Oxford Economic Papers, 71(1), 2019, 292–308
doi: 10.1093/oep/gpy035
Advance Access Publication Date: 30 August 2018

OXFORD

Bargaining and hold-up: the role of arbitration

By Yannick Gabuthy[a] and Abhinay Muthoo[b]

[a]University of Lorraine, University of Strasbourg, CNRS, BETA, Nancy, France; e-mail: yannick.gabuthy@univ-lorraine.fr
[b]Department of Economics, University of Warwick, Coventry

Abstract

This paper analyses arbitration as a surrogate for complete contracts. We embed this idea in a simple model of a long-term relationship between a firm and its workforce, in which they can make productive-enhancing, relationship-specific investments, and then negotiate over the division of the resultant surplus. It is shown that the mere presence of the arbitrator (in the background of negotiations) may enhance investment incentives *ex ante* by minimizing each party's ability to engage in hold-up behaviours *ex post*. Furthermore, we highlight notably that the partners should optimally commit to call an arbitrator ensuring a compromise by awarding a reasonable share of the surplus to the worker. Indeed, this type of arbitrator would harmonize the parties' bargaining powers and then weight their investment incentives optimally.

JEL classifications: D74, J52, K41.

> . . . the role of the arbitrator is, in effect, to remember who sunk costs in the past, and to ensure that future compensation is paid in a way that creates the correct incentives to make such investments. It is important that he can do this without infringing on the right to bargain, simply by threatening . . . an arbitral settlement that leads bargainers to negotiate, for themselves, the appropriate agreement. (Crawford, 1985, p. 375)

1. Introduction

1.1 Motivation

Arbitration can be defined by the presence of a third party who is empowered to impose a binding settlement when parties fail to agree on one via bilateral negotiations. This mechanism is now widely used in developed countries for most types of disputes, and its popularity may be explained by the increasing caseload of public courts, a preference for confidentiality, and the desire to reach agreements with minimal delay and minimal cost. For example, in labour agreement, the costs of unresolved disputes may be dramatic, as when there is a strike, or such costs may evolve more slowly as there is a steady erosion of morale and productivity in the workplace. It is precisely for these reasons that arbitration is currently used to settle wage disputes in the US public sector and is included in many contracts as an alternative to litigation (Deck *et al*., 2007).[1]

In this paper, we argue that arbitration may also serve quite a different purpose, namely, as a surrogate for complete contracts. It can act as a mechanism that indirectly increases the incentives of parties in a long-term relationship to undertake productivity-enhancing (but non contractible) relationship-specific investments. Given the presence of transaction costs, it is optimal—in terms of promoting such investments—to determine whether or not parties in a long-term relationship would engage the services of an arbitrator (if they fail to resolve any disputes by themselves) at the outset of their relationship, well before undertaking such investments. The basic argument that underlies this idea, which we develop in a simple model, runs as follows. Whether or not disputes end-up being resolved by arbitration determines the parties' respective outside options. This, in turn, affects their respective bargaining powers over conflict situations encountered during their relationship, and hence influences the distribution of the surplus. This then implies that it would indirectly influence each party's incentives to contribute to the generation of such surplus *ex ante*. In sum, the presence of an arbitrator in the background of negotiations may have efficiency consequences (as well as having distributive consequences), by mitigating the well-known hold-up problem.[2] Indeed, arbitration may be viewed as a mechanism able to protect each party's investment against opportunism by his partner (who could otherwise use the threat of disagreement to appropriate some of the returns from the other party's investment during the *ex post* bargaining stage).

1.2 Arbitration literature

This potential role of arbitration has been overlooked in the literature, which essentially studies the allocative effects of arbitration (i.e. the impact that arbitration may or may not have on the negotiated outcome). The authors analyse mainly the strategic interactions that influence how bargainers respond to arbitration, in order to study the effects of arbitration on negotiations and/or the parties' behaviour during the arbitration stage itself (Armstrong, 2004; Deck and Farmer, 2009). Arbitration is then viewed as a process whose *raison d'etre* is to enhance the likelihood of reaching an agreement. Indeed, it is commonplace in this literature that negotiated agreements are to be preferred to the arbitral settlements, and the extent to which an arbitration procedure encourages bargaining—by promoting convergence between the disputants' claims—is thus the main criterion used in evaluating the existing procedures.

Following this criterion, final-offer arbitration (FOA) is generally considered as superior to conventional arbitration (CA) as it induces disputants to make concessions and submit closer bids. CA and FOA are the two most prevalent forms of arbitration used in practice and perform as follows: CA mimics civil litigation in form since the arbitrator listens to the two sides'

1 A key difference between the arbitration of disputes and their resolution in a court concerns the nature of the fact-finder: both parties must agree to call in the arbitrator for the dispute to be settled by him (Manzini and Mariotti, 2001).

2 This problem has appeared in the economic literature under many different guises in a wide variety of situations (Grout, 1984; Grossman and Hart, 1986; Hart and Moore, 1990; Tirole, 1999).

settlement proposals and is free to impose any award of his choice, while FOA requires that each disputant proposes a final offer and the arbitrator must choose one of the two (i.e. the one that is closest to his own opinion of a fair settlement). The major criticism of CA in literature is that arbitrators are inclined—or perceived—to compromise between the parties' final positions, which encourages them to exaggerate claims and avoid concessions (Farber, 1981). To this respect, the proponents of FOA argue that this mechanism incites disputants to stake out more reasonable positions since no compromise is possible and the disputant with an extreme offer is likely to lose (Farber and Katz, 1979; Farber, 1980; Armstrong and Hurley, 2002). However, it appears empirically that arbitrators do much more than compromising, for example by referencing to some exogenous notion of an equitable award (Bloom, 1986; Farber and Bazerman, 1986). Furthermore, the fairly high frequency of negotiated settlements under FOA, which is often presented in industrial relations literature as evidence of its success, may indicate only that bargainers, while appearing to negotiate their own settlements, have correctly perceived the arbitrator's wishes and yielded to the incentives created by FOA to conform to them (Crawford, 1981).[3]

1.3 Our contribution

As mentioned above, the aim of our paper is to study the link between arbitration and bargaining in another way, by considering that arbitration may have not only allocative effects but also efficiency consequences, essentially in terms of promoting relationship-specific investments. This issue is analysed in a simple model which contains several restrictive features and assumptions that have been deliberately chosen to develop our main results and insights in a focused and simplified manner.

To fix ideas, we consider the long-term relationship that invariably exists between a firm and its workforce. The parties are contemplating undertaking relationship-specific investments, which would increase the size of the aggregate surplus (revenue) that they can generate from their relationship. Such investments may be of a wide variety and manifold. For example, investments by the firm can be in a new productive technology, while those by the workforce in productivity-enhancing work practices. However, since these investments are typically non-contractible, the parties will end up re-negotiating the terms of their relationship (that determines in particular the division of the aggregate surplus between them) after such investments are undertaken and sunk, but before the returns can be generated, and then consumed. The negotiations when conducted in the shadow of arbitration may generate an outcome which provides the parties with investment incentives that they would not have otherwise. As such, the parties consider (before undertaking any investments) whether or not to commit (via a costlessly enforceable contract) to call an arbitrator in the eventuality that they fail to reach a negotiated agreement by themselves, who would be empowered to impose a settlement. In order to get more general results, we allow the parties the option to agree to rely on arbitration randomly and we determine, on this basis, the optimal probability with which the arbitrator would be called (i.e. the optimal arbitration arrangement/regime) if *ex post* negotiations broke down. The optimal arbitration arrangement is the one which provides the best investment incentives *ex ante* and, hence, maximizes the *ex post* social surplus. In a first result, we highlight that arbitration cannot fully prevent the hold-up problem to occur since no party may be made residual claimant of

3 Aside from CA and FOA, some researchers have proposed alternative arbitration mechanisms following a normative perspective (e.g. Zeng, 2003; Dickinson, 2004).

his investment at the bargaining stage. Consequently, each partner under-invests relative to his first-best level of investment, regardless of the likelihood of arbitration. However, it is shown that the presence of the arbitrator (in the background of negotiations) may be crucial in determining the magnitude of these distorsions. The optimal regime is then the arbitration probability which minimizes partners' ability to engage in rent-seeking behaviours, by ensuring that their payoffs from disagreement are equalized and, thus, their bargaining powers are harmonized (which prevents them from making strategic use of their threat points in *ex post* bargaining). In this context, we show that some core parameters are particularly central to the determination of this optimal regime. A first one is the share of the returns generated by the parties' investments that the arbitrator will allocate to the worker in case of disagreement. This parameter captures the 'type' of the arbitrator and measures the importance of the workforce in his preferences. In this context, our model highlights that the parties should optimally commit to call an arbitrator ensuring a compromise by awarding a reasonable share of the surplus to the worker. Indeed, this type of arbitrator would maintain a balance in the parties' bargaining powers and then induce an optimal weight of their investment incentives. Furthermore, the optimal regime depends also upon technological parameters which characterize the productivities of parties' investments. In this regard, it is notably emphasized that the firm should be better protected—by weakening the likelihood of arbitration—when its investment becomes relatively more productive. The objective to balance the parties' bargaining powers should then be relaxed by adjusting for the relative productivities of their investments.

The remainder of the paper is organized as follows. The next section lays down our model and derives some preliminary results. Section 3 analyses the main issue of concern, namely, the optimal arbitration arrangement, and explores how it depends on the core parameters of the framework. Section 4 concludes. Proofs are relegated to the appendix.

2. The model

2.1 The framework

Given past investments, we consider that the firm and its workforce when working together in harmony and with full co-operation produce one unit of output per worker. Furthermore, given the most recent wage agreement, each worker's wage rate is $\overline{w}$, where $\overline{w} \in (0, 1)$. With this wage contract in place, we consider the following strategic interaction between the parties.

At date 0, the firm and its workforce choose and commit to the probability π with which they would call upon the services of an arbitrator in the eventuality that they fail to resolve any disputes by themselves in the future.

At date 1, given the choice made at date 0, the two parties consider whether or not to make some further investments that would enhance productivity. If the firm invests I_f and the workforce I_u, then the output per worker would increase by $V(I_f, I_u)$, provided, of course, that the two parties continue to work in harmony and with full co-operation.[4] It is assumed that V is concave, strictly increasing, and twice continuously differentiable.

4 Since these investments are non-contractible—because, for example, they cannot be verified by third parties (such as the courts)—this necessarily implies that they will be chosen non-cooperatively.

These investments are sunk at this date. The cost to player i ($i = f$, u) of investing I equals kI, where $k > 0$ is the common marginal cost of investment.[5]

At date 2, given the choices made at dates 0 and 1, the firm and its workforce engage in wage re-negotiations. We adopt the Nash bargaining solution (NBS) to describe the outcome of these negotiations, in which the threat point is identified with the parties' payoffs from disagreement.[6] If the parties strike a negotiated settlement on wage w, then the firm's payoff (per worker) is $1 + V(I_f, I_u) - w$ and each worker's payoff is w. However, it is possible that after the investments are undertaken and sunk, the two parties do not continue to work in harmony and with full co-operation, because they fail to strike a re-negotiated wage settlement. In that eventuality, one of two things can happen, depending on whether or not an arbitrator is called to impose a wage settlement.

If an arbitrator is not called (which happens with probability $1 - \pi$), then each worker continues to receive the old wage rate, $\overline{w}$, but in that case (for a variety of reasons) not all (and possibly none) of the returns from the investments are realized. We formalize and parameterize this possibility by assuming that only a fraction λ of the additional surplus $V(I_f, I_u)$ materializes in that case, where $\lambda \in [0, 1)$. The disagreement payoffs in the eventuality that an arbitrator is not called are as follows: the firm's payoff (per worker) is $1 + \lambda V(I_f, I_u) - \overline{w}$ and each worker's payoff is $\overline{w}$. These payoffs capture the notion that the parties stick with the existing wage contract, and consequently, disagreement entails that only a fraction λ of the benefits from the investments are obtainable. This makes sense, since if the parties do not reach an agreement on their own and have no arbitrator who would impose a settlement, then disagreement entails some degree of inefficiency (as the dispute remains unresolved and the parties do not work with full co-operation). The parameter λ (i.e. the workforce resentment factor) captures the extent to which the workforce refuses to cooperate with the firm if they end up working at the old wage contract, implying that the workforce's resentment is decreasing in λ.

If, on the other hand, an arbitrator is called (which happens with probability π), then he would be empowered to impose a wage settlement. We assume that the arbitrator's wage decision is:

$$w^A = \overline{w} + \alpha V(I_f, I_u), \tag{1}$$

where $0 < \alpha < 1$. It seems reasonable to consider that the wage given by the arbitrator to each worker is the old wage rate plus some fraction of the additional surplus V generated by the parties' investments. The additional surplus could then be interpreted as a mapping that defines the transformation of parties' investments into the arbitrator's sentence when the parties fail to resolve the dispute by themselves. In other words, the parameter α captures the type of the arbitrator and could be viewed as a measure ot the importance of the workforce in the arbitrator's preferences, implying that this importance is increasing in α. The disagreement payoffs are then as follows: the firm's payoff (per worker) is $1 + V(I_f, I_u) - w^A$, where w^A is defined in eq. (1), and each worker's payoff is w^A.

5 For simplicity and without loss of generality, we assume that the firm and its workforce have an identical and constant marginal cost of investment. This assumption has been introduced for algebraic convenience and could be relaxed without altering the gist of our arguments.

6 See Muthoo (1999) for a discussion and analyses of the strategic (or non-cooperative) foundations of the NBS, where, using various versions of Rubinstein's alternating-offers model, he addresses the questions of why, when and how to use this bargaining solution concept.

These payoffs capture the notion that since the dispute is resolved by the arbitrator, the parties would subsequently work in harmony and with full co-operation.

In consequence, the expected payoffs to the firm and each worker respectively from disagreement are as follows:

$$d_f \equiv 1 - \overline{w} + [\pi(1-\alpha-\lambda) + \lambda]V(I_f, I_u) \tag{2}$$

$$d_u \equiv \overline{w} + \pi\alpha V(I_f, I_u). \tag{3}$$

We would like to emphasize that although efficient bargaining implies that in equilibrium the parties will reach a negotiated settlement—and thus the arbitrator would not be called upon—the size of the aggregate surplus depends on the equilibrium investments made at date 1. The latter, it will be shown below, depend on the equilibrium distribution of the surplus between the two parties, which, in turn, depends on the parties' relative bargaining powers that are in general influenced by their respective expected disagreement payoffs (and hence by the probability with which they have agreed at date 0 to call upon the arbitrator).

The core parameters of our model include the parameter α that captures the arbitrator's preferences, the parameter λ that captures the extent of non-cooperation when the workforce continues to work according to the old wage contract, and the technological parameters as captured by the properties of the function V. Our main objective is to study and develop the role and impact of these parameters on the optimal arbitration arrangement (that is, the optimal value of π). The optimal value of π is the one which maximizes the net economic surplus and hence the equilibrium incentives to invest.[7]

Before analysing this main question of interest, some comments on the framework and its restrictive assumptions are in order. First, we assume that the firm and its workforce incur no cost in employing the arbitrator, which is obviously not consistent with practice. However, this assumption is made to alleviate notations and without loss of generality since our results on the efficiency consequences of arbitration would carry over if the cost of employing the services of the arbitrator would comprise a fixed cost. Indeed, the parties investment incentives are driven by their marginal returns on investment, which are unaffected by fixed costs. Nevertheless, it is clear that the introduction of such costs would have distributive implications—if these costs are different between the parties—by altering the threat points and, hence, the Nash-bargained utility payoffs. However, the analysis of such implications are beyond the scope of this paper which is focused on efficiency considerations.

Second, by using the NBS, we implicitly rule out arbitrator's intervention in equilibrium since the parties will call upon an arbitrator only if they fail to reach a negotiated settlement, which never happens: Pareto-efficiency property implies that parties always reach an agreement since disagreement entails some degree of inefficiency due to the potential workforce's resentment. This restriction is somewhat interesting since it enables us to highlight that the mere presence of the arbitrator in the background of negotiations may entirely drive their outcome and prevent the parties from behaving opportunistically *ex post* (thereby enhancing investment incentives *ex ante*). However, this framework is also puzzling

7 π could also be considered as an institutional choice variable that may be viewed as the availability of alternative dispute resolution mechanisms (such as arbitration) in the society.

since arbitration obviously occurs in practice, and an extension to tackle this issue is thus proposed in conclusion.

Third, following the literature on arbitration, we consider that the so-called 'arbitrator's idea of a fair settlement' (i.e. α) is exogenously defined and captures how the economic environment and arbitrator's reasoning influence the arbitrated outcome (see, e.g. Armstrong and Hurley, 2002). This assumption is made for simplicity and could be relaxed by considering that α is endogenously chosen by this third player. Such an extension is proposed in conclusion of the paper.

Finally, the parameter λ is also exogenous and this is naturally a restrictive assumption. However, this parameter captures the extent to which the workforce refuses to cooperate with the firm if they end up working at the old wage, and its value may be exogenously determined by both behavioural and institutional factors. Indeed, following reciprocal motives, a worker who feels he has been mistreated by a firm may engage in strike or acts of sabotage in order to punish the employer for being unfair. Fehr and Gächter (2000) provide an overview of the considerable evidence for such reciprocity/fairness-based behaviours arising in different economic and social contexts (public good games, labour market interactions . . .). The authors emphasize that the population is partly composed of such 'reciprocal types', which would be characterized by a low value of λ in our framework.[8] However, as mentioned above, this ability to retaliate may also be linked to institutional environment since it may be correlated with laws on the right to strike or with the power of trade unions in the jurisdiction concerned.

We now proceed to solve the above described three-stage dynamic game via the usual backward induction procedure. We therefore begin by characterizing the equilibrium bargaining outcome at date 2, conditional on an arbitrary set of choices made at dates 0 and 1.

2.2 Bargaining outcome

As stated above, we use the—symmetric—NBS to characterize the equilibrium bargaining outcome at date 2. Applying the NBS in which the threat points are given by eqs (2) and (3), simplifying and re-arranging terms, it follows that the Nash-bargained utility payoffs to the firm and each worker are respectively as follows:

$$u_f^N = 1 - \overline{w} + \Omega_f V(I_f, I_u) \qquad \text{and} \qquad u_u^N = \overline{w} + \Omega_u V(I_f, I_u), \tag{4}$$

where

$$\Omega_f = \frac{\pi(1-\lambda-2\alpha)+(1+\lambda)}{2} \qquad \text{and} \tag{5}$$

$$\Omega_u = \frac{\pi(2\alpha+\lambda-1)+(1-\lambda)}{2}. \tag{6}$$

These expressions can be interpreted in terms of partitions of the old and the additional surpluses. The partition of the old surplus—which is one unit of output per worker and existed before any new investments were undertaken—is defined by the old wage contract in which

8 These behavioural considerations justify that the worker does not express feelings of resentment towards the employer in case of arbitration. Indeed, his wage does not remain at its old value in this case and the worker may be grateful to the employer for his *ex ante* commitment to call an arbitrator in case of bargaining failure *ex post*.

each worker receives $\overline{w}$ units of the surplus, and the firm obtains the remainder. The additional surplus, which is $V\ (I_f,\ I_u)$ units of output per worker, is split in such a way that party i gets a fraction Ω_i of it (with $\Omega_f + \Omega_u = 1$). Notice that the influence of π on the party i's bargaining power, captured by Ω_i, depends on the type of the arbitrator (captured by the parameter α) and the workforce resentment factor (determined by the value of λ).

2.3 Equilibrium investments

Having characterized the outcome of the negotiations at date 2 over the partition of the aggregate surplus (for an arbitrary set of choices made at dates 0 and 1), we now determine the equilibrium investment levels chosen by the parties at date 1. However, we first characterize the first-best investment levels.

The first-best (or Pareto-efficient) investment levels I_f^e and I_u^e maximize the difference between the aggregate surplus and the total cost of the investments, $V(I_f, I_u) - k(I_f + I_u)$. As such I_f^e and I_u^e are the unique solutions to the following first-order conditions:

$$V_1(I_f, I_u) = k \quad \text{and} \quad V_2(I_f, I_u) = k, \tag{7}$$

where V_1 and V_2 respectively denote the first-order derivatives of V with respect to its first and second arguments. They respectively define the aggregate marginal benefits of the firm's and the workforce's investments. Efficiency requires that such marginal returns be equated to the corresponding marginal costs.

We now characterize the Nash equilibrium investment choices made at date 1. Given an arbitrary value of π, the equilibrium investment levels, denoted by I_f^* and I_u^*, comprise the unique solution to the following first-order conditions:

$$\Omega_f V_1(I_f, I_u) = k \quad \text{and} \quad \Omega_u V_2(I_f, I_u) = k \tag{8}$$

The left-hand sides of these expressions are respectively the parties' private marginal benefits from their respective investments, which are strictly less than the corresponding aggregate marginal benefits from such investments (given that $\Omega_i < 1$). This observation immediately implies that the players' Nash equilibrium investment levels are below their corresponding first-best levels, whatever the arbitration arrangement (i.e. for any $\pi \in [0, 1]$). This under-investment result is essentially due to the fact that whatever the arbitration terms a player receives less than the full marginal benefit of his investment.

Proposition 1 Whatever is the arbitration arrangement, each party under-invests relative to his first-best level of investment.

This result arises because after undertaking and sinking his investment, each party is held-up by the other party. It is precisely because of this problem that it makes economic sense for parties to commit *ex ante*—well before any investments are undertaken—on the probability with which they would call upon the arbitrator. Indeed, since I_f^* and I_u^* depend on π, it is clear that π will be crucial in determining the size and direction of the investment distortions. In other words, different values of π will induce different equilibrium investment levels, and hence different levels of the net economic surplus. The optimal arbitration arrangement (i.e. the optimal value of π) is the one which minimizes the distortions of the productivity-enhancing, relationship-specific investments from their first-best levels. We now turn to address this main issue.

3. Optimal arbitration

Moving backwards to date 0, we now turn attention to the main issue of concern, namely, the study of the determination of the optimal value of π. The optimal π is one that maximizes the equilibrium net surplus $S(I_f^*, I_u^*)$, where $S(.,.)$ is defined as follows

$$S(I_f, I_u) \equiv 1 + V(I_f, I_u) - k(I_f + I_u) \tag{9}$$

The equilibrium net surplus depends on the value of π indirectly, via its influence on the equilibrium investment levels. We write it as $S^*(\pi)$. So, the optimal arbitration arrangement, denoted by π^*, is the value of π over the interval [0, 1] that maximizes $S^*(\pi)$ and thus provides the partners with the best investment incentives. As noted above, it is clear that the optimal π will provide relatively higher equilibrium surplus, and relatively smaller distortions of the equilibrium investment levels from their respective first-best levels. If $\pi^* = 1$ then the firm and its workforce should optimally contract to call an arbitrator in the eventuality that they fail to reach a negotiated settlement, while if $\pi^* = 0$ then it is optimal for the parties to resolve their dispute without the presence of the arbitrator. If instead $0 < \pi^* < 1$ then it is optimal for the partners to call upon the arbitrator with probability π^*. The following result is useful in developing our subsequent analysis of optimal arbitration arrangement.

Lemma 1 Define $\Delta = (2\alpha + \lambda - 1)/2$. For any $\pi \in [0, 1]$,

$$\text{If } \Delta > 0, \text{ then } \frac{\partial S^*(\pi)}{\partial \pi} \gtreqless 0 \Leftrightarrow V_{22}(\Omega_u)^4 - V_{11}(\Omega_f)^4 \gtreqless 0, \tag{10}$$

$$\text{If } \Delta < 0, \text{ then } \frac{\partial S^*(\pi)}{\partial \pi} \gtreqless 0 \Leftrightarrow V_{11}(\Omega_f)^4 - V_{22}(\Omega_u)^4 \gtreqless 0, \tag{11}$$

where V_{11} and V_{22} are evaluated at the equilibrium investment levels.

Proof See Appendix 1 □

Without imposing any further restrictions on the function V, it is evident from Lemma 1 that not much can be said about π^*. In what follows, therefore, we derive a number of results about π^* in the context of some additional parametric restrictions on V. We begin with the case in which V_{11} and V_{22} are identical, and equal some strictly negative constant, but we do not impose any restriction on the sign of V_{12}.

Proposition 2 Assume that V_{11} and V_{22} are identical. Define:

$$\underline{\alpha} = \frac{1-\lambda}{2} \text{ and } \overline{\alpha} = \frac{1}{2}.$$

(*a*) If either $\lambda = 0$, or $\lambda > 0$ and $\alpha < \underline{\alpha}$, then the optimal rule is $\pi^* = 0$.

(b) if $\lambda > 0$ and $\underline{\alpha} < \alpha \leq \overline{\alpha}$, then the optimal rule is $\pi^* = 1$

(c) Otherwise (i.e. if $\lambda > 0$ and $\alpha > \overline{\alpha}$), then the firm and its workforce should optimally contract to call an arbitrator with probability π^*, where:

$$\pi^* = \frac{\lambda}{2\alpha + \lambda - 1}. \tag{12}$$

Proof See Appendix 2 □

While the hypothesis of this proposition restricts the class of applicable functions V (e.g. some subclass of the quadratic family), its implications are powerful, as they reappear in the context of a larger class of utility functions (see below) and offer a significant benchmark. This result says that the workforce resentment factor (i.e. λ) and the arbitrator's type (i.e. α) determine the optimal contractual arrangement.

The optimal value π^* from Proposition 2(c) is the unique value of π such that $\Omega_f = \Omega_u = 1/2$. In other words, it is the unique arbitration arrangement that ensures that the parties' payoffs from disagreement are equalized, that the power relations within work are harmonized, and that the net surplus is split equally. This balancing out of bargaining powers is relevant, especially in relation to the provision of investment incentives to the two parties. Indeed, the players cannot make strategic use of their threat points in bargaining situations they encounter throughout their labour relationships. Both parties are thus willing to invest optimally because neither fears expropriation by the other partner. Thus, the optimal arbitration arrangement designs a particular game between the partners, one in which the ability of individuals to engage in rent-seeking behaviour is minimized.

In summary, the terms of the optimal arbitration agreement prevent the parties from behaving opportunistically *ex post*, thereby promoting efficient second-best investments *ex ante*. Of course, the partners' equilibrium investments may differ as they also depend on the structure of the function V. Turning to the optimal value of π itself, eq. (12) highlights that it is increasing in λ. This implies that the more the firm is able to capture the benefits from the investments when the parties fail to reach an agreement (i.e. the higher is the value of λ), the larger should be the probability with which the partners would call upon the arbitrator. Although unappealing at first, this relationship is consistent with the idea that a key force underlying the determination of the optimal contractual arrangement concerns the tendency to equalize players' bargaining powers. In other words, an increase in λ, which increases the firm's bargaining power, should be (partially) offset by an increase in π.[9]

In this context, it is straightforward to interpret the other parts of Proposition 2 and the role played by the parameter α in determining the optimal arbitration arrangement. Considering that $\underline{\alpha} < \alpha < \overline{\alpha}$, the result stated in Proposition 2(b) shows that it is optimal to call upon the arbitrator ($\pi^* = 1$). This is because this kind of arbitrator, by choosing a reasonable sentence (i.e. an intermediate value of α), would help induce an optimal compromise in the provision of investment incentives to the parties. Following Proposition 2(a), the same logic holds if $\alpha < \underline{\alpha}$ since, in this case, the presence of the arbitrator decreases sharply the worker's bargaining power and does not maintain a balance in the provision of investment incentives between the two parties, which implies that the optimal behaviour is not to call the arbitrator ($\pi^* = 0$).

We now examine another widely used class of functions of the Cobb-Douglas type, namely $V = (I_f)^{\eta_f}(I_u)^{\eta_u}$, where $0 < \eta_i < 1$ and $\eta_f + \eta_u < 1$. These functions are smooth, strictly increasing in each of their two arguments, and strictly concave. Furthermore, since $V_{12} > 0$, the investments are complements. The parameters η_f and η_u capture the parties' productivities and the ratio η_i/η_j $(i \neq j)$ is therefore a measure of productive heterogeneity

9 Furthermore, notice that the value of λ determines the threshold $\underline{\alpha}$ as well as the equilibrium contractual arrangement π^*: we can notice that $\underline{\alpha}'(\lambda) < 0$ which implies that an increase in λ induces the probability that $\pi^* = 1$ to be more likely (following the idea mentioned above).

between the two players. Our main objective is to analyse the impact of these parameters (as well as the impact of the arbitrator's preferences and the workforce resentment factor) on determination of the optimal arbitration arrangement. Our main results are summarized in the following.

Proposition 3 Assume that $V = (I_f)^{\eta_f}(I_u)^{\eta_u}$, where $0 < \eta_f < 1$, $0 < \eta_u < 1$ and $\eta_{\mathrm{f}} + \eta_{\mathrm{u}} < 1$. Define:

$$\underline{\alpha} = \frac{1-\lambda}{2}, \quad \overline{\alpha} = \frac{\theta}{1+\theta}, \quad \text{and } \tilde{\lambda} = \frac{1-\theta}{1+\theta},$$

$$\text{where } \theta = \sqrt{\frac{\eta_u(1-\eta_f)}{\eta_f(1-\eta_u)}}.$$

(a) If $\lambda \leq \tilde{\lambda}$, then the optimal rule is $\pi^* = 0$.

(b) If $\lambda > \tilde{\lambda}$ and $\underline{\alpha} < \alpha \leq \overline{\alpha}$, then the optimal rule is $\pi^* = 1$.

(c) Otherwise (i.e. if $\lambda > \tilde{\lambda}$ and $\alpha > \overline{\alpha}$), then the firm and its workforce should optimally contract to call an arbitrator with probability π^* where:

$$\pi^* = \frac{\lambda}{2\alpha+\lambda-1} - \frac{(1-\theta)}{(1+\theta)(2\alpha+\lambda-1)}. \tag{13}$$

Proof See Appendix 3 □

We know that as far as the objective to balance the bargaining strengths of the parties is concerned, the first term in eq. (13) rules out opportunistic expropriation by either party by offsetting the players' bargaining powers. Consider the case in which $\eta_f = \eta_u$, that is, the function V is symmetric in partners' investments. Proposition 3(c) implies that the optimal contractual arrangement is identical to the arrangement identified in Proposition 2. Indeed, if $\eta_f = \eta_u$, then $\theta = 1$ and the second term in π^* disappears. The earlier discussion on Proposition 2 therefore applies to this case as well, even though the third partial derivatives of V here are not zero.

Turning to the case of productive heterogeneity between the players, now consider the case in which $\eta_f \neq \eta_u$. The second term in (13) relaxes the objective to harmonize the parties' *ex ante* bargaining powers by adjusting for the relative importance of the investments. To see this, substitute the optimal value π^* from Proposition 3(c) into eqs (5) and (6) to obtain:

$$\Omega_f^* = \frac{1}{1+\theta} \quad \text{and} \quad \Omega_u^* = \frac{\theta}{1+\theta},$$

which implies that:

$$\begin{aligned} \eta_f > \eta_u (\text{hence } \theta < 1) \Leftrightarrow \Omega_f^* > \Omega_u^*, \\ \eta_f < \eta_u (\text{hence } \theta > 1) \Leftrightarrow \Omega_f^* < \Omega_u^*. \end{aligned}$$

The intuition for this follows from the fact that when party i's investment is more productive than j's, then the second term of the optimal value π^* translates into a greater bargaining strength of i compared to j, which he will use to obtain a larger share of the economic surplus. Therefore, the comparative importance/productivities of the two parties for the generation of the returns is a key force determining the optimal arbitration arrangement. If the firm's investment is more productive than the workforce's investment (i.e. $\eta_f > \eta_u$), then the firm's share of the additional surplus V generated from the two parties' investments increases *ceteris paribus*. This will cause the firm's investment incentives to increase, and the workforce's incentives to decrease, so that $I_f^* > I_u^*$. The opposite is true if $\eta_f < \eta_u$. Ultimately the optimal value π^* from Proposition 3(c) ensures—by weighting the marginal returns to investing of both partners and balancing out their bargaining powers—that the sum of their contributions, and thus economic surplus, is maximized. Indeed, conducting a comparative statics analysis on π^*, it is straightforward to show that the optimal probability of arbitration is decreasing in η_f (i.e. $\partial\pi^*/\partial\eta_f < 0$), which emphasizes that the firm should be better protected—by lessening the likelihood of arbitration—when its investment becomes more productive.

The analysis in this section does not offer a general characterization of the optimal arbitration terms when both parties can invest and a stochastic contract is possible. Additional research on this issue remains to be done. However, in spite of the specific parametrizations of V, a main insight of Propositions 2–3 is worth emphasizing.[10] In equilibrium, three components matter in determining the optimal contractual arrangement (which is designed so as to balance out the bargaining powers of both parties and their investment incentives). The first is given by the arbitrator-preference parameter (i.e. α) that captures the type of the arbitrator. The second element is the workforce resentment factor (i.e. λ) that captures the extent of non-cooperation when the workforce continues to work according to the old wage contract. The third component is represented by the technological parameters (i.e. η_f and η_u) that capture the productivities of each partner's relationship-specific investment. In this perspective, giving a larger share of V to the party whose investment is relatively more important than the other party's investment is consistent with the aim of endowing this player with greater bargaining power and thus inducing him to provide the optimal level of investment. Indeed, the party i's relative bargaining power determines i's marginal returns on his investment, which, in turn, determines his marginal incentives.

4. Conclusion

Arbitration can matter. Its presence (in the background) may have not only distributive but also efficiency consequences within long-term relationships, by influencing each partner's incentives to undertake productivity-enhancing, relationship-specific investments. As such, the optimal terms of arbitration—whether or not to institute it, and if so, what type—need to be based, if not wholly, at least substantially on these observations. In this paper, we have formally explored this issue from this perspective. We show that the optimal arbitration arrangement depends on various key variables including the arbitrator' preferences, the parties' bargaining powers, and the comparative parties' productivities in generating the returns from their investments. A fundamental principle that governs the optimal arbitration arrangement is to equalize the parties' economic positions within work (thereby

10 We could finally consider a simple parametric case in which investments are independent ($V_{12} = 0$), by assuming for example that $V(I_f, I_u) = \eta_f I_f^\beta + \eta_u I_u^\beta$ where $0 < \beta < 1$. However, the results for this case would be identical to those identified in Proposition 2.

balancing out parties' bargaining powers within their relationship) and to weight their investment incentives. Hence, by giving backing to the wage re-negotiation process, the terms of the arbitration arrangement prevent parties behaving opportunistically *ex post*, thereby promoting efficient investments *ex ante*.

As noted in Introduction, to the best of our knowledge this is the first paper that applies the incomplete contracts approach to the study of arbitration. The incomplete contracts approach is powerful because it is based on the premiss that transaction costs plague human interaction, and the consequent potent observation that property rights have efficiency consequences, which, of course, means that Coase's Theorem fails to apply. However, we have only just scratched the surface of the various matters than impinge on this issue. Much more work needs to be done to improve our understanding of the efficiency implications of arbitration.

In this context, several extensions and generalizations suggest themselves. First, a simplifying but somewhat restrictive assumption that underlies the model is that disagreement (and arbitrator's intervention) is ruled out in equilibrium. Although many of our main qualitative insights would be robust to allowing arbitration to occur with positive probability, it would be useful to formally address this issue partly because if the likelihood of dispute is real then the parties' incentives would be altered from our current analysis. A natural way to make disagreement possible in equilibrium would be to introduce asymmetric information in the present setup.

Second, the above analysis concerns the optimal arbitration arrangement that, once in place, will maximize the social surplus. This approach should be interpreted by considering that π is chosen by the regulator to provide the parties with the best investment incentives. From a practical perspective, following the arguments exposed in note 7, the value of π^* may capture the actions taken by public authorities in order to promote the use of arbitration as an alternative to standard litigation. For example, it is well-known that alternative dispute resolution mechanisms are widely used in the USA, while it less common in France, and it can be related to institutional choices. The weakness of this idea is that it is silent on the question of whether, or under what circumstances, will the parties agree in equilibrium to commit to call the arbitrator (in the eventuality that they fail to reach a negotiated settlement). Therefore, it would be interesting to consider that the parties have the opportunity to negotiate a (legally enforceable) contract governing the arbitration terms.

Third, an alternative approach could involve to endogenize the arbitrator's behaviour by considering that the parameter α is chosen to maximize the equilibrium aggregate surplus. The idea behind this story would be to analyse how the arbitrator's award might shape investment incentives, and to determine what is the optimal award on this basis, depending on some crucial technological and institutional parameters (i.e. λ, π, η_f and η_u). In other words, the aim of this alternative model would be to emphasize how the arbitrator should make one party prevailing over the other one—by choosing the value of α—in order to enhance the efficiency of the labour relationship.[11]

Finally, a further step towards realism would be also to embed the present framework within a dynamic game in which parties (and arbitrator) interact repeatedly over time, learn about their investments, and build up some trust.

11 In this framework, the parameter π would be exogenously defined.

Acknowledgements

We wish to thank two anonymous referees and the editor for their useful comments.

References

Armstrong, M.J. (2004) A comparison of arbitration procedures for risk-averse disputants, *Decision Sciences*, **35**, 639–64.

Armstrong, M.J. and Hurley, W.J. (2002) Arbitration using the closest offer principle of arbitrator behaviour, *Mathematical Social Sciences*, **43**, 19–26.

Bloom, D.E. (1986) Empirical models of arbitrator behaviour under conventional arbitration, *Review of Economics and Statistics*, **68**, 578–85.

Crawford, V.P. (1981) Arbitration and conflict resolution in labor-management bargaining, *American Economic Review Papers and Proceedings*, **71**, 205–10.

Crawford, V.P. (1985) The role of arbitration and the theory of incentives, in A. Roth (ed.) *Game-Theoretic Models of Bargaining*, Cambridge University Press, Cambridge, MA.

Deck, C.A. and Farmer, A. (2009) Strategic bidding and investments in final-offer arbitration: theory and experimental evidence, *Journal of Economic Behaviour & Organization*, **70**, 361–73.

Deck, C., Farmer, A., and Zeng, D.-Z. (2007) Amended final-offer arbitration over an uncertain value: a comparison with CA and FOA, *Experimental Economics*, **10**, 439–54.

Dickinson, D.L. (2004) A comparison of conventional, final-offer, and 'combined' arbitration for dispute resolution, *Industrial and Labor Relations Review*, **57**, 288–301.

Farber, H.S. (1980) An analysis of final-offer arbitration, *Journal of Conflict Resolution*, **24**, 683–705.

Farber, H.S. (1981) Splitting-the-difference in interest arbitration, *Industrial and Labor Relations Review*, **35**, 70–7.

Farber, H. and Bazerman, M.H. (1986) The general basis of arbitrator behaviour: an empirical analysis of conventional and final-offer arbitration, *Econometrica*, **54**, 1503–28.

Farber, H.S. and Katz, H.C. (1979) Interest arbitration, outcomes, and the incentives to bargain, *Industrial and Labor Relations Review*, **33**, 55–63.

Fehr, E. and Gächter, S. (2000) Fairness and retaliation, *Journal of Economic Perspectives*, **14**, 159–81.

Grossman, S. and Hart, O. (1986) The costs and benefits of ownership: a theory of vertical and lateral integration, *Journal of Political Economy*, **94**, 691–64.

Grout, P. (1984) Investment and wages in the absence of binding contracts: a Nash bargaining approach, *Econometrica*, **52**, 449–60.

Hart, O.D. and Moore, S.J. (1990) Property rights and the nature of the firm, *Journal of Political Economy*, **98**, 1119–58.

Manzini, P. and Mariotti, M. (2001) Perfect equilibria in a model of bargaining with arbitration, *Games and Economic Behaviour*, **37**, 170–95.

Muthoo, A. (1999) *Bargaining Theory with Applications*, Cambridge University Press, Cambridge, MA.

Tirole, J. (1999) Incomplete contracts: where do we stand? *Econometrica*, **67**, 741–81.

Zeng, D.-Z. (2003) An amendment to final-offer arbitration, *Mathematical Social Sciences*, **46**, 9–19.

Appendix 1: Proof of Lemma 1

Differentiating S^* with respect to π, we obtain that:

$$\frac{\partial S^*(\pi)}{\partial \pi} = V_1 \frac{\partial I_f^*}{\partial \pi} + V_2 \frac{\partial I_u^*}{\partial \pi} - k\left(\frac{\partial I_f^*}{\partial \pi} + \frac{\partial I_u^*}{\partial \pi}\right)$$

Since, by eqs (5) and (6), $\Omega_f + \Omega_u = 1$, it follows using eq. (8) that:

$$\frac{\partial S^*(\pi)}{\partial \pi} = \Omega_u V_1 \frac{\partial I_f^*}{\partial \pi} + \Omega_f V_2 \frac{\partial I_u^*}{\partial \pi} \tag{A.1}$$

Differentiating the first-order conditions in (8) with respect to π, we find that:

$$\frac{\partial I_f^*}{\partial \pi} = \frac{\Delta}{\Sigma}\left(\frac{V_2 V_{12}}{\Omega_u} + \frac{V_1 V_{22}}{\Omega_f}\right), \tag{A.2}$$

$$\frac{\partial I_u^*}{\partial \pi} = -\frac{\Delta}{\Sigma}\left(\frac{V_1 V_{12}}{\Omega_f} + \frac{V_2 V_{11}}{\Omega_u}\right), \tag{A.3}$$

where $\Delta = (2\alpha + \lambda - 1)/2$ and $\Sigma = V_{11}V_{22} - (V_{12})^2$, with all the first-order and second-order partial derivatives evaluated at the equilibrium investment levels I_f^* and I_u^*.

After substituting for the derivatives of the equilibrium investments (using (A.2) and (A.3)) into the right-hand side of (A.1), simplifying, re-arranging terms, and finally using the first-order conditions in (8) to substitute for V_1 and V_2, we obtain that:

$$\frac{\partial S^*(\pi)}{\partial \pi} = \frac{k\Delta}{\Sigma(\Omega_f\Omega_u)^3}[V_{22}(\Omega_u)^4 - V_{11}(\Omega_f)^4].$$

The Lemma follows immediately since $k > 0$, $\Sigma > 0$, $\Omega_f > 0$, and $\Omega_u > 0$ □

Appendix 2: Proof of Proposition 2

Step 1. Consider first the case where $\Delta > 0$. Under the (additional) hypothesis of this proposition, it follows from Lemma 1 that for any $\pi \in [0, 1]$,

$$\frac{\partial S^*(\pi)}{\partial \pi} \gtreqless 0 \Leftrightarrow \Omega_f - \Omega_u \gtreqless 0.$$

Then, using (5) and (6), we obtain that:

$$\frac{\partial S^*(\pi)}{\partial \pi} \gtreqless 0 \Leftrightarrow \rho(\pi) \gtreqless 0, \text{ where } \rho := \pi(1 - \lambda - 2\alpha) + \lambda.$$

We then note that:

$$\rho(\pi) \gtreqless 0 \Leftrightarrow \pi^* \gtreqless \pi,$$

where π^* is stated in (12).

Hence, since $\lambda \in [0, 1)$, it follows that $\pi^* \nless 0$ and we obtain:

$$\pi^* \geq 0 \Leftrightarrow \lambda \geq 0 \text{ and } \pi^* \gtreqless 1 \Leftrightarrow \alpha \lesseqgtr \overline{\alpha}, \tag{A.4}$$

where $\overline{\alpha}$ is stated in Proposition 2.

This implies that, when $\Delta > 0$, π^* is the unique stationary (or turning) point of the function S^*.

Step 2. Consider now the case where $\Delta < 0$. It follows from Lemma 1 that for any $\pi \in [0, 1]$,

$$\frac{\partial S^*(\pi)}{\partial \pi} \gtreqless 0 \Leftrightarrow \Omega_u - \Omega_f \gtreqless 0.$$

Then, using (5) and (6), we obtain that:

$$\frac{\partial S^*(\pi)}{\partial \pi} \gtreqless 0 \Leftrightarrow \varphi(\pi) \gtreqless 0, \text{ where } \varphi := \pi(2\alpha + \lambda - 1) - \lambda.$$

However, since $\Delta < 0$ and $\lambda \in [0, 1)$, it is straightforward to show that $\varphi(\pi) < 0$ which implies the following result:

$$\frac{\partial S^*(\pi)}{\partial \pi} < 0 \text{ for any } \pi \in [0, 1]. \tag{A.5}$$

Proposition 2 follows immediately from the above results (in (A.4) and (A.5)) and the fact that $\Delta \gtreqless 0 \Leftrightarrow \alpha \gtreqless \underline{\alpha}$, where $\underline{\alpha}$ is stated in Proposition 2. □

Appendix 3: Proof of Proposition 3

For simplicity, we assume that $\Delta > 0$ (i.e. $\alpha > \underline{\alpha}$, where $\underline{\alpha}$ is stated in this Proposition). This assumption has been introduced in order to develop understanding about the role of the factors under study in a sharper manner and could be easily relaxed without altering the gist of our arguments.

From the first-order conditions in (8), we obtain the following relationship between the equilibrium investment levels:

$$I_f^* = \left(\frac{\Omega_f \eta_f}{\Omega_u \eta_u}\right) I_u^*.$$

It follows from an application of Lemma 1—after substituting for the equilibrium values of V_{11} and V_{22}, using the above relationship between the equilibrium investment levels, and finally substituting for Ω_f and Ω_u—that:

$$\frac{\partial S^*(\pi)}{\partial \pi} \gtreqless 0 \Leftrightarrow \psi(\pi) \lesseqgtr 0,$$

$$\text{where } \psi := \left(\frac{1-\theta}{2}\right) + \left(\frac{1+\theta}{2}\right)[\pi(2\alpha - 1) - \lambda(1 - \pi)].$$

We then note that:

$$\psi(\pi) \gtreqless 0 \Leftrightarrow \pi \gtreqless \pi^*,$$

where π^* is stated in (13). This implies that π^* is the unique stationary (turning) point of the function S^*. Proposition 3 follows immediately from the above results, the assumption that $\Delta > 0$, and the following two results:

$$\pi^* \gtreqless 0 \Leftrightarrow \lambda \gtreqless \tilde{\lambda} \text{ and } \pi^* \gtreqless 1 \Leftrightarrow \alpha \lesseqgtr \overline{\alpha},$$

where $\tilde{\lambda}$ and $\overline{\alpha}$ are stated in the Proposition. □