inherently non-negative, the numerator being a square and the $\mathbf{x}'\mathbf{x}$ in the denominator being the length of the vector $\mathbf{x}$. Therefore the whole of the right hand side is non-negative. It can be zero only if the expression in the square brackets in the numerator is zero. Therefore the expected total surplus from arbitration is at least as great as that available in the court, and typically exceeds the latter. To understand this benefit of arbitration a little further, use (13) to note that

$$S^{ARB} - S^{CRT} \sim \left(\omega^{ARB}\right)^2 .$$

Thus arbitration yields a positive benefit any time the bonus payable upon realization of the indicator verifiable only to the arbitrator is actively used at a nonzero level in the contract governed by arbitration.

Note that ω does not have to be positive in this contract. Using the normalization (8), we see that the sign of ω is the same as the sign of

$$\mathbf{w}'\,[\,\mathbf{y} - (\mathbf{y}'\mathbf{x})\ \mathbf{x}\,] = \mathbf{w}'\,[\,\mathbf{a}^{FB} - \mathbf{a}^{CRT}\,]\,,$$

using (12). Thus the extra bonus instrument available to the principal under arbitration is more useful when the vector of effects of actions on the newly verifiable measure w points is closer to being collinear with or parallel to the vector by which action would fall short of the first best under court enforcement alone. If the two vectors are positively aligned in the sense of being at an acute angle to each other, a positive bonus in the event $w = 1$ serves to move action in the right direction; if the two vectors are negatively aligned, that is, point at an obtuse angle to each other, then the bonus should be negative. Only if $\mathbf{w}$ is orthogonal to the action shortfall is arbitration no better than court enforcement, and the bonus if $w = 1$ is then zero.

The intuition can be completed by asking how the availability of the new performance measure w changes the bonus awarded on the basis of the publicly verifiable measure x. Using (9) and (13), we find

$$\xi^{ARB} - \xi^{CRT} = -(\mathbf{w}'\mathbf{x})\, \omega^{ARB}\,.$$

If the effects of action on the two variables are positively aligned, then the two measures are substitutes – for example, if the new bonus is used with a positive magnitude, its availability reduces the magnitude of the old bonus. This is again similar to the theory of omitted variables in econometrics.

When the gain in surplus from arbitration is positive, the size of this gain is an upper bound on the cost of arbitration that would still leave a positive net gain. Of course it remains to structure a game with fully specified moves and division of the net gains that will lead the parties to choose this route, but that is quite easy to do so I will omit it.

7 Concluding Comments

The main contribution of the model is to clarify and make precise the nature of the informational advantage that can make adjudication by arbitration superior to that in a court of law. The model also yields some by-products. For example, it helps us understand the amicable coexistence between state law and private institutions of arbitration. Arbitration does not detract from the state's ability to enforce contractual provisions contingent on publicly verifiable information, and the existence and primacy of the state's law need not interfere with the extra verification and more complete contracting that arbitration makes possible. The courts' forbearance for the decisions of arbitrators, and their willingness to lend the services of their monopoly of coercion to enforce the awards made by arbitrators if that becomes necessary, become immediately understandable in this light. Such synergy does not obtain for all modes of private governance. For example, relational contracts and formal law can interact dysfunctionally. Relational contracts can use better information that is observable by the parties to a contract but is not publicly verifiable. The expected outcome from a formal court-enforced contract acts as the incentive compatibility constraint on the relational contract in a repeated game. A partial improvement in the formal contract tightens this constraint, and therefore worsens the best feasible relational contract; see Baker, Gibbons, and Murphy (1994). Arbitration seems fortunate to escape such harmful interactions.

The result that arbitration allows more complete contracts to be written seems to point to a puzzle. If arbitration expands the set of contingencies on which contractual terms can be conditioned, then, other things equal, contracts governed by arbitration should be more complex than those governed by courts. However, Williamson (1996, p. 96) offers an example of a contract that stipulates arbitration if the parties cannot resolve a dispute on their own. At least to the extent we are told, the contract is very simple and incomplete. No detailed contingent actions are specified that have to be verified and enforced under arbitration. Instead, there are very general references to "inequitable conditions" that may occur and the action required "in good faith ... to cure or adjust for the inequity." The apparent paradox can be resolved by distinguishing between completeness and complexity. Arbitration contracts can be more complete and yet simpler if the various contingencies can be left implicit because they are well understood in the custom of the industry by all parties, including industry-specialist arbitrators.

The general structure of the model can be adapted to examine several variants. Here are just a few of these, suggesting ideas for future research.

[1] I assumed that the government's courts defer to the arbitrator's expertise and stand ready to enforce his verdict. If that is not the case, but w is observable to both parties, then the arbitrator's decision could be enforceable in a relational context of a repeated game, along the lines of the analyses in Baker, Gibbons, and Murphy (1994, 2002). This can be better than a purely relational contract without any third party arbitration, because the arbitrator can impose fines, whereas the two parties can only inflict smaller punishments, namely refusal to trade in the future. Thus the arbitrator can help the parties implement better carrot-and-stick punishments as in Abreu (1986). Other repeated-game models of arbitration include Milgrom, North and Weingast (1990) and Ramey and Watson (2002).

[2] I calculated the gains from arbitration assuming that the only alternative was court-enforced contracts conditioned on the publicly observable performance measure x. But if the two parties can observe w, then the alternative may be a combination of such a formal contract and a relational contract based on w.

[3] I did not model the supply side of arbitration. This would introduce such matters as a fixed cost of acquiring expertise, a possible tradeoff between this cost and the scope of transactions where the expertise yields additional verifiable information, competition between rival arbitrators, and so on.

[4] The model can be interpreted in the context of a firm's implicit contracts with its employees or managers, not merely that of exchange between independent partners. This ties in with another of Williamson's major interests, namely vertical integration. Arbitration can expand the set of feasible contracts, thus reducing transaction costs and improving the outcome, both for internal employment contracts and external supply contracts. One cannot say a priori which mode will benefit more, and therefore it is not clear whether arbitration will make vertical integration more or less desirable. The answer is likely to be context-specific; this should provide an interesting question for empirical researchers.

Appendix – A More General Model

The agent's action $\mathbf{a}$ is an ℓ-dimensional vector as before. But now the agent's cost of action $C(\mathbf{a})$ is a general function, as are the probabilities for the principal's outcome

$$\text{Prob}(\, y = 1 \,|\, \mathbf{a}\,) = Y(\mathbf{a})\,,$$

and the performance measures

$$\begin{aligned}
\text{Prob}(\, x = 1,\ w = 1 \,|\, \mathbf{a}\,) &= B(\mathbf{a})\,,\\
\text{Prob}(\, x = 1,\ w = 0 \,|\, \mathbf{a}\,) &= X(\mathbf{a}) - B(\mathbf{a})\,,\\
\text{Prob}(\, x = 0,\ w = 1 \,|\, \mathbf{a}\,) &= W(\mathbf{a})\,,\\
\text{Prob}(\, x = 0,\ w = 0 \,|\, \mathbf{a}\,) &= 1 - X(\mathbf{a}) - W(\mathbf{a})\,.
\end{aligned}$$

These functions are assumed to be twice-differentiable and subject to all the second-order conditions of maximization in the calculations below.

A special case of some interest is when, conditional on $\mathbf{a}$, the two indicators x and w are stochastically independent. The condition for this stipulates that, given $\mathbf{a}$,

$$\text{Prob}(\, w = 1 \,|\, x = 1\,) = \text{Prob}(\, w = 1 \,|\, x = 0\,)\,,$$

or

$$\frac{B(\mathbf{a})}{X(\mathbf{a})} = \frac{W(\mathbf{a})}{1 - X(\mathbf{a})}\,.$$

This simplifies to

$$B(\mathbf{a})\ [1 - X(\mathbf{a})] = X(\mathbf{a})\ W(\mathbf{a})\,. \tag{A.1}$$

(Equating the two probabilities for x conditioned on different realizations of w leads to the same condition.)

First Best

The joint surplus from the relationship is

$$S = Y(\mathbf{a}) - C(\mathbf{a})\,.$$

The first-order condition defining the first-best action $\mathbf{a}^{FB}$ to maximize S is

$$\nabla Y(\mathbf{a}) = \nabla C(\mathbf{a})\,, \tag{A.2}$$

where ∇ denotes the gradient of a function – the (column) vector of its partial derivatives. I will assume that the resulting first-best joint surplus S^{FB} exceeds the sum of the two parties' reservation payoffs, namely u_0; otherwise the whole relationship would be irrelevant.

As in the linear-quadratic case of the text, the first-best can be achieved if y is verifiable, by offering the agent a contract consisting of a salary s plus a bonus equal to 1 to be paid if $y = 1$.

Court-Enforced Contract

The state legal system or official court can verify only x. Then, as in the linear-quadratic case of the text, the most general available form of the contract will stipulate payment to the agent in the form of an unconditional salary s, and a bonus ξ to be paid if $x = 1$. Then the agent's utility will be

$$U = s + \xi\ X(\mathbf{a}) - C(\mathbf{a})\,.$$

The agent's utility-maximizing choice of $\mathbf{a}(\xi)$ is characterized by the first-order condition

$$\xi\ \nabla X(\mathbf{a}) = \nabla C(\mathbf{a})\,. \tag{A.3}$$

To find the comparative static derivative of $\mathbf{a}(\xi)$ to ξ, differentiate (A.3) totally:

$$[\nabla^2 C - \xi \nabla^2 X]\, d\mathbf{a} = \nabla X \, d\xi \,,$$

or

$$\frac{d\mathbf{a}}{d\xi} = [\nabla^2 C - \xi \nabla^2 X\,)]^{-1} \nabla X \,, \tag{A.4}$$

where the operator ∇^2 applied to a function denotes the ℓ-by-ℓ matrix of the second-order partial derivatives of that function, and the points of evaluation $\mathbf{a}(\xi)$ are omitted for brevity. The matrix within the square brackets on the right hand side of (A.4) is positive definite by the second-order condition of the agent's utility maximization problem.

The principal chooses s and ξ to induce effort $\mathbf{a}$ so as to maximize his profit

$$\Pi = Y(\mathbf{a}(\xi)) - [\, s + \xi \; X(\mathbf{a}(\xi))\,] \,,$$

subject to the agent's participation constraint $U \geq u_0$. To meet this, the principal sets

$$s = u_0 - \xi \; X(\mathbf{a}(\xi)) + C(\mathbf{a}(\xi)) \,.$$

Substituting for s, the principal's objective function becomes

$$\Pi = Y(\mathbf{a}(\xi)) - C(\mathbf{a}(\xi)) - u_0 \,. \tag{A.5}$$

Observe that when s is chosen as a function of the principal's other instruments to satisfy the agent's participation constraint, the principal's objective becomes fully consistent with social optimality. The only question is how well the available instruments can affect the agent's action. This will be the theme of much of the later analysis.

The first-order condition for the choice of ξ to maximize Π is

$$\begin{aligned} 0 &= [\nabla Y - \nabla C\,]' \, d\mathbf{a}/d\xi \\ &= [\nabla Y - \nabla C\,]' \, [\nabla^2 C - \xi \nabla^2 X\,]^{-1} \nabla X \,. \end{aligned} \tag{A.6}$$

The right hand side in (A.6) is an ℓ-by-1 (column) vector. If $\ell = 1$ (effort $\mathbf{a}$ is a scalar). The right hand side of (A.6) becomes the product of three scalars, of which the last two are nonzero. Therefore the condition reduces to $\nabla Y = \nabla C$, which is the first-best. The intuition is that the single bonus variable ξ suffices to control the scalar effort $\mathbf{a}$. At least it does so locally. There remains the question of global controllability, namely of whether the range of the function $\mathbf{a}(\xi)$ includes the first-best a^{FB}. Very little useful can be said about this with general functional forms; therefore I will not pursue this.

Arbitration

Now introduce an arbitrator who can verify the outcome of w. Since x remains publicly verifiable, the two players can condition their contract on the realizations of both x and w. This opens up the possibility of three distinct bonus payments in addition to the salary s:

$$\begin{array}{ll} \xi & \text{if } \; x = 1 \\ \omega & \text{if } \; x = 0 \text{ and } w = 1 \\ \beta & \text{if } \; x = 1 \text{ and } w = 1 \,. \end{array}$$

Let $\mathbf{v}$ denote the 3-by-1 column vector of the bonus payments ξ, ω, and β in that order.

Recall that I am assuming the arbitrator's decision to be enforceable, backed by the coercive power of the courts if necessary. Therefore there is no need for a self-enforcement requirement or a corresponding additional incentive-compatibility constraint.

With this, the agent's expected utility is

$$U = s + \xi \; X(\mathbf{a}) + \omega \; W(\mathbf{a}) + \beta \; B(\mathbf{a}) - C(\mathbf{a}) \,.$$

The agent's choice of action is characterized by the first-order condition

$$\xi \; \nabla X(\mathbf{a}) + \omega \; \nabla W(\mathbf{a}) + \beta \; \nabla B(\mathbf{a}) = \nabla C(\mathbf{a}) \,. \tag{A.7}$$

This can be written in the form

$$\Big(\nabla X(\mathbf{a}) \,\vdots\, \nabla W(\mathbf{a}) \,\vdots\, \nabla B(\mathbf{a}) \Big) \;\; \mathbf{v} = \nabla C(\mathbf{a}) \,,$$

where the partitioned matrix on the left hand side is ℓ-by-3 (consisting of three horizontally stacked ℓ-by-1 column vectors), and the vector multiplying it is 3-by-1, yielding an ℓ-by-1 matrix to match the size of the vector on the right hand side. The condition defines the agent's action as a function $\mathbf{a}(\mathbf{v})$ of the vector of bonuses $\mathbf{v}$. Its comparative static derivative can be found by total differentiation,

$$\frac{d\mathbf{a}}{d\mathbf{v}} = \Big[\; \nabla^2 C - \xi \, \nabla^2 X - \omega \, \nabla^2 W - \beta \, \nabla^2 B \; \Big]^{-1} \; \Big(\nabla X \,\vdots\, \nabla W \,\vdots\, \nabla B \Big) \,. \tag{A.8}$$

The first matrix on the left hand side is ℓ-by-ℓ, and is positive definite by the second-order condition of the agent's utility maximization problem. The left hand side, being the derivative of an ℓ dimensional function of a three-dimensional variable, is an ℓ-by-3 matrix. All the gradients etc. are evaluated at $\mathbf{a}(\mathbf{v})$.

The principal must now choose s and $\mathbf{v}$ to maximize Π. Subsuming the agent's optimal choice of $\mathbf{a}$, the participation constraint yields

$$s = u_0 - \xi \; X(\mathbf{a}(\mathbf{v})) - \omega \; W(\mathbf{a}(\mathbf{v})) - \beta \; B(\mathbf{a}(\mathbf{v})) + C(\mathbf{a}(\mathbf{v}))\,.$$

Then

$$\begin{aligned}\Pi &= Y((\mathbf{v})) - [\; s + \xi \; X(\mathbf{a}(\mathbf{v})) + \omega \; W(\mathbf{a}(\mathbf{v})) + \beta \; B(\mathbf{a}(\mathbf{v})) \;] \\ &= Y((\mathbf{v})) - C(\mathbf{a}(\mathbf{v})) - u_0\end{aligned}$$

The first-order condition for the choice of $\mathbf{v}$) to maximize this is

$$\begin{aligned}0 &= [\; \nabla Y - \nabla C \;]' \; d\mathbf{a}/d\mathbf{v} \\ &= [\; \nabla Y - \nabla C \;]' \; \Big[\; \nabla^2 C - \xi \, \nabla^2 X - \omega \, \nabla^2 W - \beta \, \nabla^2 B \;\Big]^{-1} \; \Big(\nabla X \,\vdots\, \nabla W \,\vdots\, \nabla B \Big) \quad \text{(A.9)}\end{aligned}$$

The third factor on the right hand side is an ℓ-by-3 matrix. Suppose $\ell = 3$ (the agent takes three actions). Suppose also that the partitioned matrix $\Big(\nabla X \,\vdots\, \nabla W \,\vdots\, \nabla B \Big)$, which is now 3-by-3, is nonsingular. The inverse matrix in the middle of the right hand side is positive definite. Therefore the condition (A.9) reduces to $\nabla Y = \nabla C$, which is the condition for the first-best. Subject to the problem of global controllability about which little useful can be said for general functional forms, we see that in this case the first-best is attainable. Thus arbitration can expand the range of principal-agent interactions for which the first-best is feasible.

The partitioned matrix $\Big(\nabla X \,\vdots\, \nabla W \,\vdots\, \nabla B \Big)$ is nonsingular if the vectors ∇X, ∇W, and ∇B are linearly independent, and in the present context these derivatives should be evaluated at the first best action a^{FB}. How does this linear independence relate to the stochastic independence of the indicators x and w? The condition for the latter is (A.1). Taking logarithms and differentiating, we have

$$\frac{1}{B} \; \nabla B - \frac{1}{1 - X} \; \nabla X - \frac{1}{X} \; \nabla X - \frac{1}{W} \; \nabla W = 0\,. \qquad \text{(A.10)}$$

For any given **a** where all these functions and gradients are evaluated, this gives us a linear dependence relation linking the vectors ∇X, ∇W, and ∇B. Thus, if x and w are *stochastically independent,* then the vectors of the marginal effects of the action on the three probabilities are *linearly dependent.* The partitioned matrix has rank (at most) 2. One of the three bonuses is redundant (it can be constructed out of a suitable combination of the other two), and the two independent bonuses that are available cannot control the agent's three-dimensional action fully.

It may seem paradoxical that the additional indicator w that can be verified by the arbitrator is more valuable when it is not stochastically independent of the publicly verifiable x. The apparent paradox is resolved by recognizing that it is precisely the absence of independence that provides extra information from the event where both x and w equal 1, and therefore an extra dimension or degree of freedom to structure the bonuses.

References

Abreu, Dilip. 1986. "Extremal equilibria of oligopolistic supergames." *Journal of Economic Theory,* 39 (1), June, 191–225.

Baker, George. 2002. "Distortion and risk in optimal incentive contracts." *Journal of Human Resources,* 37 (4), Fall, 728–751.

Baker, George, Robert Gibbons and Kevin J. Murphy. 1994. "Subjective performance measures in optimal incentive schemes." *Quarterly Journal of Economics,* 109 (4), November, 1125–1156.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 2002. "Relational contracts and the theory of the firm." *Quarterly Journal of Economics,* 117 (1), February, 39–84.

Bennett, Steven C. 2002. *Arbitration: Essential Concepts.* New York: American Lawyer Media.

Bernstein, Lisa. 1992. Opting out of the legal system: Extralegal contractual relations in the diamond industry. *Journal of Legal Studies* 21: 115-157.

Bernstein, Lisa. 2001. "Private commercial law in the cotton industry: Creating cooperation through rules, norms, and institutions." *Michigan Law Review,* 99 (7), June 2001, 1724–1788.

Casella, Alessandra. 1996. "On market integration and the development of institutions: The case of international commercial arbitration." *European Economic Review,* 40 (1), January, 155–186.

David, René. 1985. *Arbitration in International Trade.* Deventer, Netherlands and Boston, MA: Kluwer Law and Taxation Publishers.

Dezalay, Yves and Bryant G. Garth. 1996. *Dealing in Virtue: International Commercial Arbitration and the Construction of a Transnational Legal Order.* Chicago and London: University of Chicago Press.

Mattli, Walter. 2001. "Private justice in a global economy: From litigation to arbitration." *International Organizations,* 55 (4), Autumn, 919–947.

Milgrom, Paul R., Douglass C. North and Barry R. Weingast (1990) "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs," *Economics and Politics,* 2 (1), March, 1–23.

Ramey, Garey and Joel Watson (2002) "Contractual Intermediaries," *Journal of Law, Economics, and Organization,* 18 (2), October, 362–384.

Williamson, Oliver E. (1996) *The Mechanisms of Governance,* Oxford, UK and New York: Oxford University Press.