Edinburgh School of Economics

Discussion Paper Series

Number 74

A Dynamic Theory of Holdup

Yeon-Koo Che (University of Wisconsin-Madison)
Jozsef Sakovics (University of Edinburgh)

Date
November 2001

Published by
School of Economics
University of Edinburgh
30 -31 Buccleuch Place
Edinburgh EH8 9JT
+44 (0)131 650 8361
http://www.ed.ac.uk/schools-departments/economics

THE UNIVERSITY *of* EDINBURGH

A Dynamic Theory of Holdup

Yeon-Koo Che[*] József Sákovics[†]

November 20, 2001

Abstract

The holdup problem arises when parties negotiate to divide the surplus generated by their ex ante noncontractable investments. We study this problem in a model which, unlike the stylized static model, allows the parties to continue to invest until they agree on the terms of trade. These possible investment dynamics overturn the conventional wisdom dramatically. First, the holdup problem need not entail underinvestment-type inefficiencies when the parties are sufficiently patient. Second, inefficiencies can arise unambiguously, but the reason for their occurrence differs from the one recognized by the static model. This latter finding sheds new light on the design of contracts and organizations.

Key words: Investment, Bargaining with an endogenous pie, contribution games.

[*]Department of Economics, University of Wisconsin-Madison.

[†]University of Edinburgh.

Both authors are grateful for comments from Jennifer Arlen, Ian Gale, Don Hausch, Bart Lipman, Bentley MacLeod, Leslie Marx, Steven Matthews, John Moore, Larry Samuelson, Jonathan Thomas and the seminar participants at the USC Law School (John Olin Lecture), the University of Edinburgh, the University of Texas and the University of Washington. Part of this research was conducted while the first author was visiting, and the second author was employed at, the Institut d'Anàlisi Econòmica (CSIC) in Barcelona. The former author wishes to acknowledge their hospitality as well as the financial support provided through the Spanish Ministry of Education and Culture. He also acknowledges financial support from the Wisconsin Alumni Research Foundation.

1 Introduction

A recurring problem in economics concerns an agent who must make a sunk investment without being guaranteed to receive its social return. One such phenomenon arises when trading partners negotiate to divide their trade surplus after making relationship-specific investments. This problem, referred to as holdup, is inherent in many bilateral exchanges. For instance, workers and firms often invest in firm-specific assets prior to negotiating for wages. Manufacturers and suppliers often customize their equipment and production processes to the special needs of their partners, knowing well that future (re)negotiation will confer part of the benefit from customization to their partners. While these investments increase the value of their transaction, they often occur without contractual protection, thus putting investing parties at the risk of holdup. The conventional wisdom is that the risk of holdup would lead parties to underinvest in specific assets (see Grout, 1984; and Tirole, 1986). The purpose of this paper is to reexamine this conventional wisdom.[1]

Our point of departure is the observation that the stylized model predicting underinvestment does not capture some rich dynamic interaction present in many trading relationships. For instance, the standard two-stage model assumes that the trading partners invest only once and that bargaining can only begin after all investments are completed. In practice, however, the timing of investment and bargaining is — at least to some extent — chosen endogenously by the parties, and the investment and bargaining stages are often intertwined. In particular, when one makes a specific investment targeted at a particular partner, it is plausible for him to approach his partner to negotiate trade terms, before his investment is completed.

In the current paper, we develop a model that introduces dynamic interaction between investment and bargaining, by allowing the parties to continue to invest until they agree on how to divide the trading surplus. Specifically, our extensive form has the following structure. In each period, both parties choose how much to invest, and then a (randomly chosen) party offers the terms of trade. If that offer is accepted, then trade occurs according to the agreed terms and the game ends. If that offer is rejected, however, the game moves on to the next period

[1] For this reason, we do not ask why the parties do not contract prior to investments and what contract would be most effective at alleviating the inefficiencies — the central questions raised in the recent debates on the incomplete contracts paradigm (see Tirole (1999) for a survey). These questions would be relevant only after finding an affirmative answer to our question: Do inefficiencies necessarily arise in the absence of any contracts? See our remarks on the incomplete contracts literature later in this section and Section 6.

without trade, and the parties can further invest to add to the existing stock of investments, which is again followed by a round of bargaining of the same form, and the same process is repeated until there is agreement. Except for the investment dynamics, our model retains the essential features of the static model of holdup: we assume that no ex ante contracts exist and that the trade partners invest before they begin negotiating the terms of trade and complete their investments before they agree on trade.

We obtain two novel results. First, despite the holdup problem, investments can be efficient when the parties discount the future very little (or when the interval between periods shrinks to zero). This asymptotic efficiency result follows from the dynamic nature of our model. On the equilibrium path, the parties split their joint surplus in proportion to their relative bargaining power, so holdup arises just as in the static model. But holdup need not entail poor incentives for investment since the incentives also depend on how they will split the surplus if a party deviates from his equilibrium level of investment. The deviator's payoff in turn depends on what he plans to do with his investment should they fail to agree in the current period. Suppose that the deviator is expected to make up for the shortfall by investing up to the original equilibrium level in the next period. Then, his continuation value will include the (discounted) cost of this additional investment. Consequently, when it is his partner's turn to offer, she will rationally subtract from her standard offer (discounted value of his share of the future pie) the investment cost he will save if they agree in the current period. This feature generates an additional penalty against underinvestment (than the standard sharing of the pie would), thus generating a stronger incentive for investment than in the static game. This improved incentive for investment in turn makes credible the original belief that the investor will make up for the shortfall in the future if he deviates. In fact, if the parties discount the future very little, each party will have almost no incentive to deviate from any investment level below the first-best one since a dollar underinvestment results in an offer from his partner (when he makes an offer) that is almost a dollar lower.[2]

Second, we identify circumstances under which holdup entails underinvestment unambiguously. In our dynamic model, inefficiencies may arise from the imbalance between a party's bargaining power and the (relative) social value of his investment. As mentioned above, on the equilibrium path the parties split the joint surplus according to their bargaining power.

[2] When the party in question makes an offer, he internalizes the marginal social return by virtue of being a residual claimant in that case.

Hence, a party cannot be induced to make an investment that costs more than he expects to receive from bargaining. A socially desirable investment will thus not be made if the investor does not have a strong enough bargaining power to recoup his investment cost.[3] In this case, the investment would be inefficient, but the reason for the inefficiency is different from the one recognized by the standard model, as will become clear through our analysis.

Our model is related to three branches of literature. First, the current paper sheds some new light on the incomplete contracts literature. This literature, largely taking the underinvestment result as given, has focused on how contractual and organizational safeguards can mitigate the holdup problem: Some authors have proposed vertical integration as a solution (Klein, Crawford and Alchian, 1978; and Williamson, 1979), while others have proposed allocating asset ownership (Grossman and Hart, 1986; and Hart and Moore, 1990), financial control rights (Aghion and Bolton, 1992; and Dewatripont and Tirole, 1994), hierarchical authority (Aghion and Tirole, 1997), to name a few. This literature has been criticized for failing to consider the full set of feasible arrangements. While the empirical relevance of contractual incompleteness is widely acknowledged, there is a sense in which the incomplete contracts literature has not provided sufficient justifications for the predicted forms of contractual incompleteness (see Tirole (1999) for the recent survey). The current paper can be seen as reconciling this apparent tension, since our main result would imply that the underlying incentive problem may not be too worrisome even in the absence of any ex ante contracts.[4]

Second, our model is closely related to non-cooperative bargaining models. In particular, it can be seen as an extension of Rubinstein's (1982) model to allow for the pie to be endogenously determined through investments. Given a reasonable restriction on Markov Perfect equilibria, our model predicts no delay in agreement, as in the Rubinstein model. His uniqueness result does not carry over to the current model, though, since different future investment plans generate different incentives for investments, which in turn makes those plans sustainable. Busch and Wen (1995) obtain multiple Perfect equilibria in a bargaining model in which negotiators play a normal-form game repeatedly whenever they disagree. Their disagreement game does not affect the pie, so inefficiency arises only through delay. Further, their multiple

[3]This is also true in the static model. This issue never arises there, however, since any static equilibrium trivially satisfies this feature. In our dynamic model, a candidate equilibrium may sustain a much higher investment than in the static equilibrium, so the issue of individual rationality becomes relevant.

[4]Section 6 contains more detailed discussion on the implications of our results for the incomplete contracts view on organization.

equilibria correspond to different bargaining shares, which are supported by threats contingent on payoff irrelevant histories. By contrast, the inefficiency and multiplicity of equilibria in our model follows from the endogeneity of the bargaining stake. Also, threats contingent on payoff irrelevant histories play no role in our model, since we focus on Markov Perfect equilibria.

Last but not least, our model is related to the literature on "contribution games" which studies the incentives for voluntary contribution to public projects (see Marx and Matthews (2000), Gale (2000), Lockwood and Thomas (2001), Admati and Perry (1991) and Bagnoli and Lipman (1989)).[5] Marx and Matthews (2000) and Lockwood and Thomas (2001) show that, if contributors are allowed to contribute over time, the standard free-riding problem can be almost overcome by gradual accumulation strategies and the accompanying dynamic threat. The holdup problem is similar to the free-riding problem arising in the contribution games, but our model differs crucially on several accounts. First, the parties explicitly bargain to split the surplus in our model, rather than following some exogenous sharing rule (implied by the public good technology). This extra strategic interaction is largely responsible for the efficiency result we obtain. Second, there is a difference in the way surplus is realized. In our model, surplus can arise *only once when the parties trade*, even though the *level* of surplus realizable from trade increases continuously with investments. Hence, future accumulation of investments can be achieved only by foregoing current trade, i.e., by postponing surplus realization. By contrast, the contribution models assume that the *timing* of surplus realization as well as the *level* of surplus depends completely on the accumulated investments. Hence, a future accumulation does not require the postponement of surplus realization.[6] This difference in environment implies that the gradual accumulation strategies proposed by the contribution game literature would be unsustainable in our game.[7] In fact, investments take place all at

[5]Gale (2000) studies monotone games with positive spill-over, while Lockwood and Thomas (2001) consider repeated games with irreversible action variables. Both include contribution games as a special case.

[6]This remark applies even in the so-called "payoff jump" case considered by Marx and Matthews (2000), in which contributors realize no flow surplus until reaching a certain accumulation target. Clearly, there is no current surplus to be sacrificed to enable future accumulation, prior to reaching the target.

[7]Suppose that the parties split the surplus according to some *exogenous* sharing rule (rather than through bargaining), but that the surplus is realized only when they agree to trade (hence keeping our second feature). Since the *realizable* surplus increases continuously with investments in our model, the gradual investment strategies would involve ever-shrinking investment increments toward the accumulation target (see Marx and Matthews (2000) and Lockwood and Thomas (2001)). This latter feature means that the additional surplus that would be obtained by delaying trade becomes arbitrarily small as the target is approached, relative to the

once in our (selected) equilibria. Finally, the contribution models consider investments that are perfectly substitutable or symmetric in their effect on the surplus. We consider a wide range of cases in which parties' investments are imperfectly substitutable or even complementary, and more importantly, have asymmetric effects on the surplus, including the extreme case in which only one party invests. Our results hold regardless of the underlying technologies.

The rest of the paper is organized as follows. The next section introduces the model and the solution concept. Section 3 illustrates the main ideas through an example. The general model is then analyzed in Section 4 (and Section 5 which relaxes our refinement). Section 6 concludes.

2 The model

2.1 The description of the dynamic holdup game

Two risk-neutral parties, a buyer and a seller, make sunk investments for the purpose of a potential trade of a good. Time flows in discrete periods of equal length, $t = 0, 1, 2, ...$, and the players discount future utility by the common per-period discount factor, $\delta \in [0, 1)$. Trade can occur in any period and the parties can invest in any period up to (i.e., including) the period of trade. The parties can add to the existing stock of investments but they cannot disinvest. Investments are measured by the costs incurred, and the costs are incurred at the time of investments. A trade that occurs in period t after the cumulative investments of $b \geq 0$ by the buyer and of $s \geq 0$ by the seller create a joint surplus of $\phi(b, s)$ for the parties in that period if they trade (which amounts to $\delta^t \phi(b, s)$ in period 0 terms). No trade yields no surplus.

Trade surplus, ϕ, can be interpreted in a number of ways, including the following:

- **Bilateral trade:** Trade may require costly production on the part of the seller. Suppose that the seller can deliver $q \in Q \subset \mathcal{R}_+$ units of a good to the buyer at the cost of $c(q, b, s)$, yielding the gross surplus of $v(q, b, s)$ for the buyer. Assuming that $v(0, b, s) = c(0, b, s) = 0$, we can take $\phi(b, s) := \max_{q \in Q}\{v(q, b, s) - c(q, b, s)\}$ as the maximized joint surplus (that realizes as a result of an efficient quantity choice), with the standard assumption

cost of postponing surplus realization, so future accumulation (which would require postponing trade) becomes no longer credible, thus unraveling the gradual accumulation strategies. Gradual accumulation is sustainable in the contribution models since it does not require postponing surplus realization, regardless of whether payoffs jump at the target or not.

that quantity is contractible. Such a specification has been frequently adopted in many incomplete contract articles.[8]

- **Partnership/Team Production:** The parties may be partners in a joint project. They initially make investments (b, s), which could include developing business plans, purchasing buildings and equipment, hiring employees. They then bargain to split the joint surplus, $\phi(b, s)$, resulting from their joint project. Unlike the trade example, partners' individual revenue contributions may be difficult to identify, but, as in the trade example, their project requires a joint decision.

To capture the idea that the parties can invest until they conclude the negotiation, we adopt the following extensive form. Each period is divided into two stages: investment and bargaining. In the investment stage, the parties simultaneously choose (incremental) amounts to invest. Once the investments are sunk, they become public. In the bargaining stage, a party is chosen randomly to offer to his partner a share of the surplus that would result from trade at that point. We assume that the buyer is chosen with probability $\alpha \in [0, 1]$, and the seller is chosen with the remaining probability.[9] If the offer is accepted, then trade takes place, the surplus is split according to the agreed-upon shares between the two parties, and the game ends. If the offer is rejected, then the game moves on to the next period without trade, and the same process is repeated; i.e., the players can make incremental investments, which is followed by a new bargaining round with a random proposer. Note that, if the game ends after the first period (or equivalently $\delta = 0$), our model will coincide with the standard static model. For future reference, this one-period truncation of our game will be referred to as the *static holdup game.*

We make several assumptions about the social surplus function, ϕ. First, we assume that $\phi(\cdot, \cdot)$ is strictly positive, twice continuously differentiable, strictly increasing and strictly con-

[8]Note also that this specification allows the investments to be both cooperative and selfish. (A selfish investment directly benefits the investor while a cooperative investment directly benefits the trading partner of the investor. See Che and Hausch (1999).) Che and Hausch (1999) showed that cooperative investments limit the ability of contracting to solve the holdup problem. Our results below do not depend on the nature of investments.

[9]This part of the game represents a simple modification of the Rubinstein game, suggested by Binmore (1987). This model separates the issue of relative bargaining power from the discount factor and eliminates the (arbitrary) bias associated with who becomes the first proposer. The subsequent results will remain qualitatively the same, in particular when the parties discount very little, if one adopts the Rubinstein model.

cave,[10] and $\lim_{b\to\infty} \frac{\phi(b,\cdot)}{b} = 0$ and $\lim_{s\to\infty} \frac{\phi(\cdot,s)}{s} = 0$. These assumptions ensure existence of a unique pair of efficient investments. Further, we require that either $\phi_{bs}(b,s) \leq 0$ for all $(b,s) \geq (0,0)$ or $\phi_{bs}(b,s) > 0$ for all $(b,s) \geq (0,0)$. This last assumption means that investments are either complements or substitutes, globally. It simplifies the subsequent analyses and the interpretation of our results.

We now establish some benchmark outcomes. The first-best investment pair is defined as

$$(b^*, s^*) = \arg\max_{b,s}\{\phi(b,s) - b - s\},$$

which is unique due to the strict concavity of ϕ. We assume that $(b^*, s^*) > (0,0)$.[11] It is useful to consider the following (hypothetical) best response functions for the buyer[12]

$$B_\delta(s) = \arg\max_{b}\{\alpha\phi(b,s) - (1-(1-\alpha)\delta)b\}$$

and for the seller

$$S_\delta(b) = \arg\max_{s}\{(1-\alpha)\phi(b,s) - (1-\alpha\delta)s\}.$$

Observe that $B_1(\cdot)$ and $S_1(\cdot)$ are the socially efficient responses. Thus, (b^*, s^*) satisfies $b^* = B_1(s^*)$ and $s^* = S_1(b^*)$. Consider the other extreme case, with $\delta = 0$: $B_0(\cdot)$ and $S_0(\cdot)$ represent the best response functions of the buyer and the seller, respectively, in the static holdup game.

Notice that the above objective functions exhibit strict increasing differences in $(b;\delta)$ and in $(s;\delta)$, respectively, for $\alpha \in (0,1)$. Hence, $B_0(s) \leq B_\delta(s) \leq B_1(s)$ for any $s \geq 0$ and $\delta \in (0,1)$, and $S_0(b) \leq S_\delta(b) \leq S_1(b)$ for all $b \geq 0$ and $\delta \in (0,1)$. Meanwhile, strict concavity of $\phi(\cdot,\cdot)$ implies that

$$\phi_{bb}\phi_{ss} > \phi_{bs}^2 \Leftrightarrow B'_\delta S'_\delta < 1,$$

from which it follows that, for any $\delta \in (0,1)$, $B_\delta(\cdot)$ intersects $S_\delta(\cdot)$ only once. Let (b_δ, s_δ) denote such an intersection (that is, $b_\delta = B_\delta(s_\delta)$ and $s_\delta = S_\delta(b_\delta)$). Again $(b_1, s_1) = (b^*, s^*)$, while (b_0, s_0) characterize the subgame perfect equilibrium investments of the static holdup game.

[10] The strict concavity assumption rules out the case of perfect substitutable investments (i.e., $\phi(b,s) = \phi(b+s)$), which is plausible in many public good provision problem (see Marx and Matthews (2000), for instance). While we assume strict concavity for ease of exposition, all subsequent results hold for the case of perfectly substitutable investments. See Remark 1.

[11] Allowing one component of the first-best pair to be zero incorporates the possibility that only one party invests.

[12] The best response functions are well-defined due to the strict concavity of ϕ.

2.2 The equilibrium conditions

In our model, the stake of bargaining at any given time depends only on the stock of investments accumulated prior to that time. Consequently, the stock of aggregate investments constitutes the only payoff-relevant part of the history. A *Markov strategy* profile specifies for each period (incremental) investment choices by the two parties as functions of the pair of total investments accumulated up to the *last period*, and, for the bargaining stage, a price offer and a response rule for each party — a function mapping from an offer to {accept, reject} — both as functions of the *current-period* stock of total investments. We look for a subgame perfect equilibrium in Markov strategies, called *Markov Perfect equilibrium.*

Formally, we can characterize the equilibrium by a pair of *value functions*, $\langle \beta, \sigma \rangle : \mathcal{R}^2_+ \to \mathcal{R}^2$, that map from the *current period* stock of investments into the (continuation) values for the buyer and the seller at that period (before a proposer is randomly chosen), and by a *policy rule*, $E \equiv \langle x, y \rangle : \mathcal{R}^2_+ \to \mathcal{R}^2_+$, that maps the *last*-period stock of investments into the current-period stock of investments by the buyer and the seller.

The necessary and sufficient conditions for a Markov Perfect equilibrium are that, for each $(b, s) \in \mathcal{R}^2_+$, and for $(b', s') = E(b, s)$,

$$\begin{aligned}\beta(b,s) = \alpha \max\{\phi(b,s) - \delta[\sigma(b',s') - (s'-s)], \delta[\beta(b',s') - (b'-b)]\} \\ +(1-\alpha)\delta[\beta(b',s') - (b'-b)],\end{aligned} \tag{1}$$

$$\begin{aligned}\sigma(b,s) = (1-\alpha) \max\{\phi(b,s) - \delta[\beta(b',s') - (b'-b)], \delta[\sigma(b',s') - (s'-s)]\} \\ +\alpha\delta[\sigma(b',s') - (s'-s)],\end{aligned} \tag{2}$$

$$b' \in \arg\max_{\tilde{b} \geq b} \{\beta(\tilde{b}, s') - (\tilde{b} - b)\}, \tag{3}$$

$$s' \in \arg\max_{\tilde{s} \geq s} \{\sigma(b', \tilde{s}) - (\tilde{s} - s)\}. \tag{4}$$

The buyer's value function is explained as follows. Given an existing stock of (b, s), a buyer is chosen with probability α to make an offer. He can either make an offer which is unacceptable to the seller, in which case he obtains his net discounted continuation value, $\delta[\beta(b', s') - (b' - b)]$, or he can make the lowest offer acceptable to the seller, which leaves the seller with her discounted continuation value, $\delta[\sigma(b', s') - (s' - s)]$, whichever is more profitable. With probability $1 - \alpha$, the buyer becomes a responder. Based on the above reasoning, he will earn his discounted continuation value, $\delta[\beta(b', s') - (b' - b)]$, no matter how the seller resolves her trade-off. The seller's value function has a similar explanation. The next two equations, (3) and (4), state that the players choose their incremental investment pair as mutual best responses, taking into account the future evolution of the game. Clearly, these conditions are necessary for Markov Perfection. They are also sufficient, since there is no profitable (possibly non-Markovian) deviation, given the single-period deviation principle.

As will be seen, there are many Markov Perfect equilibria. Throughout, we will focus on Markov Perfect equilibria that satisfy the following properties:

Refinement (P): *(i) If* $E(b,s) \geq (\tilde{b},\tilde{s}) \geq (b,s)$ *then* $E(b,s) = E(\tilde{b},\tilde{s})$*, and (ii) if* $E(b,s) = (b,s)$*, then* $x(b+\epsilon, s) = b+\epsilon$ *and* $y(b, s+\epsilon) = s+\epsilon$*, for all* $\epsilon > 0$.

The first property requires an equilibrium policy rule to select investment reactions consistently. To see this, let $(b', s') = E(b, s)$. By (3) and (4), (b', s') is a mutual best response when they start with a stock, (b, s). Then, (b', s') is also a mutual best response if they start with a stock of $(\tilde{b}, \tilde{s})$, $(b, s) \leq (\tilde{b}, \tilde{s}) \leq (b', s')$. If this is the only mutual best response pair, then clearly $(b', s') = E(\tilde{b}, \tilde{s})$, so this refinement would impose no restriction. **(P)**-(i) imposes a restriction, though, when there are multiple best responses, in which case it requires selecting the same pair, $(b', s') = E(\tilde{b}, \tilde{s})$. The second property says that if both parties stop investing further starting from a certain stock, then each of them should stop investing further, when starting from a stock containing a higher sunk investment of his own and the same sunk investment of his opponent. Combined together, these properties impose a *passivity* restriction on the players' out-of-the equilibrium beliefs: any deviation by an investor from an equilibrium plan triggers the least amount of revisions on the beliefs about what that investor will do next period if no agreement is reached in the current period. It is worth emphasizing that this refinement can only limit the set of equilibria. In other words, the refinement plays no role in sustaining an efficient investment pair. Further, as will be seen, these properties preserve the flavor of the static holdup game.

We call a Markov perfect equilibrium satisfying **(P)** a *regular* equilibrium, or simply an equilibrium. Section 5 will explore the consequences of relaxing **(P)**, which includes a delay in agreement and an inefficiently low investment — even lower than the underinvestment that would arise in the static holdup game.

3 A Motivating Example

Suppose that only the seller invests and her investment decision is binary: "invest" and "not invest." Much of the insight for our results can be gained in this simplified model, which also serves to highlight that our main results hold even when the investments are discrete. Assume that total surplus is $\phi = 100$ if the seller invests and it is $\phi = 40$ if she does not invest and that the cost of investment for the seller is $C \in (30, 50]$. Hence, the investment is socially desirable. Assume that $\alpha = \frac{1}{2}$. The following table describes the payoff the seller receives from bargaining in the static holdup game and in (an equilibrium) of our dynamic model.

Investment Decision	Total surplus	S's gross payoff (**Static model**)	S's gross payoff (**Dynamic model**)
Invest	100	50	50
Not invest	40	20	$\max\{20 - \delta\frac{C}{2}, \delta(50 - C)\}$
Difference	60	30	$\geq C$ if $\delta \geq \delta^*(C) := 2 - \frac{60}{C}$

As described in the third column, in the static holdup game, the seller receives one half of the total surplus, whether she invests or not. Hence, she enjoys the marginal return of $30 = 50 - 20$ from the investment. Since this return falls short of the investment cost of $C > 30$, the seller will never invest in equilibrium. This describes the classic "underinvestment" problem arising from holdup.

Now consider our dynamic model. Suppose first that the seller invests. Then, the realizable total surplus is fixed at 100 for the rest of the game, since no higher investment is feasible and disinvestment is not allowed. The ensuing subgame then simply coincides with the standard (random-proposer) Rubinstein game with a fixed pie. Hence, along this path, the seller's payoff (gross of the investment cost) is one half of the total surplus, 50. Notice that the seller is held up even in the dynamic model, in terms of her *absolute* payoff.

The holdup need not imply that the seller will choose "not invest" in our model, however. The seller's incentive also depends on the payoff she receives when she chooses "not invest." That payoff, in turn, depends on the belief held by the players regarding whether the seller will invest in the next period if no agreement is reached in the current period. Suppose that the seller is expected to *invest* in the next period. This means that the seller will incur the cost of C and split 100 equally (as argued above), so her (discounted) continuation value will be $\delta(50-C)$. Likewise, the buyer's continuation value following "not invest" and rejection will be 50δ. With probability $\frac{1}{2}$, the buyer makes an offer. In this event, he offers either the seller's continuation value of $\delta(50-C)$ or an amount that will lead to rejection, whichever is more profitable. The resulting payoff to the seller is $\delta(50-C)$ either way. Likewise, if the seller makes an offer (which occurs with probability $\frac{1}{2}$), she offers 50δ or any rejectable amount, which yields the payoff of $\max\{40-50\delta, \delta(50-C)\}$ for the seller.

Combining the terms, the expected (gross) payoff for the seller is $\max\{20-\delta\frac{C}{2}, \delta(50-C)\}$ following "not invest," *given the belief that she will invest next period.* Hence, the seller's marginal return from investment is $50-\max\{20-\delta\frac{C}{2}, \delta(50-C)\} = \min\{30+\delta\frac{C}{2}, (1-\delta)50+\delta C\}$, which exceeds the cost of C if $\delta \geq \delta^*(C) = 2 - \frac{60}{C}$. (Note $\delta^*(C) < 1$ for any C in the relevant range.) If $\delta \geq \delta^*(C)$, the seller will indeed have an incentive to invest given the aforementioned belief, which in turn validates the belief itself. In sum, "invest" can be sustained as an equilibrium in our dynamic model if $\delta \geq \delta^*(C)$. In this equilibrium, the seller is held up when she invests, but she is held up even more severely when she does not invest, due to the unfavorable, but credible, belief. For instance, $\delta^*(C) = \frac{1}{2}$ if $C = 40$. In that case, investment would not occur in the static holdup game, but it is sustainable for any $\delta \geq \frac{1}{2}$.

Several remarks are worth making. First, "not invest" is also supported as an equilibrium, by the belief that the seller will never invest in the future if they disagree. Our claim is thus not that holdup will *never* entail inefficiency, but rather that it *need not*. Second, the example shows that efficiency is achieved even when the investment decision is binary and it is one sided. This feature, as mentioned above, contrasts with the efficiency result obtained in Marx and Matthews, which requires the presence of multiple investors with more than binary investment choices. Finally, the example can be modified to illustrate the other main theme of the current paper: Holdup can unambiguously lead to inefficiency when there is imbalance between a party's bargaining power and his investment cost. Suppose the investment cost is $C \in (50, 60)$. In this case, the seller's investment is socially desirable, yet it cannot be sustained in our dynamic model, for any $\delta \in [0, 1)$.[13] This example shows that the investor's individual rationality, i.e., whether her equilibrium payoff at least covers her investment cost, plays an important role for the efficiency consequence of holdup.

[13] The seller would earn $50 - C < 0$ if she invests, whereas if she does not invest, she can guarantee herself at least zero payoff by repeatedly rejecting any offer. This point is also seen by a profitable one-period deviation: The seller would earn from deviate no less than $\delta(50-C) > 50-C$.

4 The Analysis of the Dynamic Holdup Game

We now turn to our general model. We characterize the set of regular equilibria, beginning with the necessary conditions.

4.1 Necessary Conditions

Let $\langle \beta, \sigma, E \rangle$ be a (regular) equilibrium. A direct consequence of **(P)**-(i) is that $E(b, s)$ is a fixed point of $E(\cdot, \cdot)$: $E\left(E(b, s)\right) = E(b, s)$. This means that, on a (regular) equilibrium path, the parties invest only in the first period, if at all. Delays are inefficient when no further investments are expected, so trade will occur immediately, as is shown below.

Lemma 1 *(No Delay) Starting from any stock, trade must occur immediately (i.e., in the same period) on the equilibrium path.*

Proof. Let (b, s) be the stock by the end of period t and let $(b', s') = E(b, s)$. The fixed point property implies that $E(b', s') = (b', s')$. Suppose, contrary to the claim, that no trade occurs in period $t + 1$. Then,

$$\phi(b', s') \leq \delta[\beta(b', s') + \sigma(b', s')]. \tag{5}$$

Hence, it follows from (1) and (2) that

$$\beta(b', s') = \delta\beta(b', s') \text{ and } \sigma(b', s') = \delta\sigma(b', s'),$$

which implies that

$$\beta(b', s') = \sigma(b', s') = 0. \tag{6}$$

However, (6) contradicts (5) since $\phi(\cdot, \cdot) > 0$. Consequently, trade must occur in period $t + 1$ in equilibrium. Q.E.D.

Note that delay can still occur off the equilibrium path if further investments can raise ϕ sufficiently to cover both the additional investment expenses and the delay costs (i.e., discounting of future surplus). We next consider further necessary conditions.

Suppose that $E(0, 0) = (\hat{b}, \hat{s})$, i.e., $(\hat{b}, \hat{s})$ is an equilibrium pair, starting with a stock of $(0, 0)$. Since $E(\hat{b}, \hat{s}) = (\hat{b}, \hat{s})$ and trade occurs at $(\hat{b}, \hat{s})$ by Lemma 1, the game at that point reduces to the standard (random-proposer) bargaining game with an exogenously given pie equal to $\phi(\hat{b},\hat{s})$. The latter game yields the unique (subgame-perfect) equilibrium payoffs of

$$\beta(\hat{b},\hat{s}) = \alpha\phi(\hat{b},\hat{s}) \text{ and } \sigma(\hat{b},\hat{s}) = (1-\alpha)\phi(\hat{b},\hat{s}) \tag{7}$$

for the buyer and the seller, respectively. (These payoffs can also be obtained directly, by substituting $E(\hat{b},\hat{s}) = (\hat{b},\hat{s})$ into (1) and (2) using Lemma 1.) Note that these are precisely the payoffs that the parties would receive in the static holdup game, if they had invested $(\hat{b},\hat{s})$ in equilibrium. As we will show below, this resemblance to the static holdup outcome does not imply that the parties will have the same incentives to invest in the dynamic holdup game, since the bargaining terms need not satisfy (7) off the equilibrium path, i.e., when a party deviates from his equilibrium investment level.

Since each party has the option of making no investment and ensuring himself at least zero payoff, we must have that $\alpha\phi(\hat{b},\hat{s}) - \hat{b} \geq 0$ and $(1-\alpha)\phi(\hat{b},\hat{s}) - \hat{s} \geq 0$. By the strict concavity of ϕ, this implies that

$$\hat{b} \in [0, \underline{B}(\hat{s})] \text{ and } \hat{s} \in [0, \underline{S}(\hat{b})] \tag{8}$$

where $\alpha\phi(\underline{B}(s), s) - \underline{B}(s) \equiv 0$ and $(1-\alpha)\phi(b, \underline{S}(b)) - \underline{S}(b) \equiv 0$. We refer to these two conditions as *individual rationality constraints.* The set of pairs satisfying these constraints is denoted by $IR(\alpha)$, and is graphically described as an area upper-bounded by $\underline{B}(\cdot)$ and $\underline{S}(\cdot)$ in Figures 1 and 2 for the case in which the buyer's bargaining power is large and for the case it is small (α close to 0), respectively.[14]

[Insert Figures 1 and 2 about here.]

Note that $\underline{B}(\cdot)$ and $\underline{S}(\cdot)$ are nondecreasing and that $IR(\alpha)$ varies with α. As the bargaining power of the buyer (α) decreases, the condition is easier to satisfy for the seller but more difficult to satisfy for the buyer. For instance, the investment pair (b_δ, s_δ) satisfies the buyer's individual rationality in Figure 1 but not in Figure 2. In particular, as α tends to 0, no positive investment by the buyer can be sustained.

Next, we derive some incentive constraints. Again, let $E(0,0) = (\hat{b},\hat{s})$ and assume $\hat{b} > 0$. Suppose that the buyer deviates unilaterally by investing less than the target, say $b < \hat{b}$ (while the seller chooses $\hat{s}$). Given **(P)**-(i), $E(b,\hat{s}) = (\hat{b},\hat{s})$. That is, following the deviation, the seller expects that the buyer would invest the remainder $(\hat{b} - b)$ in the next period if no agreement were reached in the current period. If b is sufficiently close to $\hat{b}$, then trade will

[14] Note that both figures assume $\phi_{bs} \equiv 0$.

occur immediately since the delay costs exceed the gains from further investments. Using $E(E(b,\hat{s})) = E(b,\hat{s}) = (\hat{b},\hat{s})$ (which again follows from **(P)**-(i)), we can compute the buyer's payoff following the deviation to $b < \hat{b}$ (for b sufficiently close to $\hat{b}$):

$$\begin{aligned}\beta(b,\hat{s}) &= \alpha[\phi(b,\hat{s}) - \delta(1-\alpha)\phi(\hat{b},\hat{s})] + (1-\alpha)\delta[\alpha\phi(\hat{b},\hat{s}) - (\hat{b}-b)] \\ &= \alpha\phi(b,\hat{s}) - (1-\alpha)\delta(\hat{b}-b).\end{aligned}$$

The first term in the right-hand side represents a share of the (reduced) pie resulting from the deviation (i.e., underinvestment relative to $\hat{b}$). This term appears in the static holdup game as well, and it constitutes the only incentive for the buyer's investment in that game. The second term is new in the dynamic game and it reflects the buyer's weakened bargaining position resulting from the equilibrium belief that he will make up for the shortfall of $(\hat{b}-b)$ next period, in the event of disagreement. The presence of this second term will be seen later as sustaining a higher investment level than in the static holdup game.

Since the deviation must be unprofitable, $\beta(\hat{b},\hat{s}) - \hat{b} - [\beta(b,\hat{s}) - b] \geq 0$ for $b < \hat{b}$. Dividing both sides by $\hat{b} - b$ and letting b approach $\hat{b}$ from below, we obtain the buyer's *incentive constraint*:

$$\alpha\phi_b(\hat{b},\hat{s}) - [1-(1-\alpha)\delta] \geq 0 \quad \Longleftrightarrow \quad \hat{b} \leq B_\delta(\hat{s}). \tag{9}$$

Applying the same argument for the seller, we obtain:

$$(1-\alpha)\phi_s(\hat{b},\hat{s}) - [1-\alpha\delta] \geq 0 \quad \Longleftrightarrow \quad \hat{s} \leq S_\delta(\hat{b}). \tag{10}$$

The incentive constraints, (9) and (10), impose a further upper bound on the total investments sustainable in equilibrium, in addition to individual rationality, (8). As will be seen in the next proposition, **(P)**-(ii) serves to further narrow down the set of regular equilibria, in this case providing lower bounds on investments. Combining all the observations, the following proposition shows that any regular equilibrium investment pair must be in the set:

$$\Omega := \{(b,s) \in \mathcal{R}^2_+ | b \in [B_0(s), \min\{\underline{B}(s), B_\delta(s)\}] \text{ and } s \in [S_0(b), \min\{\underline{S}(b), S_\delta(b)\}]\},$$

which is graphed as shaded areas in Figures 1 and 2. (Ω depends on (α,δ) but this dependence is suppressed for simplicity.)

Proposition 1 *In (regular) equilibrium, $E(0,0) \in \Omega$.*

Proof. Fix a regular equilibrium and let $(\hat{b}, \hat{s}) = E(0, 0)$. We have already argued that $(\hat{b}, \hat{s})$ must lie weakly below $\min\{\underline{B}(\cdot), B_\delta(\cdot)\}$ and $\min\{\underline{S}(\cdot), S_\delta(\cdot)\}$.

Here we show that, in addition, $(\hat{b}, \hat{s})$ must lie weakly above $B_0(\cdot)$ and $S_0(\cdot)$. Suppose, to the contrary, that $\hat{b} < B_0(\hat{s})$. Since $(\hat{b}, \hat{s}) = E(\hat{b}, \hat{s})$, by **(P)**-(i), the equilibrium payoff for the buyer is $\beta(\hat{b}, \hat{s}) - \hat{b} = \alpha\phi(\hat{b}, \hat{s}) - \hat{b}$. Consider a unilateral deviation by the buyer to $b \in (\hat{b}, B_0(\hat{s})]$. By **(P)**-(ii), $x(b, \hat{s}) = b$, and by **(P)**-(i), $E(b, y(b, \hat{s})) = (b, y(b, \hat{s}))$. Using these two facts, the deviation payoff satisfies:

$$\begin{aligned} \beta(b, \hat{s}) - b &\geq \alpha\phi(b, \hat{s}) + \alpha\delta(y(b, \hat{s}) - \hat{s}) - b \\ &\geq \alpha\phi(b, \hat{s}) - b \\ &> \alpha\phi(\hat{b}, \hat{s}) - \hat{b} \\ &= \beta(\hat{b}, \hat{s}) - \hat{b}, \end{aligned}$$

where the first inequality holds since trading immediately after the deviation may not be optimal, the second holds since $y(b, \hat{s}) \geq \hat{s}$, and the last (strict) inequality holds since $b \in (\hat{b}, B_0(\hat{s})]$ and $\phi(\cdot, \hat{s})$ is strictly concave. Hence, the deviation is strictly profitable, which is a contradiction to the fact that $(\hat{b}, \hat{s}) = E(0, 0)$. Q.E.D.

Note that Ω is nonempty, since $B_0(\cdot) \leq \min\{\underline{B}(\cdot), B_\delta(\cdot)\}$ and $S_0(\cdot) \leq \min\{\underline{S}(\cdot), S_\delta(\cdot)\}$.

4.2 Sufficient Conditions

We now provide sufficient conditions for a (regular) equilibrium. In the case of (weakly) substitutable investments, the necessary conditions for a regular equilibrium characterized in Proposition 1 are also sufficient.

Proposition 2 *If the investments are (weak) substitutes, then any pair in Ω can be sustained as a (regular) equilibrium outcome. In such an equilibrium, the buyer and the seller invest $(\hat{b}, \hat{s}) \in \Omega$ in the first period and agree to trade immediately. Their equilibrium payoffs are $\alpha\phi(\hat{b}, \hat{s}) - \hat{b}$ and $(1 - \alpha)\phi(\hat{b}, \hat{s}) - \hat{s}$ for the buyer and the seller, respectively.*

Proof. See the Appendix.

To see the intuition behind this result, consider a special case where the investments are independent (i.e., $\phi_{bs}(\cdot, \cdot) = 0$). In this case, $B_\delta(\cdot) \equiv b_\delta$ and $S_\delta(\cdot) \equiv s_\delta$, i.e., the best response curves are perfectly inelastic. Fix any $(\hat{b}, \hat{s}) \in \Omega$. That pair, $(\hat{b}, \hat{s})$, can be sustained as an equilibrium by the policy rule: $E(b, s) = (\max\{b, \hat{b}\}, \max\{s, \hat{s}\})$. Given this rule, the buyer, for example, invests up to $\hat{b}$ when starting from a smaller level and stops when starting from a higher level. This is an equilibrium strategy since the buyer faces the payoff function graphed in Figure 3.

[Insert Figure 3 about here.]

Starting with any $b < \hat{b}$, the belief that he will make up the shortfall next period means that the buyer faces the marginal payoff described in (9):

$$\alpha\phi_b(b, \hat{s}) - [1 - (1 - \alpha)\delta],$$

which is strictly positive since $b < \hat{b} \leq b_\delta$. Starting with $b > \hat{b}$, the belief that he will stop investing further means that the buyer will face the marginal payoff:

$$\alpha\phi_b(b, \hat{s}) - 1,$$

which is strictly negative since $b > \hat{b} \geq b_0$. Since the same argument works for the seller, the policy rule, $E(\cdot, \cdot)$, is sustainable.

Unlike in the case of substitutability, if investments are complementary, then not all pairs in Ω can be sustained as a (regular) equilibrium. In particular, the static holdup outcome, (b_0, s_0), can never be sustained as a regular equilibrium with complementary investments, as shown below.

Proposition 3 *If investments are strictly complementary at* (b_0, s_0)*, then the static holdup outcome,* (b_0, s_0)*, cannot be sustained in any regular equilibrium.*

Proof. See the Appendix.

Note that, with complementarities, the joint surplus associated with (b_0, s_0) is the lowest among all pairs in Ω. Consequently, any regular equilibrium pair in our dynamic model must be more efficient than that outcome. We now study the sustainability of the most efficient pair in Ω. Let (b_{CE}, s_{CE}) denote such a pair. It follows from the strict concavity of ϕ and the complementarity of the investments that (b_{CE}, s_{CE}) is the largest vector in Ω. In particular, if $(b_\delta, s_\delta) \in IR(\alpha)$, then $(b_{CE}, s_{CE}) = (b_\delta, s_\delta)$. Next, we provide sufficient conditions for (b_{CE}, s_{CE}) to be sustained in equilibrium with complementary investments.

Proposition 4 *If investments are strictly complementary, it is an equilibrium for the parties to choose* (b_{CE}, s_{CE}) *and trade in the first period if*

$$\sup\left\{\left|\frac{\phi_{sb}}{\phi_{ss}}\right|, \left|\frac{\phi_{sb}}{\phi_{bb}}\right|\right\} < M \textit{ for some } M > 0;$$

i.e., the degree of complementarities is sufficiently small. In particular, if $(b_\delta, s_\delta) \in IR(\alpha)$, *then the condition reduces to*

$$\left|\frac{\phi_{sb}}{\phi_{ss}}\right| \le \frac{1-\alpha}{\alpha} \textit{ and } \left|\frac{\phi_{sb}}{\phi_{bb}}\right| \le \frac{\alpha}{1-\alpha}.$$

Proof. See the Appendix.

These conditions put a bound on the degree of complementarities. In particular, if (b_δ, s_δ) satisfies the individual rationality constraint, the bound has an intuitive expression. Strict concavity of ϕ implies that

$$\phi_{bb}\phi_{ss} > \phi_{bs}^2 \Leftrightarrow \left|\frac{\phi_{sb}}{\phi_{ss}}\right|\left|\frac{\phi_{sb}}{\phi_{bb}}\right| < 1,$$

so, if $|\frac{\phi_{sb}}{\phi_{ss}}| > \frac{1-\alpha}{\alpha}$, then $|\frac{\phi_{sb}}{\phi_{bb}}| < \frac{\alpha}{1-\alpha}$, implying that at least one of the two latter conditions must hold. In fact, if both parties have equal bargaining power and $\phi(b, s)$ is symmetric in (b, s), then all the sufficient conditions are satisfied.

Corollary 1 *If* $\alpha = \frac{1}{2}$ *and* $\phi(b, s)$ *is symmetric in* (b, s), *then it is an equilibrium for the parties to choose* (b_δ, s_δ) *and trade in the first period.*

Proof. Given $\alpha = \frac{1}{2}$ and the symmetry, $b_\delta = s_\delta$. Since $\phi(b_\delta, s_\delta) - b_\delta - s_\delta > 0$, this implies that $(b_\delta, s_\delta) \in IR(\alpha)$. Further, by concavity of ϕ, $\left|\frac{\phi_{sb}}{\phi_{ss}}\right|\left|\frac{\phi_{sb}}{\phi_{bb}}\right| \le 1$, which, together with the symmetry, implies that $\left|\frac{\phi_{sb}}{\phi_{ss}}\right| = \left|\frac{\phi_{sb}}{\phi_{bb}}\right| \le 1$. Thus, the claim follows directly from Proposition 4. Q.E.D.

We are now in a position to state the main result of this paper. Recall that $(b_\delta, s_\delta) \to (b^*, s^*)$ as $\delta \to 1$. Hence, the first-best investment pair can be sustained as a (regular) equilibrium in the limit as δ goes to 1, provided that (b^*, s^*) satisfies the individual rationality constraint (and the investments are not too complementary). This shows the importance of the individual rationality constraint for the prediction of inefficiencies. Let

$$\mathcal{A} \equiv \{\alpha \in [0, 1] | (b^*, s^*) \in IR(\alpha)\}$$

denote the set of α's at which (b^*, s^*) satisfies the individual rationality constraint. Since the joint net payoff is strictly positive at (b^*, s^*), individual rationality must hold for at least one party for any α, and it must hold for both for some values of α. Hence, $\mathcal{A}$ is nonempty. Further, it is convex (i.e., forms an interval), and if $\alpha \in \mathcal{A}$, then $(b_\delta, s_\delta) \in IR(\alpha)$ for all $\delta \in [0, 1)$. Most importantly, $\mathcal{A}$ does not include the extreme values of $\alpha = 0$ and $\alpha = 1$, if $(b^*, s^*) >> (0, 0)$. Intuitively, individual rationality fails when one party has a very weak bargaining power relative to the cost of his first-best investment decision;, i.e., when there is an imbalance between one's bargaining power and the social value of his investment. Our results are summarized in the following proposition:

Proposition 5 *If $\alpha \in \mathcal{A}$ and $-\frac{\phi_{sb}}{\phi_{ss}} \leq \frac{1-\alpha}{\alpha}$ and $-\frac{\phi_{sb}}{\phi_{bb}} \leq \frac{\alpha}{1-\alpha}$, then the first-best equilibrium pair can be sustained asymptotically as $\delta \to 1$. If $\alpha \notin \mathcal{A}$, then any regular equilibrium investment pair is bounded away from (b^*, s^*) for any $\delta \in [0, 1)$. In particular, any regular equilibrium pair (if it exists) approaches to the static one if either $\alpha \to 0$ or $\alpha \to 1$.*

Two implications can be drawn from this result. First, the holdup problem need not entail inefficiencies in our dynamic environment. The artificial separation between the investment and bargaining stages is what rules out efficiency in the static holdup game. At first glance, such an artificial separation appears only to serve as a modelling convenience. Yet, once we dispense with that feature and analyze holdup in a more realistic dynamic model, inefficiencies may vanish altogether. Second, the result sheds new light on the conditions needed for an unambiguous prediction of inefficiencies. In particular, it is the failure of the individual rationality constraint that causes all equilibria to be bounded away from efficiency. While this failure of individual rationality can be attributed to the imbalance between the investor's bargaining power and the social value of his investment, this insight for inefficiencies differs from the standard one offered by the static model. In particular, there need not be any connection between the extent to which a party is held up and the level of investment he makes in equilibrium. For instance, if a party's individual rationality constraint is almost binding at the first-best pair, he can still be induced to invest efficiently even though by doing so he will earn almost zero payoff (and will thus be much more severely held up than he would be in the static equilibrium).

Remark 1 *In public good provision, only the total contribution of the agents matters, so the investments are perfect substitutes. While our strict concavity assumption rules out this case,* *our result continues to hold even in that case. In fact, the result is strengthened: Asymptotic efficiency holds for any $\alpha \in [0,1]$ if the investments are perfect substitutes. To see this, suppose that $\phi(b,s) = \psi(b+s)$ for some strictly concave function $\psi(\cdot)$. Letting $z^* := \arg\max_{z\geq 0} \psi(z) - z$, any (b,s) such that $b+s = z^*$ constitutes the first-best outcome. Let Z^* denote the set of all such pairs. Clearly, for any $\alpha \in [0,1]$, there exists a first-best pair $(b,s) \in Z^*$ satisfying individual rationality. As before, Proposition 2 holds and implies that any pair in Ω is sustainable. It can be then easily checked that, any pair in $Z^* \cap IR(\alpha)$ is approached by a pair in Ω as $\delta \to 1$.*[15] *Intuitively, if one agent's investment is just as good as the other's, they can allocate the investment responsibilities to reflect their relative bargaining positions, i.e., by assigning a higher investment responsibility to the agent with a stronger bargaining power.*

Remark 2 *While investments and trade take place in the first period in our equilibria (thus as in static models), the sustainability of these equilibria rests on the infinite horizon. For instance, our asymptotic efficiency result depends on the out-of-equilibrium belief that any party who deviates to invest less than his target level will make up the short-fall in the next period. In a finite horizon model, such a belief will not be credible in the last period. Hence, the asymptotic efficiency result cannot be sustained in a finite-horizon model. Curiously, the same issue arises in the contribution games literature.*[16] *This feature implies that these types of results are most relevant in the circumstances where the time horizon is either infinite or uncertain.*

5 Nonregular equilibria

Our refinement, **(P)**, has so far limited the role of non-passive out-of-equilibrium beliefs. We show here that relaxing this refinement admits additional equilibria that involve delay and inefficiently low investments, even lower than the level predicted in the static holdup game.

5.1 Sub-static Investments

Relaxing **(P)**-(ii) admits a nonregular equilibrium in which a party invests strictly less than in the static holdup game (i.e., outside the set of potential regular equilibria identified in

[15] Note that the sets $\{(b,s)|b = B_\delta(s)\}$ and $\{(b,s)|s = S_\delta(b)\}$ both converge to Z^* as $\delta \to 1$.

[16] The asymptotic efficiency result of Marx and Matthews in the "no-payoff jump" case (which corresponds to the situation considered in our model) also unravels in the finite-horizon setting.

Proposition 1). Such an equilibrium can be sustained by the (out-of-equilibrium) belief that, following an upward deviation, the deviator will invest even more.[17]

This point can be illustrated in a model with one-sided investment. Suppose that only the buyer invests. Let $\phi(b)$ represent the surplus of trade following the total investment of b by the buyer and define b_δ and $\underline{B}$ (which is now a constant) as before. Now consider the set of possible investments:

$$\bar{\Omega} := [b_L, \min\{b_\delta, \underline{B}\}],$$

where $b_L := \min\{b \geq 0 \mid \alpha\phi(b) - b \geq \max\{\alpha\phi(b_\delta) - b_\delta, 0\}\}$. Clearly, $b_L < b_0$ whenever $b_0 > 0$ since $\alpha\phi(b_0) - b_0 = \max_b\{\alpha\phi(b) - b\} > \max\{\alpha\phi(b_\delta) - b_\delta, 0\}$. In short, the set $\bar{\Omega}$ contains investment levels lower than the static equilibrium level, b_0.

One can easily check that any investment level in this set, say $\hat{b} \in \bar{\Omega}$, can be sustained as a Markov Perfect equilibrium, with the policy rule:

$$x(b) = \begin{cases} \hat{b} & \text{if } b \leq \hat{b}, \\ \min\{b_\delta, \underline{B}\} & \text{if } b \in (\hat{b}, \min\{b_\delta, \underline{B}\}] \\ b & \text{if } b > \min\{b_\delta, \underline{B}\}. \end{cases}$$

Notice that any upward deviation is followed by a further jump to the highest sustainable level, $\min\{b_\delta, \underline{B}\}$, which punishes such a deviation more than would any response permitted by **(P)**-(ii).[18] Hence, a lower investment than the static equilibrium level can be sustained in a non-regular equilibrium. To be concrete, suppose $\phi(b) = 1 + 2\sqrt{b}$ and $\alpha = \delta = \frac{1}{2}$. Then, the set of regular equilibria is $\Omega = [\frac{1}{4}, \frac{4}{9}]$, while the set of Markov Perfect equilibria is $\bar{\Omega} = [\frac{1}{9}, \frac{4}{9}]$.

5.2 Delay

Relaxing **(P)**-(i) can give rise to a delayed agreement. The reason can be described as a coordination failure. When her partner plans to make a substantial investment in the future, a given party finds it optimal to delay trade until after that investment has been made, which makes it optimal to delay her own investment. Condition **(P)**-(i) rules out such a coordination failure.

[17] Recall that such a belief is not allowed by **(P)**-(ii), which restricts the deviator to stop investing further after an upward deviation.

[18] In the one-sided investment case, jumping to $\min\{b_\delta, \underline{B}\}$ constitutes the worst sustainable punishment, which implies that indeed $\bar{\Omega}$ is the full set of Markov Perfect equilibria.

Proposition 6 *Assume that investments are weak substitutes (i.e., $\phi_{bs} \leq 0$).*[19] *There exists $\hat{\delta} < 1$ such that for any $\delta > \hat{\delta}$, it is a Markov Perfect equilibrium for the parties to initially invest some sufficiently small amounts, $(\nu, \epsilon) << (b_0, s_0)$, disagree at the end of that period, and invest next period to (b_0, s_0) and trade.*

Proof. See the Appendix.

6 Conclusion

The current paper has reexamined an important foundation of modern organizational theory: the holdup problem. When parties negotiate ex post to determine the terms of trade, they split the ex post trading surplus, so a party can be seen as appropriating only a fraction, say a half, of the return to his sunk investment in terms of *absolute* payoff. It would then follow, according to the conventional wisdom, that the investor would appropriate only half of the *marginal* return to his investment, from which underinvestment would follow. We have shown that this link between *absolute* and *marginal* appropriability is an artifact of the rigid separation of the investment and bargaining stages assumed in the static model. Once we allow the two stages to be intertwined in a realistic fashion (by permitting the investments to continue until bargaining concludes), the fact that the parties split their trading surplus need not imply poor marginal incentives for their investment decisions. In particular, in the limit when the parties are extremely patient, the first-best investment decisions can be supported as an equilibrium as long as the parties recoup their investment costs from the negotiation.

Several implications can be drawn from these results. First, the holdup problem may not be as worrisome as may have been thought previously, even in the absence of ex ante contracts. In particular, our asymptotic efficiency result implies that agents may choose not to contract ex ante even in the presence of negligible transaction costs. This may explain why contracts are incomplete, in particular, why business transactions seldom rely on explicit contracts (see Macaulay (1963)).[20] Second, our theory suggests that, even when underinvestment occurs,

[19] A similar argument can be made for complementary investments, subject to the extra complication noted earlier.

[20] A similar result has been established by Che and Hausch (1999), Segal (1999) and Hart and Moore (1999). These papers show that contracts can do little to overcome the inefficiencies caused by the holdup problem. By contrast, our result rests on the finding that investment dynamics alone can solve the incentive problem, thus making ex ante contracts unnecessary. Remarkably, the features that make contracts ineffective in the

it may result from an investor not earning enough from bargaining to recoup the cost of his socially desirable investment. This suggests that “individual-rationality-type considerations” should be an important part of contract and organization design. In particular, inefficiencies may be eliminated simply by restoring one’s bargaining position to guarantee sufficiently high bargaining revenue for that party. The latter may be accomplished by simple instruments, such as (re)allocating decision rights and/or (re)drawing the boundary of agents’ authority within an organization, or by the courts’ assigning the legal default rights appropriately. This seems to suggest that our dynamic model can serve as a tighter, and thus more successful, foundation for understanding the observed nature of contractual incompleteness, than its extant counterpart.

aforementioned papers (e.g., “cooperative investments” and/or environmental “complexities”) do not disrupt efficiency here even in the absence of ex ante contracts!

7 Appendix A: Proofs

We first establish sufficient conditions under which a pair in Ω can be implemented as a regular equilibrium. For a pair, $(\hat{b}, \hat{s}) \in \Omega$, suppose that there exist functions $\hat{B}(\cdot) : [\hat{b}, \infty) \to \mathcal{R}_+$ and $\hat{S}(\cdot) : [\hat{s}, \infty) \to \mathcal{R}_+$ that satisfy the following conditions:

$(A1)$ $\quad \hat{B}(\hat{s}) = \hat{b}$ and $\hat{S}(\hat{b}) = \hat{s}$;

$(A2)$ $\quad \hat{B}(s) \in [B_0(s), \min\{\underline{B}(s), B_\delta(s)\}]$ for any $s \geq \hat{s}$ and $\hat{S}(b) \in [S_0(b), \min\{\underline{S}(b), S_\delta(b)\}]$ for all $b \geq \hat{b}$;

$(A3)$ $\quad \hat{B}(\cdot)$ and $\hat{S}(\cdot)$ are differentiable almost everywhere, and $\hat{B}'(\cdot)\phi_{bs}(\cdot) \geq 0$ and $\hat{S}'(\cdot)\phi_{bs}(\cdot) \geq 0$.

$(A4)$ $\quad$ If $\phi_{bs}(\cdot) \geq 0$ (i.e., complementary investments), then

$$\alpha\left[\phi_b(b, \hat{S}(b)) + \delta \hat{S}'(b)\right] - 1 \leq 0 \text{ and } \alpha\delta\left[\phi_b(b, \hat{S}(b)) + \phi_s(b, \hat{S}(b))\hat{S}'(b)\right] - 1 \leq 0 \text{ for } b \geq \hat{b},$$

and

$$(1-\alpha)[\phi_s(\hat{B}(s), s) + \delta \hat{B}'(s)] - 1 \leq 0 \text{ and } (1-\alpha)\delta[\phi_s(\hat{B}(s), s) + \phi_b(\hat{B}(s), s)\hat{B}'(s)] - 1 \leq 0 \text{ for } s \geq \hat{s}.$$

The first three conditions are trivial to satisfy by themselves for any $(\hat{b}, \hat{s}) \in \Omega$. The last condition imposes a nontrivial restriction on the degree of investment complementarities. Consider the following policy rule:

$$(A5) \qquad \hat{E}(b, s) = \begin{cases} (b, s) & \text{if } b \geq \hat{B}(s) \text{ and } s \geq \hat{S}(b), \\ (\hat{b}, \hat{s}) & \text{if } (b, s) \leq (\hat{b}, \hat{s}), \\ (b, \hat{S}(b)) & \text{if } b > \hat{b} \text{ and } s \leq \hat{S}(b), \\ (\hat{B}(s), s) & \text{if } b \leq \hat{B}(s) \text{ and } s > \hat{s}. \end{cases}$$

Lemma A1 *For any $(\hat{b}, \hat{s}) \in \Omega$, if there exist $\hat{B}(\cdot)$ and $\hat{S}(\cdot)$ satisfying $(A1) - (A4)$, then it is equilibrium behavior for the parties to invest according to the rule, $\hat{E}(\cdot, \cdot)$, and to trade immediately.*

Proof. We prove that a deviation is unprofitable in each case.

(i) $b \geq \hat{B}(s)$ and $s \geq \hat{S}(b)$: If both parties follow the investment rule, $\hat{E}(b, s) = (b, s)$. Substituting this into (1) gives the equilibrium payoff of the buyer (excluding sunk investment cost): $\alpha\phi(b, s)$. Suppose now that the buyer unilaterally deviates by choosing $\tilde{b} > b$. Applying the one-period deviation principle, trade must occur (by Lemma 1) and the resulting payoff for the buyer is $\alpha\phi(\tilde{b}, s) - (\tilde{b} - b)$. Such a deviation is unprofitable since $\alpha\phi(\tilde{b}, s) - \tilde{b} < \alpha\phi(b, s) - b$, for $\tilde{b} > b > \hat{B}(s) \geq B_0(s)$. A symmetric argument holds for the seller's deviation.

(ii) $(b, s) \leq (\hat{b}, \hat{s})$: Since $\hat{E}(\hat{b}, \hat{s}) = (\hat{b}, \hat{s})$, the equilibrium payoff for the buyer is $\beta(\hat{b}, \hat{s}) - \hat{b} = \alpha\phi(\hat{b}, \hat{s}) - \hat{b}$. Suppose now that the buyer deviates by investing just $\tilde{b} < \hat{b}$ (such a deviation is feasible only if $b < \hat{b}$). Suppose that trade occurs following such a deviation. Substituting $E^n(\tilde{b}, \hat{s}) = (\hat{b}, \hat{s})$, $n \geq 1$, into (1) then gives the buyer's deviation payoff:

$$\beta(\tilde{b}, \hat{s}) - (\tilde{b} - b) = \alpha\phi(\tilde{b}, \hat{s}) - (1 - \alpha)\delta(\hat{b} - \tilde{b}) - (\tilde{b} - b)$$

Such a deviation is unprofitable since the right-hand side is strictly increasing in $\tilde{b}$ (since $\tilde{b} < \hat{b} \leq B_\delta(\hat{s})$) and equals $\alpha\phi(\hat{b}, \hat{s}) - (\hat{b} - b)$ when $\tilde{b} = \hat{b}$. If trade does not occur following the deviation, then, since $\hat{E}^2(\tilde{b}, \hat{s}) = \hat{E}(\tilde{b}, \hat{s}) = (\hat{b}, \hat{s})$, the deviation payoff is simply

$$\begin{aligned}
\beta(\tilde{b}, \hat{s}) - (\tilde{b} - b) &= \delta(\alpha\phi(\hat{b}, \hat{s}) - (\hat{b} - \tilde{b})) - (\tilde{b} - b) \\
&\leq \alpha\phi(\hat{b}, \hat{s}) - (\hat{b} - b) \\
&= \beta(\hat{b}, \hat{s}) - (\hat{b} - b),
\end{aligned}$$

where the inequality follows from $(\hat{b}, \hat{s}) \in IR(\alpha)$ and $\delta < 1$. We again conclude that such a deviation is unprofitable.

Now consider a deviation to $\tilde{b} > \hat{b}$. Again, if trade occurs after the deviation, then the resulting payoff for the buyer is

$$\begin{aligned}
\beta(\tilde{b}, \hat{s}) - (\tilde{b} - b) &= \alpha\phi(\tilde{b}, \hat{s}) + \alpha\delta(\max\{\hat{S}(\tilde{b}) - \hat{s}, 0\}) - (\tilde{b} - b) \\
&= \alpha\phi(\hat{b}, \hat{s}) - (\hat{b} - b) + \int_{\hat{b}}^{\tilde{b}} [\alpha\phi_b(b', \hat{s}) - 1 + \alpha\delta \max\{\hat{S}'(b'), 0\}]\, db' \\
&\leq \alpha\phi(\hat{b}, \hat{s}) - (\hat{b} - b) \\
&= \beta(\hat{b}, \hat{s}) - (\hat{b} - b),
\end{aligned}$$

where the inequality follows since $\alpha\phi_b(b', \hat{s}) - 1 \leq 0$ and $\hat{S}'(\cdot) \leq 0$ if $\phi_{bs} \leq 0$, and since, for any $b' \geq \hat{b}$,

$$\alpha\phi_b(b', \hat{s}) - 1 + \alpha\delta\hat{S}'(b') \leq \alpha\phi_b(b', \hat{S}(b')) - 1 + \alpha\delta\hat{S}'(b') \leq 0,$$

if $\phi_{bs} \geq 0$.[21] The above result proves that such a deviation is unprofitable. Suppose now that trade does not occur after the deviation. Since $\hat{E}^2(\tilde{b}, \hat{s}) = \hat{E}(\tilde{b}, \hat{s}) = (\tilde{b}, \hat{S}(\tilde{b}))$, the deviation payoff is

$$\delta\alpha\phi(\tilde{b}, \hat{S}(\tilde{b})) - (\tilde{b} - b),$$

[21]The first inequality holds since $\hat{S}'(\cdot) \geq 0$ and the second inequality holds by $(A4)$.

which, by $(A4)$, is nonincreasing in $\tilde{b}$ and attains a value less than $\beta(\hat{b},\hat{s})-(b-\hat{b})$ when $\tilde{b}=\hat{b}$.

A symmetric argument holds for s with respect to both types of deviation.

(iii) $b>\hat{b}$ and $s\le \hat{S}(b)$: The equilibrium payoff for the buyer is $\alpha\phi(b,\hat{S}(b))$ and $(1-\alpha)\phi(b,\hat{S}(b))-(\hat{S}(b)-s)$ for the seller. Consider any deviation by the buyer to $\tilde{b}>b$. Again, $(A4)$ ensures that no profitable deviation exists for the buyer. Its proof follows almost exactly the same line of argument as in the second part of (ii) and is omitted.

Suppose now that the seller deviates to $\tilde{s}$. Assuming an immediate trade, the resulting payoff is

$$\sigma(b,\tilde{s})-(\tilde{s}-s)=\begin{cases}(1-\alpha)\phi(b,\tilde{s})-\alpha\delta(\hat{S}(b)-\tilde{s})-(\tilde{s}-s) & \text{if } \tilde{s}<\hat{S}(b),\\ (1-\alpha)\phi(b,\tilde{s})-(\tilde{s}-s) & \text{if } \tilde{s}>\hat{S}(b).\end{cases}$$

Again, such a deviation is unprofitable since the first line of the RHS is increasing in $\tilde{s}$ (since $\tilde{s}<\hat{S}(b)\le S_\delta(b)$), and its second line is decreasing in $\tilde{s}$ (since $\tilde{s}>\hat{S}(b)\ge S_0(b)$), and they both equal the equilibrium payoff when $\tilde{s}=\hat{S}(b)$. Trade must occur after a deviation if $\tilde{s}>\hat{S}(b)$. If $\tilde{s}<\hat{S}(b)$, then no trade is a possibility but its resulting payoff is simply $\delta[(1-\alpha)\phi(b,\hat{S}(b))-(\hat{S}(b)-\tilde{s})]-(\tilde{s}-s)\le(1-\alpha)\phi(b,\hat{S}(b))-(\hat{S}(b)-s)$ by $(A2)$, so again no such deviation is profitable.

(iv) This case mirrors (iii), so the same proof applies with roles of the seller and the buyer switched. Q.E.D.

Proof of Proposition 2: When the investments are weak substitutes ($\phi_{bs}\le 0$), only $(A1)$-$(A3)$ need to be satisfied. Note that for any $(\hat{b},\hat{s})$ in Ω, there exists $(\delta_1,\delta_2)\in[0,\delta]^2$ such that $\hat{b}=B_{\delta_1}(\hat{s})$ and that $\hat{s}=S_{\delta_2}(\hat{b})$. Define $\hat{B}(s)=B_{\delta_1}(s)$ for $s\ge\hat{s}$ and $\hat{S}(b)=S_{\delta_2}(b)$ for $b\ge\hat{b}$. Then, $\hat{B}$ and $\hat{S}$ are nonincreasing and differentiable almost everywhere, so they satisfy $(A1)$ and $(A3)$. Since $(\delta_1,\delta_2)\in[0,\delta]^2$ and $\hat{B}$ and $\hat{S}$ are nonincreasing (whereas $\underline{B}$ and $\underline{S}$ are nondecreasing), $(A2)$ also holds. Therefore, the proposition holds by Lemma A1. Q.E.D.

Proof of Proposition 3: Suppose to the contrary that $(b_0,s_0)=E(0,0)$ for some regular equilibrium $\langle\beta,\sigma,E\rangle$. In equilibrium, the buyer receives $\beta(b_0,s_0)=\alpha\phi(b_0,s_0)-b_0$. Consider a unilateral deviation by the buyer to $b>b_0$. By **(P)**-(ii), we have $x(b,s_0)=b$ and, by **(P)**-(i), we have $E(E(b,s_0))=E(b,s_0)$. Hence, the (gross) deviation payoff satisfies

$$\begin{aligned}\beta(b,s_0) &\ge \alpha\phi(b,s_0)+\alpha\delta(y(b,s_0)-s_0)\\ &\ge \alpha\phi(b,s_0)+\alpha\delta(S_0(b)-S_0(b_0)),\end{aligned}$$

where the first inequality holds since trade may not follow the deviation and the second inequality follows since, by an argument like that of Proposition 1, $y(b, s_0) \geq S_0(b)$. Since the above inequality must hold for any $b > \hat{b}$, we have

$$\begin{aligned}\liminf_{b \downarrow b_0} \frac{\beta(b, s_0) - \beta(b_0, s_0)}{b - b_0} &\geq \alpha\phi_b(b_0, s_0) + \alpha\delta S_0'(b_0) \\ &> \alpha\phi_b(b_0, s_0) \\ &= 1,\end{aligned}$$

where the second inequality follows from $S_0'(b_0) > 0$ (implied by the complementarity) and the equality follows from $b_0 = B_0(s_0)$. These inequalities imply that there exists b ($> b_0$) to which the buyer strictly prefers to deviate. We thus have a contradiction. Q.E.D.

Proof of Proposition 4: Set $(\hat{b}, \hat{s}) := (b_{CE}, s_{CE})$ and let

$$\hat{B}(s) := \max\{b_{CE}, B_0(s)\} \ \forall s \geq s_{CE} \text{ and } \hat{S}(b) := \max\{s_{CE}, S_0(b)\} \ \forall b \geq b_{CE}.$$

Then, the associated policy rule, $\hat{E}$, defined in $(A5)$ satisfies $(A1)$-$(A3)$, so it suffices to check $(A4)$. For $b \in [b_{CE}, S_0^{-1}(s_{CE})]$, $\hat{S}(b)$ is flat, so we have

$$\alpha[\phi_b(b, \hat{S}(b)) + \delta\hat{S}'(b)] - 1 = \alpha\phi_b(b, \hat{S}(b)) - 1 \leq 0,$$

where the last inequality follows from $b \geq B_0(\hat{S}(b))$. Thus the condition is satisfied. For $b > S_0^{-1}(s_{CE})$, the expression in question is

$$\alpha[\phi_b(b, S_0(b)) + \delta S_0'(b)] - 1 < \alpha\phi_b(b, S_0(b)) - 1 + \alpha M. \tag{11}$$

Observe that $\alpha\phi_b(b, S_0(b)) - 1 < 0$ at $b = S_0^{-1}(s_{CE}) > B_0(s_{CE})$. Hence, for some $M > 0$, the RHS in (11) is nonpositive. Since $\alpha\phi_b(b, S_0(b))$ is nonincreasing (by the concavity of ϕ), for the same $M > 0$, the expression in (11) remains nonpositive for all $b > S_0^{-1}(s_{CE})$. Therefore, there exists $M > 0$ such that the first part of $(A4)$ holds. By the symmetric argument, the second part of $(A4)$ holds for some $M > 0$.

We now turn to the case in which $(b_\delta, s_\delta) \in IR(\alpha)$. Set $\hat{B}(s) = B_\delta(s)$ for $s \geq s_\delta$ and $\hat{S}(b) = S_\delta(b)$ for $b \geq b_\delta$. Again, it suffices to check $(A4)$.

$$\begin{aligned}\alpha[\phi_b(b, \hat{S}(b)) + \delta\hat{S}'(b)] - 1 &= \alpha[\phi_b(b, S_\delta(b)) + \delta S_\delta'(b)] - 1 \\ &\leq \alpha[\phi_b(b_\delta, s_\delta)) + \delta S_\delta'(b)] - 1\end{aligned}$$

$$
\begin{aligned}
&\leq \alpha\left[\phi_b(b_\delta, s_\delta)) + \delta\frac{1-\alpha}{\alpha}\right] - 1 \\
&= \alpha\phi_b(b_\delta, s_\delta)) - (1-(1-\alpha)\delta) \\
&= 0,
\end{aligned}
$$

where the first inequality follows from $b \geq B_\delta(S_\delta(b))$ (by our assumption of a unique intersection of B_δ and S_δ), and the second follows from the hypothesis.

Likewise,

$$
\begin{aligned}
\delta\alpha[\phi_b(b,\hat{S}(b)) + \phi_s(b,\hat{S}(b))\hat{S}'(b)] - 1 &= \delta\alpha[\phi_b(b,S_\delta(b)) + \phi_s(b,S_\delta(b))S'_\delta(b)] - 1 \\
&= \delta\alpha\left[\phi_b(b,S_\delta(b)) + \frac{1-\alpha\delta}{1-\alpha}S'_\delta(b)\right] - 1 \\
&\leq \delta\alpha\left[\phi_b(b_\delta,s_\delta) + \frac{1-\alpha\delta}{\alpha}\right] - 1 \\
&= \delta\left[\alpha\phi_b(b_\delta,s_\delta) - \frac{1-\delta}{\delta} - \alpha\delta\right] \\
&\leq \delta\left[\alpha\phi_b(b_\delta,s_\delta) - (1-(1-\alpha)\delta)\right] \\
&= 0.
\end{aligned}
$$

Symmetric proofs hold for the derivatives with respect to s. The proof is then complete by invoking Lemma A1. Q.E.D.

Proof of Proposition 6: Set $(\hat{b},\hat{s}) := (b_0, s_0)$ and let

$$\hat{B}(s) := B_0(s)\ \forall s \geq s_0 \text{ and } \hat{S}(b) := S_0(b)\ \forall b \geq b_o.$$

Then, the associated policy rule, $\hat{E}$, defined in $(A5)$ is sustainable as a Markov Perfect equilibrium.[22] Consider now a modified policy rule:

$$\hat{E}_d(b,s) \equiv \langle \hat{x}_d(b,s), \hat{y}_d(b,s)\rangle \equiv \begin{cases} (\nu,\epsilon) & \text{if } (b,s) < (\nu,\epsilon), \\ \hat{E}(b,s) & \text{otherwise,} \end{cases}$$

for some $(\nu,\epsilon) << (b_0,s_0)$.

Given this rule, the parties first invest to (ν,ϵ), delay agreement for one period, and then they invest up to (b_0,s_0) next period and trade immediately. We show that there exists $\hat{\delta} < 1$ such that, for any $\delta \geq \hat{\delta}$, $\hat{E}_d$ is sustainable as a Markov Perfect equilibrium.

[22]Observe that $\hat{B}(s)$ and $\hat{S}(b)$ satisfy $(A1)$-$(A4)$ (with $(A4)$ being satisfied by default since $\phi_{bs} \leq 0$), so the sustainability of $\hat{E}$ follows from Lemma A1.

Observe first that, since $(\nu, \epsilon) << (b_0, s_0)$, there exists $\hat{\delta}_1 < 1$ such that, for any $\delta \geq \hat{\delta}_1$,

$$\phi(\nu, \epsilon) \leq \delta[\phi(b_0, s_0) - (b_0 - \nu) - (s_0 - \epsilon)].$$

This means that, for such a δ, the parties will indeed choose to delay agreement after arriving at (ν, ϵ).

We now show that there is no profitable deviation from $\hat{E}_d$, for a sufficiently high δ. Consider any starting stock (b, s). If $(b, s) \not< (\nu, \epsilon)$, then the policy rule, $\hat{E}_d(b', s')$ coincides with $\hat{E}(b', s')$ for any $(b', s') \geq (b, s)$. Then, Lemma A1 proves immediately that no deviation from $\hat{E}_d$ is strictly profitable. Therefore, we restrict attention to an arbitrary starting stock $(b, s) < (\nu, \epsilon)$ and show that there is no profitable deviation from $\hat{E}_d$ in the first period. (If no deviation occurs in the first period (i.e., if they choose (ν, ϵ)), then $\hat{E}$ again applies, so no deviation is profitable from the second period onward.) We focus on the buyer's incentive. If no party deviates in the first period, the buyer's equilibrium payoff would be

$$\beta(\nu, \epsilon) - (\nu - b) = \delta[\alpha\phi(b_0, s_0) - (b_0 - \nu)] - (\nu - b). \tag{12}$$

Suppose that he deviates to $\tilde{b} \neq \nu$. Assume first that the deviation is followed by no trade in that period. Then, the buyer receives:

$$\begin{aligned} &\beta(\tilde{b}, \epsilon) - (\tilde{b} - b) \\ = &\begin{cases} \delta\{\delta[\alpha\phi(b_0, s_0) - (b_0 - \nu)] - (\nu - \tilde{b})\} - (\tilde{b} - b) & \text{if } \tilde{b} < \nu, \\ \delta[\alpha\phi(b_0, s_0) - (b_0 - \tilde{b})] - (\tilde{b} - b) & \text{if } \tilde{b} \in (\nu, b_0] \\ \delta\alpha\phi(\tilde{b}, \max\{S_0(\tilde{b}), \epsilon\}) - (\tilde{b} - b) & \text{if } \tilde{b} > b_0. \end{cases} \end{aligned}$$

This deviation payoff is no greater than the RHS of (12). This holds clearly if $\tilde{b} < \nu$ since $\delta < 1$ (and since the expression inside the curled brackets is nonnegative). The same holds if $\tilde{b} \geq \nu$ since the deviation payoff is nonincreasing in $\tilde{b}$ and equals to the RHS of (12) when $\tilde{b} = \nu$. We thus conclude that deviation is unprofitable in this case.

Assume next that the deviation is followed by an immediate trade. Then, the buyer's deviation payoff equals

$$\begin{aligned} &\beta(\tilde{b}, \epsilon) - (\tilde{b} - b) \\ = &\begin{cases} \alpha\phi(\tilde{b}, \epsilon) + \alpha\delta^2(s_0 - \epsilon) - (1-\alpha)\delta^2(b_0 - \nu) - (1-\alpha)\delta(\nu - \tilde{b}) - (\tilde{b} - b) & \text{if } \tilde{b} < \nu, \\ \alpha\phi(\tilde{b}, \epsilon) + \alpha\delta(\max\{s_0 - \epsilon, 0\}) - (1-\alpha)\delta(b_0 - \tilde{b}) - (\tilde{b} - b) & \text{if } \tilde{b} \in [\nu, b_0], \\ \alpha\phi(\tilde{b}, \epsilon) + \alpha\delta(S_0(\tilde{b}) - \epsilon) - (\tilde{b} - b) & \text{if } \tilde{b} > b_0, \end{cases} \end{aligned}$$

We first show that this deviation payoff is no greater than

$$\alpha\phi(b_0,\epsilon)+\alpha\delta(s_0-\epsilon)-(b_0-b). \tag{13}$$

This result holds if $\tilde{b}<\nu$ since

$$\begin{aligned}
& \alpha\phi(\tilde{b},\epsilon)+\alpha\delta^2(s_0-\epsilon)-(1-\alpha)\delta^2(b_0-\nu)-(1-\alpha)\delta(\nu-\tilde{b})-(\tilde{b}-b) \\
\leq\ & \alpha\phi(\nu,\epsilon)+\alpha\delta^2(s_0-\epsilon)-(1-\alpha)\delta^2(b_0-\nu)-(\nu-b) \\
\leq\ & \alpha\phi(b_0,\epsilon)+\alpha\delta^2(s_0-\epsilon)-(b_0-b) \\
\leq\ & \alpha\phi(b_0,\epsilon)+\alpha\delta(s_0-\epsilon)-(b_0-b)
\end{aligned}$$

where the first inequality holds since the first line is nondecreasing in $\tilde{b}$ for $\tilde{b}\leq\nu\leq b_0\leq B_0(\epsilon)$ and it equals the second line when $\tilde{b}=\nu$, and the second inequality holds since the second line is nondecreasing in ν for $\nu\leq b_0\leq B_0(\epsilon)$ and it equals the third line when $\nu=b_0$, and the last inequality follows from $\delta<1$. Suppose now $\tilde{b}\geq\nu$. In this case, the deviation payoff is nondecreasing in $\tilde{b}$ for $\tilde{b}\leq b_0$ and nonincreasing in $\tilde{b}$ for $\tilde{b}\geq b_0$, and it equals the expression in (13) when $\tilde{b}=b_0$. Hence, again the deviation payoff is no greater than the expression in (13).

It now remains to show that, for sufficiently large δ, the expression in (13) is no greater than the RHS of (12), the equilibrium payoff. This holds since there exists $\hat{\delta}_2<1$ such that, for any $\delta\geq\hat{\delta}_2$,

$$\begin{aligned}
& \alpha\phi(b_0,\epsilon)+\alpha\delta(s_0-\epsilon)-(b_0-b) \\
\leq\ & \delta\alpha\phi(b_0,s_0)-(b_0-b),
\end{aligned}$$

where the inequality follows from the fact that $\alpha\phi(b_0,\epsilon)+\alpha\delta(s_0-\epsilon)-(b_0-b)$ increases strictly as ϵ rises to s_0. Since the last line equals the equilibrium payoff described in (12), the deviation is unprofitable.

Next, applying a symmetric argument, there exists $\hat{\delta}_3<1$ such that, for any $\delta\geq\hat{\delta}_3$, there is no profitable deviation from $\hat{E}_d$ by the seller.

Last, combining all observations, we conclude that for $\delta>\hat{\delta}\equiv\max\{\hat{\delta}_1,\hat{\delta}_2,\hat{\delta}_3\}$, the policy rule, $\hat{E}_d$, constitutes a Markov Perfect equilibrium and a delay occurs on that path. Q.E.D.

References

[1] Admati, A.R., and Perry, M. (1991). "Joint Projects without Commitment," *Review of Economic Studies*, 58, 259-276.

[2] Aghion, P., and Bolton, P. (1992). "An Incomplete Contracting Approach to Financial Contracting," *Review of Economic Studies*, 59, 473-493.

[3] Aghion, P., and Tirole, J. (1997). "Formal and Real Authority in Organizations" *Journal of Political Economy*, 105, 1-29.

[4] Bagnoli, M., and Lipman, B. (1989). "Provision of Public Goods: Fully Implementing the Core through Private Contributions," *Review of Economic Studies,* 56, 583-601.

[5] Binmore, K. (1987). "Perfect Equilibria in Bargaining Models" *The Economics of Bargaining* (Binmore and Dasgupta eds.) Basil Blackwell, Oxford, 77-105.

[6] Che, Y.-K., and Hausch, D.B. (1999). "Cooperative Investments and the Value of Contracting," *American Economic Review,* 125-147.

[7] Dewatripont, M., and Tirole, J. (1994). "Theory of Debt and Equity: Diversity of Securities and Manager-Shareholder Congruence," *Quarterly Journal of Economics*, 109, 1027-1054.

[8] Gale, D. (2000). "Monotone Games with Positive Spillovers," mimeo., New York University.

[9] Grossman, S., and Hart, O. (1986). "The Costs and Benefits of Ownership: A Theory of Lateral and Vertical Integration," *Journal of Political Economy,* 94, 691-719.

[10] Grout, P.A. (1984). "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach," *Econometrica*, 52, 449- 460.

[11] Hart, O.D., and Moore, J.D. (1990). "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98, 1119-1158.

[12] Hart, O.D., and Moore, J.D. (1999) "Foundations of Incomplete Contracts," *Review of Economic Studies*, 66, 115-138.

[13] Klein, B., Crawford, R., and Alchian, A. (1978). "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law and Economics*, 21, 297-326.

[14] Lockwood, B., and Thomas, J. (2001). "Gradualism and Irreversibility," *Review of Economic Studies*, forthcoming.

[15] Macaulay, S. (1963). "Non-Contractual Relations in Business: A Preliminary Study," *American Sociological Review*, 28, 55-70.

[16] Marx, L.M., and Matthews, S.A. (2000). "Dynamic Voluntary Contribution to a Public Project," *Review of Economic Studies*, 67, 327-358.

[17] Rubinstein, A. (1982). "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97-109.

[18] Shaked, A. (1994). "Opting Out: Bazaars versus Hi-Tech Markets," *Investigaciones Económicas,* 421-432.

[19] Segal, I.R., (1999) "Complexity and Renegotiation: A Foundation for Incomplete Contracts," *Review of Economic Studies*, 66, 57-82.

[20] Sutton, J. (1986). "Noncooperative Bargaining Theory: An Introduction," *Review of Economic Studies,* 53, 709-724.

[21] Tirole, J. (1986). "Procurement and Renegotiation," *Journal of Political Economy,* 94, 235-259.

[22] Tirole, J. (1999). "Incomplete Contracts: Where Do We Stand?" *Econometrica,* 67, 741-781.

[23] Williamson, O. (1979). "Transactions-Cost Economics: The Governance of Contractual Relations," *Journal of Law and Economics*, 22, 233-262.