A Dynamic Theory of Holdup

Yeon-Koo Che* József Sákovics†

April 25, 2003

Abstract

The holdup problem arises when parties negotiate to divide the surplus generated by their relationship specific investments. We study this problem in a dynamic model of bargaining and investment which, unlike the stylized static model, allows the parties to continue to invest until they agree on the terms of trade. The investment dynamics overturns the conventional wisdom dramatically. First, the holdup problem need not entail underinvestment when the parties are sufficiently patient. Second, inefficiencies can arise unambiguously in some cases, but they are not caused by the sharing of surplus *per se* but rather by a failure of an individual rationality constraint.

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

---

*Department of Economics, University of Wisconsin-Madison, Madison, WI 53706, USA.

†Edinburgh School of Economics, University of Edinburgh, 50 George Square, Edinburgh, EH8 9JY, UK.

Both authors are grateful for comments from Andy Postlewaite (the Editor in charge) and two anonymous referees as well as from Jennifer Arlen, Hongbin Cai, Ian Gale, Don Hausch, Jinwoo Kim, Bart Lipman, Bentley MacLeod, Leslie Marx, Steven Matthews, John Moore, Larry Samuelson, Bill Sandholm, Jonathan Thomas and the seminar participants at the Universities of Arizona, Bristol, Edinburgh, Florida, Heriot-Watt, Southampton, Southern California (John Olin Law School Lecture), Texas, UCLA, and Washington, and Duke, Northern Illinois, Northwestern (Kellogg School of Management), Ohio State Universities, and University College London. 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

Economic agents often need to make sunk investments whose returns are vulnerable to *ex post* expropriation by their partners. One such phenomenon arises when trading partners negotiate to divide their trade surplus after making relationship-specific investments. This problem, known as holdup, is inherent in many bilateral exchanges. For instance, workers and firms often invest in firm-specific assets whose returns are shared through subsequent wage negotiation. Manufacturers and suppliers customize their equipment and production processes to their partners, knowing well that the benefit will be shared through future (re)negotiation. Assuring specific investments is critically important in modern manufacturing, with the increased need for coordination across different production stages and the availability of the technologies facilitating such coordination.[1] Elements of holdup are also present in other settings such as team production, reallocation/dissolution of partnership assets, and even in political lobbying (e.g., a campaign contribution can be seen as a "sunk investment").

The risk of these investors being held up has inspired much of the modern contracts and organizations theory. Various remedies have been proposed as safeguards against holdup, ranging from vertical integration (Klein, Crawford and Alchian, 1978; Williamson, 1979), property rights allocation (Grossman and Hart, 1986; Hart and Moore, 1990), contracting on renegotiation rights (Chung, 1991; Aghion, Dewatripont and Rey, 1994), option contracts (Nöldeke and Schmidt, 1995, 1998), production contracts (Edlin and Reichelstein, 1996), relational contracts (Baker, Gibbons and Murphy, 2002), financial rights allocation (Aghion and Bolton, 1992; Dewatripont and Tirole, 1994; Dewatripont, Legros, and Matthews, 2002) and hierarchical authority (Aghion and Tirole, 1997) to injecting market competition (MacLeod and Malcomson, 1993; Acemoglu and Shimer, 1999; Cole, Mailath and Postlewaite, 2001; Felli and Roberts, 2000; Che and Gale, 2003). Underlying all these theories is the premise that, without these special protections, the holdup problem will lead parties to underinvest in specific assets (see Grout, 1984; and Tirole, 1986). The purpose

[1]For example, Asanuma (1989) describes how suppliers customize parts for buyers even when "specific investments...have to be incurred to implement such customization." Xerox incorporates supplier-designed components into many of its products, which requires idiosyncratic adaptations of production lines and procedures to individual suppliers (Burt, 1989). Some Japanese automakers pay for consultants to work with suppliers, possibly for months, to improve production methods (Dyer and Ouchi, 1993). Vauxhall "regularly works in partnerships with suppliers to improve efficiency and trim costs. It helped suppliers reduce costs by 30-40%" (*The Engineer*, 1996).

of the current paper is to reexamine this “conventional wisdom.”[2]

Our point of departure is the observation that the stylized model predicting underinvestment does not capture the rich dynamic interaction present in many trading relationships. For instance, the standard two-stage model assumes that the trading partners invest only once, at a pre-specified time, 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 agent makes a specific investment targeted at a particular partner, it is plausible for him to approach this partner to negotiate trade terms even before his investment is completed.

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 (more) to invest, and then a (randomly chosen) party offers some terms of trade. If the offer is accepted, then trade occurs according to the agreed terms and the game ends. If the offer is rejected, however, the game moves on to the next period 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.

Our analysis focuses on the sustainability of an outcome that approaches efficiency as the parties' discount factor approaches 1 (or equivalently, as the time interval shrinks to zero). Whether such *asymptotic efficiency* holds depends on a version of *individual rationality*, which, roughly speaking, requires that each party recoup his investment costs at the efficient investment pair *when* the parties split the gross trade surplus in proportion to their relative bargaining power (i.e., just as they would in the static equilibrium). Clearly, this condition is not sufficient for the efficient investment to arise in the static model. In our dynamic model, however, it is sufficient for the existence of a (Markov Perfect) equilibrium whose outcome approaches the first-best outcome as the parties' discount factor tends to 1 (Proposition 1). In this equilibrium, holdup still arises on the equilibrium path in that

[2]Hence, rather than looking for additional possible remedies, we investigate whether such remedies are warranted in the first place. In Section 8 we discuss several implications of our results for the incomplete contracts view on organization.

a party receives only a fraction of the gross surplus commensurate with his bargaining power. Yet, this does not stop the party from investing efficiently. The key reason is the investment dynamics. Suppose that a party invests low today, but that he is expected to raise his investment tomorrow in case no agreement is reached today. Then, there will be more surplus to divide tomorrow than there is today. Since the cost of tomorrow's investment will be borne solely by the investor, the prospect of the investor raising his investment tomorrow causes his partner to demand more to settle today. The investment dynamics thus results in a worse bargaining position for the party upon investing low, and thus creates a stronger incentive for raising investment, than would be possible if such investment dynamics were not allowed. (This point will be illustrated in an example in the next section.)

Next, we show that the individual rationality condition (in a slightly weaker version) is also necessary for asymptotic efficiency, when the parties' investments are weakly substitutable (in a sense to be made precise later): If the individual rationality constraint fails, then the cumulative investment pair sustainable in any subgame perfect equilibrium is bounded away from the first-best pair, regardless of the discount factor (Proposition 3). If the investments are complementary, however, the parties may be able to exploit the investment dynamics to adjust their bargaining shares and thus attain asymptotic efficiency even when our individual rationality condition fails.

These results have several broad implications. First, they suggest that a simple — and reasonable — investment dynamics alone can virtually eliminate the inefficiencies, as long as the holdup problem is not too severe to fail a certain individual rationality condition. This suggests that the holdup problem may not be such a worrisome source of inefficiencies, thus calling into question our reliance on the holdup problem as a rationale for organization theory.

Second, they also explain why the parties may not need contractual protection to achieve efficiency. This may explain why business transactions seldom rely on explicit contracts (Macaulay, 1963). Since the absence of contracts is an extreme form of contractual incompleteness, our findings can also provide a new foundation for incomplete contracts, perhaps better than its extant counterparts.[3]

Third, even when inefficiencies arise from the holdup problem, the warranted organizational re-

[3]The existing foundations require either some unjustifiable notion of indescribable contingencies (see the criticism of the latter in Tirole (1999) and Maskin and Tirole (1999)) or a very strong notion of contract renegotiability (assumed for instance in Che and Hausch (1999), Segal (1999) and Hart and Moore (1999)). The current result avoids such criticisms since it rests on the result that efficiency is virtually achievable even without contracts.

sponses may be different from those prescribed by the existing literature. In particular, inefficiencies need not be caused by the sharing of surplus *per se* but rather by the failure of a certain individual rationality constraint, thus suggesting its relaxation as an important role of contract/organization design. This has a specific implication, for instance, for the effectiveness of the institutional tools influencing parties' relative bargaining power.[4] The existing wisdom due to Holmstrom (1982) is that such tools alone cannot be useful if both parties make specific investments, since there always exists a party who appropriates less than full marginal return to his investment.[5] In our model, individual rationality always holds for *some* bargaining shares, so such tools can be effective at least from the asymptotic efficiency perspective.

Last, our finding warrants a thorough reexamination of the remedies that have been proposed in the literature. For instance, the nature of inefficiencies found in this paper may provide new insight into how the parties should allocate ownership of critical assets, what type of ex ante trading contracts they may sign, and how the courts should allocate default rights in contract disputes.

The rest of the paper is organized as follows. The next section illustrates the main idea using a simple example. Section 3 presents the model. The existence of an asymptotically efficient equilibrium is shown in Section 4. In Section 5, we establish the necessity of individual rationality for asymptotic efficiency, for weakly substitutable investments. In Section 6, we make some interesting observations about the case of complementary investments. Section 7 discusses related literature. Section 8 concludes. The Appendices contain the proofs not presented in the main body of the paper.

2 A motivating example

The model and the main intuition behind our results can be illustrated via a simple example. Suppose only the seller can invest, and she is faced with a binary choice: either to "not invest" or to "invest". Investment costs her $C > 0$. The gross trade surplus is ϕ_I if she invests and ϕ_N if she does not. Assume that $\phi_I - C > \phi_N$, so that it is efficient to invest. Suppose that the parties have equal bargaining power, meaning that each party becomes the proposer with equal probability. In

[4] Aghion, Dewatripont, and Rey (1994) discusses ways of shifting bargaining power using cash bonds, for instance.

[5] Chung (1991) and Aghion, Dewatripont and Rey (1994) show that efficiency is achievable with a contract that shifts bargaining power, but in this case the contract *also* affects the status quo payoffs.

the static model, the seller will not invest if

$$\tfrac{1}{2}\phi_I - C < \tfrac{1}{2}\phi_N. \tag{1}$$

In our dynamic model, however, there exists an equilibrium in which the seller invests, provided that the parties' common discount factor, δ, is sufficiently large and that the investment is *individually rational* for the seller given equal sharing of the pie:[6]

$$\tfrac{1}{2}\phi_I - C \geq 0. \tag{2}$$

Consider the strategy by the seller to "invest whenever no investment has been made before." If the seller indeed invests, then, since no further investment is possible, the ensuing subgame coincides with the standard (random-proposer) bargaining game with a fixed surplus. Consequently, the parties will split ϕ_I equally (on average) in its unique equilibrium. Hence, the seller's equilibrium payoff will be $\frac{1}{2}\phi_I - C$, just as in the static analysis. That is, the seller would be held up in terms of her absolute payoff even in the dynamic model.

Suppose now that the seller deviates and chooses "not invest." Invoking the one-period deviation principle, she *will invest* next period if no agreement is reached this period. Given this, the buyer's continuation payoff following no agreement is $\delta(\frac{1}{2}\phi_I)$, so he will never agree to trade, following the seller's deviation, unless he receives at least this amount. Thus, the seller's payoff from "not invest" is at most $\max\{\phi_N - \delta(\frac{1}{2}\phi_I), \delta(\frac{1}{2}\phi_I - C)\}$: the former payoff is received if the seller offers $\delta(\frac{1}{2}\phi_I)$, which the buyer will accept; the latter payoff is received if the seller offers a lower amount (which the buyer will reject) or if the buyer becomes the proposer (in which case he will offer the seller's net continuation value, $\delta(\frac{1}{2}\phi_I - C)$). Given (2), for δ close to 1, both payoffs are less than $\frac{1}{2}\phi_I - C$ — the payoff the seller will receive by investing now. Since the (one-period) deviation is not profitable, it is a subgame perfect equilibrium for the seller to invest, for a sufficiently large δ, given (2) (even when (1) holds).[7]

This example illustrates how the simple dynamics — the mere *possibility* of adding investment later — can create stronger incentives than in the static model. It also highlights the *necessity* of

[6]The zero reservation payoff on the RHS is not given by the seller's outside option but rather by her internal option of "not investing and perpetually inducing rejection." Note also that $\phi_N > 0$ is necessary for both (1) and (2) to hold simultaneously. This feature is needed here only to illustrate our result in binary investment setting. No such assumption is needed in a general model since the inefficiency need not involve zero investments.

[7]There is another equilibrium in which the seller does not invest, supported by the pessimistic belief that she will never invest in the future.

the individual rationality constraint for sustaining efficiency. Clearly, the efficient outcome would not be sustained if (2) failed. In the sequel, we show the sufficiency and necessity of individual rationality for sustaining asymptotic efficiency in a more general environment.

3 The model

Two risk-neutral parties, a buyer and a seller, make sunk investments to increase the gains from their potential trade of a good. Time flows in discrete periods of equal length, $t = 1, 2, ...$, 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. Let $\mathcal{X} := [0, \bar{b}]$ and $\mathcal{Y} := [0, \bar{s}]$ be the feasible sets of cumulative investments for the buyer and the seller, respectively, for some large $\bar{b}$ and $\bar{s}$. (It is also useful to define $\mathcal{X}(z) := \{b \in \mathcal{X} | b \geq z\}$ and $\mathcal{Y}(z) := \{s \in \mathcal{Y} | s \geq z\}$.) If the parties trade in period t, with the cumulative investments of $b \geq 0$ by the buyer and of $s \geq 0$ by the seller, then they realize the joint surplus of $\phi(b, s)$ in that period (which amounts to $\delta^t \phi(b, s)$ in period 1 terms). They realize zero gross payoffs if no trade occurs.

We make several assumptions on ϕ. First, we assume that $\phi(\cdot, \cdot)$ is twice continuously differentiable, strictly increasing and strictly concave.[8] Further, we require that ϕ is either sub- or super-modular: either $\phi_{bs}(b, s) \leq 0$ or $\phi_{bs}(b, s) > 0$ for all $(b, s) \geq (0, 0)$. This last assumption means that investments by the two parties are either (weak) substitutes or (strict) complements, globally. It simplifies the subsequent analyses and the interpretation of our results.

This basic model applies to a broad range of circumstances. For instance, the trade negotiation may involve various decisions such as the types of the goods traded, and their quantity and quality, as long as they are all verifiable. Let $q \in Q$ denote such a (possibly multidimensional) decision and let $v(q, b, s)$ and $c(q, b, s)$ denote the buyer's gross surplus and the seller's production costs from that decision, respectively. Then, ϕ can be seen as the result of optimizing on q; i.e., $\phi(b, s) := \max_{q \in Q}\{v(q, b, s) - c(q, b, s)\}$. Since the subsequent results will depend only on ϕ, how the investments affect v and c will not be an issue. In particular, our result will not depend on

[8] The strict concavity assumption rules out the case of perfectly substitutable investments (i.e., $\phi(b, s) = \phi(b + s)$), which is plausible in many public good provision problems (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 Corollary 3 following Proposition 1.

whether investment is selfish or cooperative.[9] Our model is also readily extendable beyond bilateral trade settings, for instance to team production, or optimal reallocation/dissolution of partnership assets, with an arbitrary number of agents.[10]

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.[11] 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 if $\delta = 0$), 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 use Subgame Perfect Equilibrium (SPE) as our solution concept. That is, we require that the players' strategies — a pair of functions mapping from observed histories to the investment and bargaining behavior — should form a Nash equilibrium following any feasible history. Sometimes, we consider SPE in Markov strategies or simply Markov Perfect Equilibria (MPE). The strategies in MPE are functions only of payoff-relevant histories, which in our model are the cumulative investment pairs arrived at in each period.

[9] A selfish investment directly benefits the investor while a cooperative investment directly benefits the trading partner of the investor (see Che and Hausch (1999)). They showed that cooperative investments limit the ability of contracting to solve the holdup problem.

[10] Our positive result (Proposition 1) would readily generalize to the environment with more than two agents. The negative result (Proposition 3) will require some restrictions on the equilibrium strategies (such as stationarity) since even a pure bargaining game with more than two agents is known to admit multiple equilibria (see Osborne and Rubinstein (1990), pp. 63-65.).

[11] 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.

It is useful to describe several benchmarks. The following notations will prove useful for this purpose as well as for the subsequent analysis. For $\delta \in [0,1]$, define some (hypothetical) payoff functions,

$$U_\delta^B(b,s;\alpha) := \alpha\phi(b,s) - [1-(1-\alpha)\delta]b, \text{ and } U_\delta^S(b,s;\alpha) := (1-\alpha)\phi(b,s) - [1-\alpha\delta]s,$$

respectively for the buyer and the seller, and let

$$B_\delta(s) := \arg\max_b U_\delta^B(b,s;\alpha) \text{ and } S_\delta(b) := \arg\max_s U_\delta^S(b,s;\alpha)$$

be the corresponding best response functions. (They are well defined since $U_\delta^i(\cdot,\cdot)$ is strictly concave. The dependence on α will be suppressed from now on unless necessary.) Notice that these payoff functions exhibit increasing differences in $(b;\delta)$ and in $(s;\delta)$, respectively. Since the best responses are unique, then $B_0(s) \le B_\delta(s) \le B_1(s)$ for all $s \ge 0$ and $\delta \in [0,1]$, and $S_0(b) \le S_\delta(b) \le S_1(b)$ for all $b \ge 0$ and $\delta \in [0,1]$. Meanwhile, strict concavity of $\phi(\cdot,\cdot)$ implies that

$$\phi_{bb}\phi_{ss} > \phi_{bs}^2 \Leftrightarrow B_\delta' S_\delta' = \left(-\frac{\phi_{bs}}{\phi_{bb}}\right)\left(-\frac{\phi_{bs}}{\phi_{ss}}\right) < 1,$$

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 this intersection (i.e., $b_\delta = B_\delta(s_\delta)$ and $s_\delta = S_\delta(b_\delta)$).

Although the significance of (b_δ, s_δ) will not be immediate for $\delta \in (0,1)$, it can be seen clearly for the extreme values of δ. Consider first $\delta = 1$. Note that $\frac{U_1^B(b,s)}{\alpha} - s = \frac{U_1^S(b,s)}{1-\alpha} - b = \phi(b,s) - b - s$, the joint payoff of the parties. Hence, $B_1(\cdot)$ and $S_1(\cdot)$ are the socially efficient responses. The first-best pair is thus the intersection, (b_1, s_1), of these curves. Consider the other extreme case, with $\delta = 0$. In this case, the payoffs for the parties reduce to $U_0^B(b,s) = \alpha\phi(b,s) - b$ and $U_0^S(b,s) = (1-\alpha)\phi(b,s) - s$, which are their payoffs in the static holdup game. Since $B_0(\cdot)$ and $S_0(\cdot)$ represent the best response functions of the buyer and the seller, respectively, their intersection, (b_0, s_0), will be the subgame perfect equilibrium of that game. To avoid a trivial uninteresting case, we assume that either (b_0, s_0) or (b_1, s_1) is in the interior of $\mathcal{X} \times \mathcal{Y}$, which implies that $(b_0, s_0) \neq (b_1, s_1)$; i.e., the static outcome is inefficient. Finally, observe that (b_δ, s_δ) converges to the first-best pair (b_1, s_1) as $\delta \to 1$. The next section will show that (b_δ, s_δ) is sustainable in equilibrium for a sufficiently high δ if a certain individual rationality condition holds.

4 Asymptotic efficiency

In this section, we will investigate the circumstances under which the first-best outcome can be approximated arbitrarily closely in equilibrium, as $\delta \to 1$. Specifically, we will establish that the pair (b_δ, s_δ) can be implemented as an MPE for a sufficiently large δ, given the following condition:

$$(SIR_\alpha) \qquad\qquad U_0^B(b_1, s_1; \alpha) > 0 \text{ and } U_0^S(b_1, s_1; \alpha) > 0.$$

In words, this condition says that the parties recoup their investment costs at the efficient pair, *if they split the gross surplus precisely the same way as they would in the static holdup game.* Whether (SIR_α) holds depends on the relative bargaining power as well as the investment expenditure required for the first-best pair: the higher α and lower b_1, the easier the condition is to satisfy for the buyer and the harder for the seller. More importantly, (SIR_α) would never be sufficient for the first-best pair to be sustainable in the static model. Indeed, the unique static outcome is inefficient in our model even with (SIR_α). By contrast, (SIR_α) is sufficient for asymptotic efficiency:

Proposition 1 *(Asymptotic Efficiency) Given* (SIR_α)*, there exists a* $\delta^* < 1$ *such that, for all* $\delta \geq \delta^*$*, there exists a MPE in which the parties choose* (b_δ, s_δ) *and trade in the first period. That is, given* (SIR_α)*, there exists an MPE that implements the first-best arbitrarily closely as* $\delta \to 1$*.*

Proof. See Appendix A.

While the general proof is relegated to the appendix, its intuition can be illustrated easily in the case of weakly substitutable investments. Consider the Markovian investment strategies, $I : \mathcal{X} \times \mathcal{Y} \mapsto \mathcal{X} \times \mathcal{Y}$, which maps from the previous period cumulative investment pair into the current period pair:

$$I_s(b,s) = \begin{cases} (b_\delta, s_\delta) & \text{if } (b,s) \leq (b_\delta, s_\delta) \quad [\textit{region (i)}\ ], \\ (b, S_\delta(b)) & \text{if } b > b_\delta \text{ and } s \leq S_\delta(b) \quad [\textit{region (ii)}], \\ (B_\delta(s), s) & \text{if } s > s_\delta \text{ and } b \leq B_\delta(s) \quad [\textit{region (iii)}], \\ (b,s) & b \geq B_\delta(s) \text{ and } s \geq S_\delta(b) \quad [\textit{region (iv)}\ ]. \end{cases}$$

The associated bargaining strategies are for each proposer to choose between offering the discounted continuation payoff of the receiver, given the future investment path, and offering any rejectable offer, whichever is more profitable, and for the receiver to accept if and only if the offer is no less than his discounted continuation payoff, given the future investment path.

Figure 1(a) depicts the investment strategies in "phase diagram" form.

[Insert Figure 1(a) and 1(b) around here.]

This strategy has the same flavor as the one in our motivating example, in that a party, say the buyer, whenever coming up short of the target, b_δ, will invest up to that target when he gets a chance to invest in the next period.

To see the relevance of (SIR_α), suppose both parties follow the equilibrium strategies and choose (b_δ, s_δ). According to I_s, the parties never subsequently invest (i.e., $I_s(b_\delta, s_\delta) = (b_\delta, s_\delta)$). Hence, the ensuing subgame collapses to a pure bargaining game with a fixed surplus, $\phi(b_\delta, s_\delta)$. The equilibrium shares are then uniquely determined as $(\alpha, 1-\alpha)$ (see Binmore (1987)), so the parties will receive $U_0^B(b_\delta, s_\delta; \alpha)$ and $U_0^S(b_\delta, s_\delta; \alpha)$. Since $(b_\delta, s_\delta) \to (b_1, s_1)$ as $\delta \to 1$, (SIR_α) ensures that

$$U_0^B(b_\delta, s_\delta; \alpha) \geq 0 \text{ and } U_0^S(b_\delta, s_\delta; \alpha) \geq 0, \tag{3}$$

for sufficiently large δ. Since each party has an option of making no investment and ensuring himself at least zero payoff, (3) is necessary for I_s to be an equilibrium. In fact, (3) (and thus (SIR_α)) is sufficient for I_s to be sustainable.[12]

To illustrate, suppose that the buyer deviates to $b > b_\delta$ while the seller chooses his target $s = s_\delta$. Then, they find themselves in region (iv). Since no further investment is prescribed by the strategy, again the ensuing subgame becomes a pure bargaining game with a fixed surplus of $\phi(b, s_\delta)$, from which the buyer receives

$$\alpha\phi(b, s_\delta) - b = U_0^B(b, s_\delta), \tag{4}$$

Recall that this payoff coincides with the payoff that the buyer would have received in the static model. Since $U_0^B(b, s_\delta)$ declines in b for $b \geq b_\delta = B_\delta(s_\delta) \geq B_0(s_\delta)$ (as depicted in Figure 1(b)), the buyer will never deviate to $b > b_\delta$.

Suppose next that the buyer deviates to $b < b_\delta$. In this case, the strategy of him investing back to the target next period (if no agreement is reached in the current period) means that the payoff facing the buyer will no longer coincide with $U_0^B(b, s_\delta)$. Instead, his payoff given the above strategy turns out to be

$$\max\{U_\delta^B(b, s_\delta) - \delta(1-\alpha)b_\delta,\ \delta U_0^B(b_\delta, s_\delta) - (1-\delta)b\}, \tag{5}$$

where the first term corresponds to the case in which the deviation is followed by an immediate agreement to trade and the second term corresponds to the case in which the deviation is followed

[12] As can be seen from the proof of Proposition 1, (3) is sufficient for asymptotic efficiency. In many cases, however, a weak inequality version of (SIR_α), (IR_α) introduced in the next section, is sufficient for (3), in which case (IR_α) will be necessary and sufficient for asymptotic efficiency, as will be seen in the next section.

by a rejection.[13] As depicted in Figure 1(b), the incentives associated with this payoff are steeper than the one generated in the static model. Specifically, the second term can never dominate the equilibrium payoff of $U_0^B(b_\delta, s_\delta)$ as long as it is nonnegative, which will hold for a sufficiently large δ given (SIR_α). Meanwhile, the first term of (5) is strictly increasing in b and attains its maximum at $b = b_\delta$ and equals $U_0^B(b_\delta, s_\delta)$. Hence, again buyer has no incentive to deviate to $b < b_\delta$.

If the investments are strictly complementary, the above strategies may not work. In that case, $B_\delta(\cdot)$ and $S_\delta(\cdot)$ are strictly increasing, so, for instance, the buyer may wish to raise her investment in region (ii) to credibly force the seller to invest even further, which may lower the latter's continuation payoff and improve the buyer's current-period bargaining position. In that case, the strategies should be modified to control such incentives. A large part of the proof in the appendix addresses this issue.

A few implications of the asymptotic efficiency result can be drawn by investigating some special cases for which (SIR_α) is expected to hold. First, it must be clear that (SIR_α) holds for at least some values of α.

Corollary 1 *(Fair bargaining shares) There exists* $\alpha \in [0, 1]$ *for which asymptotic efficiency is achievable.*

Proof. For any $\alpha \in [0, 1]$, we have

$$U_0^B(b_1, s_1) + U_0^S(b_1, s_1) = \phi(b_1, s_1) - b_1 - s_1 > 0.$$

The result holds since $U_0^i(b_1, s_1)$, $i = B, S$, is continuous in α, takes a negative value for some α, and takes a positive value for some other values of α. ▌

[13] If there is no agreement in the current period, then the strategy prescribes the buyer to invest to b_δ next period, which as noted above will trigger a pure bargaining game with a pie of $\phi(b_\delta, s_\delta)$. Hence, the buyer's discounted continuation payoff is $\delta[\alpha\phi(b_\delta, s_\delta) - (b_\delta - b)]$, while that of the seller is $\delta(1-\alpha)\phi(b_\delta, s_\delta)$. Thus, the buyer's deviation payoff from choosing $b < b_\delta$ is

$$\alpha \max\{\phi(b, s_\delta) - \delta(1-\alpha)\phi(b_\delta, s_\delta), \delta[\alpha\phi(b_\delta, s_\delta) - (b_\delta - b)]\} + (1-\alpha)\delta[\alpha\phi(b_\delta, s_\delta) - (b_\delta - b)]. \qquad (*)$$

This is explained as follows. Upon choosing $b < b_\delta$, the buyer is chosen with probability α to become the proposer. In this case, the buyer can make either a minimal acceptable offer, matching the seller's continuation payoff, $\delta(1-\alpha)\phi(b_\delta, s_\delta)$, or a rejectable offer and thus collecting his own discounted continuation payoff $\delta[\alpha\phi(b_\delta, s_\delta) - (b_\delta - b)]$, whichever is more profitable. With probability $1 - \alpha$, the buyer becomes a responder, in which case the buyer will be held down to his discounted continuation payoff, no matter how the seller resolves her trade-off. Payoff $(*)$ simplifies to (5).

This result contrasts with the standard holdup model in which inefficiencies must arise regardless of the relative bargaining power if both parties have continuous investments to make (recall our result with $\delta = 0$). This also implies that efficiency would be achievable if the parties have the institutional tools to manipulate the relative bargaining power, α. If the two parties have a symmetric technology, then the equal bargaining power turns out to be the right one:

Corollary 2 *(Symmetry) Asymptotic efficiency is acheivable if $\mathcal{X} = \mathcal{Y}$, $\phi(b,s) = \phi(s,b)$ and $\alpha = \frac{1}{2}$.*

Proof. By symmetry, we have $b_1 = s_1$. Hence,

$$\tfrac{1}{2}\phi(b_1,s_1) - b_1 = \tfrac{1}{2}\phi(b_1,s_1) - s_1 = \tfrac{1}{2}[\phi(b_1,s_1) - b_1 - s_1] > 0,$$

implying that $(SIR_{1/2})$ holds. ▌

In the public good provision problem, 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:

Corollary 3 *(Perfect Substitutability) Asymptotic efficiency is achievable for any $\alpha \in [0,1]$, if $\phi(b,s) = \psi(b+s)$ for some strictly concave function $\psi(\cdot)$ with $\psi(0) = 0$ and $\lim_{z\downarrow 0}\psi'(z) > 1$ (where the limit is along the points of differentiability).*

Proof. Given the assumption, any (b,s) such that $b+s = z^* := \arg\max_z\{\psi(z) - z\}$ constitutes the first-best outcome. Since $\psi(z^*) - z^* > 0$, for any $\alpha \in [0,1]$, there exists a first-best pair (b,s) with $b+s = z^*$ satisfying (SIR_α). Pick any such pair and call it (b_1,s_1). Next, note that for each δ, both $B_\delta(\cdot)$ and $S_\delta(\cdot)$ are negative 45 degrees lines, and $B_\delta(\cdot)$ lies outside (inside) $S_\delta(\cdot)$ if $\alpha \geq \frac{1}{2}$ ($\alpha \leq \frac{1}{2}$). Assume, without loss of generality, that $\alpha \geq \frac{1}{2}$. Since $S_\delta(\cdot)$ converges to the line, $s = z^* - b$, we can choose (b_δ, s_δ) on $S_\delta(\cdot)$ such that $(b_\delta,s_\delta) \to (b_1,s_1)$ as $\delta \to 1$. Define next the investment strategies,

$$I_{ps}(b,s) = \begin{cases} (b_\delta, s_\delta) & \text{if } (b,s) \leq (b_\delta,s_\delta), \\ (b, S_\delta(b)) & \text{if } b > b_\delta \text{ and } s \leq S_\delta(b), \\ (\min\{b_\delta, B_\delta(s)\}, s) & \text{if } s > s_\delta \text{ and } b \leq \min\{b_\delta, B_\delta(s)\}, \\ (b,s) & \text{otherwise.} \end{cases}$$

The proof of Proposition 1 holds with I_{ps}, which proves that (b_δ, s_δ) is implementable as an MPE. ∎

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 more bargaining power. Consequently, in contribution games individual rationality is not an issue (see Section 7 for further discussion of this literature).

Several further remarks are worth making.

Remark 1 *For weakly substitutable investments, the proof of Proposition 1 also does not rely on the fact that the set of feasible investments is continuous. Hence, the result holds just as well if the feasible levels of investment are discrete.*[14] *Indeed, as the motivating example illustrates, the first-best outcome can be implemented precisely (and not just approximated) for a large $\delta < 1$. The ability to handle the discrete investments case contrasts with the dynamic voluntary contribution literature (c.f. Section 7).*

Remark 2 *One may wonder if Proposition 1 is a result of some folk theorem. A folk theorem does not hold in our model,*[15] *for the same reason that it does not hold in the Rubinstein bargaining model. In these models, there is a single realization of bounded surplus, which limits the power of punishment.*[16] *By contrast, folk theorems hold in models that admit an unbounded stream of surplus, whose destruction can be used as an effective punishment. This difference distinguishes*

[14] We conjecture that this is also true for complementary investments. The required strategies may involve randomization in some cases.

[15] To see this, note that (SIR_α) is stronger than the standard individual rationality invoked in the folk theorem: (SIR_α) requires the first-best pair to yield positive payoffs *conditional on dividing the pie* in the $(\alpha, 1-\alpha)$ ratio, whereas the standard individual rationality would require the parties to receive positive payoffs for *some* feasible sharing rule. Since $\phi(b_1, s_1) - b_1 - s_1 > 0$, this latter condition holds trivially. Hence, if a folk theorem were to hold, (b_1, s_1) should be implementable for a large $\delta < 1$, even when (SIR_α) fails to hold. As must be clear from the motivating example (Section 2) and will be shown formally in Section 5, however, the bargaining shares differing from $(\alpha, 1-\alpha)$ are sometimes unsustainable, in which case (b_1, s_1) cannot be implemented if our version of individual rationality fails. Hence, the folk theorem does not hold.

[16] This feature renders the folk theorem of Dutta (1995) inapplicable in our model. Dutta shows that, for any $\epsilon > 0$, any individually rational payoffs, normalized by a factor of $1-\delta$, can be approximated within ϵ by a perfect equilibrium, as $\delta \to 1$. Given our bounded surplus, however, any feasible payoff *normalized* by $(1-\delta)$ would converge to zero as $\delta \to 1$, so the result of ε-approximation has no bite in our model.

our efficiency result from that obtainable in repeated game models (see Baker et. al. (2002), Halonen (2002) and MacLeod and Malcomson (1989)).

Remark 3 *The equilibrium constructed is not only Markov perfect, but it also satisfies a certain passivity property of the beliefs: i.e., a deviation triggers a minimal revision of the equilibrium investment plan. Che and Sákovics (2001) show that, with weakly substitutable investments, any investment pair* (b, s) *is implementable by a SPE satisfying the refinement, if and only if* $(b, s) \in [B_0(s), B_\delta(s)] \times [S_0(b), S_\delta(b)]$ *and satisfies* $U_0^i(b, s) \geq 0$*, for* $i = B, S$*. Note that the latter set includes* (b_δ, s_δ) *if* $U_0^i(b_\delta, s_\delta) \geq 0$ *(even for a low* δ*), and it includes the static outcome,* (b_0, s_0)*. The sustainability of the static outcome depends crucially on the substitutability of the investments, though. As shown in Section 6, if the investments are complementary, the static outcome may not be implementable even by a SPE (and not just a MPE satisfying the refinement), for a large* δ*.*

Remark 4 *While investments and trade take place in the first period in our asymptotically efficient equilibria (just 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.*[17] *At the same time, our model's prediction may still differ from the static result even if it is truncated after two periods (c.f. the example in Section 6).*

5 The necessity of individual rationality for asymptotic efficiency when investments are substitutes

In this section, we establish a necessary condition for asymptotic efficiency. Specifically, we will argue that, for (weakly) substitutable investments, any equilibrium pair of cumulative investments is bounded away from the efficient pair for all $\delta \in (0, 1)$, unless we have[18]

$$(IR_\alpha) \qquad U_0^B(b_1, s_1; \alpha) \geq 0 \text{ and } U_0^S(b_1, s_1; \alpha) \geq 0.$$

[17] The asymptotic efficiency result of Marx and Matthews (2000) in the "no-payoff jump" case (which corresponds to the situation considered in our model) also unravels in the finite-horizon setting.

[18] Note this (IR_α) is a weak version of (SIR_α). Recall from footnote 12 that the difference stems from the fact that (IR_α) may not be sufficient for (3).

The necessity of (IR_α) for asymptotic efficiency would be immediate if every equilibrium had the feature, like I_s, that the parties never invest further once they arrive at a target pair, say (b', s'), at which trade occurs. Given this feature, the bargaining shares at the target pair will coincide with $(\alpha, 1-\alpha)$, so the equilibrium payoff for party $i = B, S$ will never exceed $U_0^i(b', s')$.[19] Hence, if (b', s') is sufficiently close to (b_1, s_1), then the equilibrium would be sustainable only if (IR_α) holds. Not all equilibria may have this feature, however. In principle, bargaining shares different from $(\alpha, 1-\alpha)$ may be implementable at the target pair, if some nontrivial investment were to follow (out-of-equilibrium) disagreement at that target pair. In fact, with complementary investments, Section 6 will show that such an investment path exists, so that an efficient pair may be implementable even when (IR_α) fails. Hence, establishing the necessity of (IR_α) for asymptotic efficiency is not trivial even for (weakly) substitutable investments. In particular, it requires identifying all possible investment paths that can be implemented after the target pair is reached.

To proceed formally, we need a few notations. For any $(b, s) \in \mathcal{X} \times \mathcal{Y}$, consider a subgame that follows immediately after (b, s) is reached (but before the proposer is chosen). Let $\overline{w}_\delta^i(b, s)$ and $\underline{w}_\delta^i(b, s)$ then denote respectively the supremum and the infimum SPE continuation payoffs for party $i = B, S$ at that subgame. We then consider their upper and lower envelopes by defining

$$\overline{\sigma}_\delta(b, s) := \limsup_{(b',s') \to (b,s)} \overline{w}_\delta^S(b', s') \text{ and } \underline{\sigma}_\delta(b, s) := \liminf_{(b',s') \to (b,s)} \underline{w}_\delta^S(b', s'),$$

for the seller, and similarly $\overline{\beta}_\delta(b, s)$ and $\underline{\beta}_\delta(b, s)$ for the buyer. We shall suppress the dependence of these functions on δ, unless it becomes relevant. We shall simply call $\overline{\sigma}(b, s)$ and $\underline{\sigma}(b, s)$ the highest and the lowest sustainable continuation payoffs for the seller at (b, s), and similarly for the buyer.

Next, we let $\mathcal{I}(b, s)$ denote the set of all (cumulative) investment pairs that can be reached from (b, s) in any SPE. As was seen in the previous section, this set is nonempty. For our purpose, it is useful to consider its closure, $\bar{\mathcal{I}}(b, s)$.

Consider the highest possible continuation payoff for the seller when starting from a stock of (b', s'):

$$V^S(b', s') := \max_{(b'',s'') \in \bar{\mathcal{I}}(b',s')} \overline{\sigma}(b'', s'') - (s'' - s'). \tag{6}$$

The maximum is well defined since $\overline{\sigma}$ is upper-semicontinuous (see Theorem A6.5 of Ash (1972),

[19] If the parties arrive at (b', s') in the first period, then their payoffs equal $U_0^B(b', s')$ and $U_0^S(b', s')$, precisely. If they reach the pair later, their payoffs fall short of these amounts.

pp. 389-390) and $\bar{\mathcal{I}}(b', s')$ is compact.[20] Consider now its limit superior as $(b', s') \to (b, s)$,

$$\limsup_{(b',s')\to(b,s)} V^S(b', s'), \tag{7}$$

and the sequence of maximizers of (6) that attain this value in the limit, and let $(\hat{x}(b, s), \hat{y}(b, s))$ be a limit point of that sequence, (which is well defined since the maximizers in the sequence lie in the compact set, $\mathcal{X} \times \mathcal{Y}$). For brevity, we will suppress the arguments of $(\hat{x}, \hat{y})$ from now on, unless the arguments are different from (b, s). Likewise, we can similarly define $V^B(b, s)$ for the buyer and a limit point, $(\tilde{x}, \tilde{y})$, of a sequence of maximizers attaining $\limsup_{(b',s')\to(b,s)} V^B(b', s')$.

We now characterize the highest and the lowest sustainable continuation payoffs through Bellman equation type conditions, which will facilitate our analysis.

Lemma 1 *For any $(b, s) \in \mathcal{X} \times \mathcal{Y}$, we must have*

$$\begin{aligned}\overline{\sigma}(b, s) \le (1-\alpha) \max\Big\{&\phi(b, s) - \delta[\min_{s'\in\mathcal{Y}(s)} \underline{\beta}(\hat{x}, s') - (\hat{x} - b)], \delta[\overline{\sigma}(\hat{x}, \hat{y}) - (\hat{y} - s)]\Big\} \\ &+\alpha\delta[\overline{\sigma}(\hat{x}, \hat{y}) - (\hat{y} - s)].\end{aligned} \tag{8}$$

and

$$\begin{aligned}\underline{\beta}(b, s) \quad &\ge \quad \alpha \max\Big\{\phi(b, s) - \delta[\overline{\sigma}(\hat{x}, \hat{y}) - (\hat{y} - s)], \delta[\min_{s'\in\mathcal{Y}(s)} \underline{\beta}(\hat{x}, s') - (\hat{x} - b)]\Big\} \\ &\qquad +(1-\alpha)\delta[\min_{s'\in\mathcal{Y}(s)} \underline{\beta}(\hat{x}, s') - (\hat{x} - b)] \\ &\ge \quad \alpha\phi(b, s) - \alpha\delta[\overline{\sigma}(\hat{x}, \hat{y}) - (\hat{y} - s)] + (1-\alpha)\delta[\min_{s'\in\mathcal{Y}(s)} \underline{\beta}(\hat{x}, s') - (\hat{x} - b)].\end{aligned} \tag{9}$$

A symmetric characterization hold for $\overline{\beta}$ and $\underline{\sigma}$, relative to $(\tilde{x}, \tilde{y})$.

Proof. See Appendix B.

Some intuition can be provided for these conditions, assuming that V^S and V^B are attained by $(\hat{x}, \hat{y})$ and $(\tilde{x}, \tilde{y})$ precisely. (That this assumption is not necessarily valid accounts for much of the proof.) Consider (8) for instance. Once the parties arrive at (b, s), the seller becomes the proposer with probability $1 - \alpha$. In this case, the lowest offer that the buyer would accept cannot be lower than $\delta[\min_{s'\in\mathcal{Y}(s)} \underline{\beta}(\hat{x}, s') - (\hat{x} - b)]$, since the latter is a lower bound for the buyer's minmax value.[21]

---

[20]Likewise, the minima are well defined for $\underline{\sigma}$ and $\underline{\beta}$, which are used in Lemma 1.

[21] The buyer's minmax value is

$$\min_{s'\in\mathcal{Y}(s)} \sup_{b'\in\mathcal{X}(b)} \underline{\beta}(b', s') - (b' - b),$$

which is no less than $\min_{s'\in\mathcal{Y}(s)} \underline{\beta}(\hat{x}, s') - (\hat{x} - b)$, since $\hat{x} \ge b$.

Hence, the highest continuation payoff for the seller cannot exceed $\phi(b,s) - \delta[\min_{s' \in \mathcal{Y}(s)} \underline{\beta}(\hat{x}, s') - (\hat{x} - b)]$ if she wishes to make an acceptable offer. Alternatively, the seller can make a rejectable offer, in which case her highest continuation payoff cannot exceed $V^S(b,s)$, or $\delta[\overline{\sigma}(\hat{x},\hat{y}) - (\hat{y} - s)]$, given our assumption. Clearly, the seller's continuation payoff cannot exceed the bigger of the two payoffs, which explains the first term. With probability α, the seller becomes a responder. In this case, the buyer will never offer more than $V^S(b,s)$ $(= \delta[\overline{\sigma}(\hat{x},\hat{y}) - (\hat{y} - s)])$, so her continuation payoff can never exceed this amount, which explains the second term. In sum, the highest payoff sustainable for the seller at (b,s), $\overline{\sigma}(b,s)$, cannot exceed the RHS of (8). Similar explanations apply to $\underline{\beta}, \overline{\beta}$ and $\underline{\sigma}$.

These conditions can be used to characterize the extreme continuation payoffs for the parties. While a party's highest and lowest continuation payoffs do not coincide in general, they do so for sufficiently large pairs of cumulative investments, leading to unique continuation payoffs in those cases.

Lemma 2 *Assume that the investments are weak substitutes ($\phi_{bs}(b,s) \leq 0$). For any $(b,s) \in \Omega_\delta := \{(b,s) \in \mathcal{X} \times \mathcal{Y} \mid s > S_\delta(b) \text{ and } b > B_\delta(s)\}$, we have, for all $\delta \in [0,1)$,*

$$\overline{\sigma}_\delta(b,s) = \underline{\sigma}_\delta(b,s) = (1-\alpha)\phi(b,s) \text{ and } \overline{\beta}_\delta(b,s) = \underline{\beta}_\delta(b,s) = \alpha\phi(b,s).$$

Proof. See Appendix B.

Lemma 2 has an immediate implication on the implementability of some investment pairs. It can be seen that no pair in Ω_δ, including the first-best pair, is reachable from in any SPE.

Proposition 2 *Given weak substitutable investments ($\phi_{bs}(b,s) \leq 0$), no pair in Ω_δ, including the first-best pair, is implementable in any SPE, for any $\delta \in [0,1)$.*

Proof. Suppose to the contrary that a pair $(b,s) \in \Omega_\delta$ is implementable for some $\delta \in [0,1)$. Then, there must exist $(b',s') < (b,s)$ such that $(b,s) \in \mathcal{I}(b',s')$. Without loss of generality, assume $b' < b$. That $(b,s) \in \mathcal{I}(b',s')$ requires that, for any $b'' \in [b', b)$ with $(b'',s) \in \Omega_\delta$, we must have

$$\overline{\beta}_\delta(b,s) - (b - b') - \left[\underline{\beta}_\delta(b'',s) - (b'' - b')\right] \geq 0.$$

Yet, by Lemma 2, we have

$$\overline{\beta}_\delta(b,s) - (b - b') - \left[\underline{\beta}_\delta(b'',s) - (b'' - b')\right] = U_0^B(b,s) - U_0^B(b'',s) < 0,$$

where the inequality follows from $b'' < b$, $b'' \geq B_\delta(s) \geq B_0(s)$, and from the concavity of U_0^B. Hence, we have obtained a contradiction. ▌

Remark 5 *Lemma 2 can be proven in the case of discrete investments, by a similar, and in fact more straightforward, argument. In the discrete case, the first-best pair is on the boundary of Ω_δ, but not in Ω_δ, for a large value of δ, so Proposition 2 does not apply. If (IR_α) fails, however, the efficient investment pair can never be implemented in the discrete investment case.*

The proposition implies that the parties' discounting of the future is a clear cause of inefficiencies. The proposition does not preclude asymptotic efficiency, however. As was shown in Proposition 1, given (SIR_α), one can find an equilibrium that implements the first-best arbitrarily closely as $\delta \to 1$. To show that (IR_α) is indeed *necessary* for asymptotic efficiency, the following lemma proves useful.

Lemma 3 *Given weakly substitutable investments ($\phi_{bs}(b,s) \leq 0$),*

$$\limsup_{(b,s)\to(b_1,s_1)} \sup_{\delta\in[0,1)} \overline{\sigma}_\delta(b,s) \leq (1-\alpha)\phi(b_1,s_1) \text{ and } \limsup_{(b,s)\to(b_1,s_1)} \sup_{\delta\in[0,1)} \overline{\beta}_\delta(b,s) \leq \alpha\phi(b_1,s_1),$$

Proof. See Appendix B.

Finally, we are in the position to state and prove the main result of this section.

Proposition 3 *Assume that the investments are weak substitutes ($\phi_{bs}(b,s) \leq 0$). If (IR_α) fails, then there exists an open set, $\mathcal{O}$, containing (b_1,s_1), such that any investment pair in $\mathcal{O}$ can never be implemented in any SPE, for any $\delta \in [0,1)$.*

Proof. Suppose that (IR_α) fails. Without loss of generality, suppose $\alpha\phi(b_1,s_1) - b_1 < 0$. Choose $\epsilon > 0$ such that $\alpha\phi(b_1,s_1) + \epsilon - b_1 < 0$. Then, by Lemma 3, there exists ν-ball, B_ν, with a center at (b_1,s_1) and a radius $\nu > 0$, such that for all $(b,s) \in B_\nu$, $\sup_{\delta\in[0,1)} \overline{\beta}_\delta(b,s) \leq \alpha\phi(b_1,s_1) + \epsilon$. This means that if trade occurs at any $(b,s) \in B_\nu$, then the buyer will obtain strictly negative payoff (no matter when the trade occurs). Since the buyer's minmax value is zero, this cannot occur in equilibrium. Hence, there exists no equilibrium in which the parties reach $(b,s) \in B_\nu$. ▌

This result suggests that inefficiencies are unavoidable when the individual rationality constraint fails. This confirms that, in some cases, the holdup problem provides some rationale for organizational remedies, although the scope of the circumstances is likely to be much narrower than has been recognized, and the nature of remedies warranted may be quite different from those proposed in the existing literature. We come back to this point in Conclusion.

6 Complementary investments: an example

Proposition 1 shows that (SIR_α) is sufficient for asymptotic efficiency, for both substitutable and complementary investments. The necessity of (IR_α) for asymptotic efficiency was however established in Proposition 3 only for weakly substitutable investments. In this section, we show via an example that the necessity of (IR_α) does not extend to complementary investments: i.e., the failure of (IR_α) need not imply unattainability of the first-best outcome if the investments are complementary. The same example will be used to make another interesting point: with complementarity, the static equilibrium may not be implementable for a large δ, even when the game must end by period two.

Consider two parties, S and B, with equal bargaining power ($\alpha = \frac{1}{2}$). S ("row" player) has the option of investing nothing or 20. B ("column"player) can invest 0, 10 or 20. The gross surplus, ϕ, depends on their investment decisions in the following way:

$S \setminus B$	*0*	*10*	*20*
0	**2**	**18**	**18**
20	**2**	**24**	**46**

It is useful first to describe the parties' payoffs in the static game (i.e., with $\delta = 0$).

$S \setminus B$	*0*	*10*	*20*
0	$(1, 1)$	$(9, -1)$	$(9, -11)$
20	$(-19, 1)$	$(-8, 2)$	$(3, 3)$

Observe first that the only equilibrium is $(0, 0)$ in the static game. Next, the first-best pair is $(0, 10)$ (at which the joint payoffs are maximized). We then make the following two observations.

- *The static equilibrium pair,* $(0, 0)$*, is not sustainable in any SPE for* $\delta > \frac{22}{23}$.

To see this, suppose to the contrary that the static equilibrium pair $(0, 0)$ is implementable. In any such equilibrium, S's payoff is at most 2, since the outcome must be individually rational for B. Suppose now S deviates to 20 unilaterally in the first period. With $\delta > \frac{22}{23}$, the parties will then disagree (regardless of who becomes the proposer) and let B invest 20 in the following period, which is indeed his best response.[22] Hence, the deviation would give S a net payoff of $\delta 23 - 20$ which exceeds 2 if $\delta > \frac{22}{23}$. Thus, $(0, 0)$ is not implementable in any SPE with $\delta > \frac{22}{23}$. It is worth

[22]By choosing 0 and 10, B can net at most 1 and 2, respectively.

noting that the above argument does not depend on the assumption that our game has infinite horizon. It applies even in a two-period truncation of our model.

This observation again reinforces the fact that the folk theorem does not apply in our model. It also implies that, despite the multiplicity, the predictions of the static and dynamic models *must* sometimes differ. That is, in general, we cannot apply some special refinement in the dynamic game and get back the static result.

- *The failure of* (IR_α) *need not imply that efficiency is unattainable.*

As can be seen in the payoff matrix, (IR_α) does not hold in this example with $\alpha = \frac{1}{2}$: At the first-best pair, $(0, 10)$, B realizes -1 from splitting the gross surplus equally. *In our dynamic game, however, the first-best pair can be implemented in an MPE if* $\delta \geq \frac{3}{5}$. To see this, consider the investment strategies:

$$I^*(b, s) = \begin{cases} (0, 10) & \text{if } (b, s) = (0, 0), \\ (20, 20) & \text{if } (b, s) \neq (0, 0). \end{cases}$$

With this strategy profile, the parties initially move to the first best pair $(0, 10)$, and, if no agreement is reached, they move to $(20, 20)$. Since $(0, 10)$ is the first best, they will never disagree after reaching $(0, 10)$, so the move to $(20, 20)$ never takes place on the equilibrium path. Yet, this out-of-equilibrium move serves to manipulate the bargaining share at $(0, 10)$ in favor of B. Once the parties reach $(20, 20)$, they will split the gross surplus of 46 equally (on average). Hence, the continuation payoffs from rejecting at the first-best pair $(0, 10)$ are $\delta 3 = \delta(23 - 20)$ and $\delta 13 = \delta(23 - 10)$ for S and B, respectively. (That the additional cost to be incurred is smaller for B increases his continuation payoff.) Hence, the optimal bargaining behavior at $(0, 10)$ is for B to offer $\delta 3$, when becoming the proposer, and accept S's offer, when becoming a responder, if and only if it is no less than $\delta 13$, and similarly for S. Given this bargaining behavior, S an B's net payoffs from reaching $(0, 10)$ are $9 - \delta 5$ and $-1 + \delta 5$, respectively.[23] With $\delta \geq \frac{3}{5}$, the parties' continuation payoffs at each pair, given I^* and the associated optimal bargaining strategies, are described as follows:

[23] B receives $18 - \delta 3$ and $\delta 13$ with equal probability, so his expected payoff at $(0, 10)$ is $\frac{1}{2}(18 - \delta 3) + \frac{1}{2}\delta 13 = 9 + \delta 5$. Subtracting from this B's investment cost of 10 gives his net payoff, $-1 + \delta 5$. S receives the remainder of the net social surplus of 8, $9 - \delta 5$.

$S \setminus B$	*0*	*10*	*20*
0	$(\delta(9-\delta 5), \delta(-1+\delta 5))$	$(9-\delta 5, -1+\delta 5)$	$((9-\delta 10) \vee \delta 3, (-11+\delta 10) \vee (\delta 23 - 20))$
20	$(\delta 23 - 20, \delta 3)$	$(\delta 23 - 20, \delta 13 - 10)$	$(3,3)$

One can easily check that the suggested investment strategies constitute an MPE if $\delta \geq \frac{3}{5}$.[24] In particular, it is mutual best response for the parties to choose the first-best pair $(0,10)$ and trade immediately. Most importantly, the new equilibrium satisfies individual rationality at $(0,10)$ by shifting the S and B's bargaining shares from $(\frac{1}{2}, \frac{1}{2})$ to $(\frac{9-\delta 5}{18}, \frac{9+\delta 5}{18})$. In fact, there is a transfer of $\delta 5$ from S to B. Such a transfer is created by the out-of-equilibrium investment move from $(0,10)$ to $(20,20)$. This move favors B's bargaining position at the expense of S's, since the move imposes a higher investment cost to S than to B, thus making rejection more unattractive for S than for B.

7 Related literature

Our model and results are related to several branches of the literature. First, our model can be seen as Rubinstein's bargaining model (1982) (more precisely, Binmore (1987)'s random proposer variant) in which the pie may grow endogenously through investments. While some insight from the pure bargaining model carries over to our model, the interaction of investment and bargaining has presented a new problem. For instance, a continuation payoff (upon reaching some cumulative investment pair, say) is no longer unique, but rather depends on the subsequent investment path, which must be in turn incentive compatible relative to the continuation payoffs. It turns out that a range of different bargaining shares are sustainable by different (self-enforcing) investment paths. In a similar vein, Busch and Wen (1995) support multiple bargaining shares in a 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. By contrast, the inefficiency and multiplicity of equilibria in our model follows from the endogeneity of the bargaining stake.

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 (2001), Lockwood and Thomas (2002), Admati and Perry (1991), Bagnoli and Lipman (1989) and Pitchford and

[24]The discount factor needs to be greater than $\frac{3}{5}$ to sustain the out-of-equilibrium move, $I^*(0,20) = (20,20)$. The associated bargaining behavior is analogous to that for $(0,10)$. Given $\delta \geq \frac{3}{5}$, disagreement/delay arises upon arriving at $(0,0)$, $(20,0)$, or at $(20,10)$. Arrival at $(0,10)$ is followed by an immediate agreement and trade. Arriving at $(0,20)$ results in delay if and only if $\delta \geq \frac{9}{13}$.

Snyder, 2001)).[25] Marx and Matthews (2000) and Lockwood and Thomas (2002) 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.[26] This difference in environment implies that the gradual accumulation strategies proposed by the contribution game literature would be unsustainable in our game.[27] In fact, investments take place all at once in our asymptotically efficient equilibria. As a consequence of not having to rely on gradual investment strategies, our results are equally valid for a discrete set of feasible investments, unlike the above mentioned papers. Finally, the contribution models consider investments that are

[25] Gale (2001) studies monotone games with positive spill-over, while Lockwood and Thomas (2002) consider repeated games with irreversible actions. Both include contribution games as a special case. Pitchford and Snyder (2001) is particularly relevant since it involves a transfer payment and interprets the gradual investment as a solution to the holdup problem. Except for this feature, their model is isomorphic to Marx and Matthews (2000), for instance.

[26] 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.

[27] 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 (2002)). 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 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.

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.

Gul (2001) also establishes asymptotic efficiency in a holdup model without ex ante contracts. In that model, a seller makes repeated one-sided offers to a buyer who has made unobservable investment. Unlike in our model, though, investment dynamics are not allowed since all investment must take place in the first period. Rather, the unobservablity plays a key role:[28] The required self selection constraint means that the price facing the buyer is independent of his investment choice. This feature makes the buyer a residual claimant on his investment whenever he is induced to purchase. Hence, efficiency would result if the equilibrium trade is efficient, which in Gul (2001) results from the Coasian effect as the time interval between offers shrinks. By contrast, our result works with observable investments and rests on the investment dynamics. Further, the asymptotic efficiency in Gul (2001) works only when the investment is selfish (in the sense of benefiting purely the investor), whereas our result requires no restriction on the nature of investments.

8 Conclusion

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 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 for simple, realistic investment dynamics, the parties' sharing of trade 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.

As mentioned in the introduction, our positive findings suggest that explicit contracts are un-

[28] If the investment were observable, then the seller would extract the entire surplus, leading to no investment on the part of the buyer in equilibrium.

necessary in many plausible situations. This finding is consistent with similar results put forth by Che and Hausch (1999), Segal (1999) and Hart and Moore (1999). These papers have shown 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. Remarkably, the features that make contracts ineffective in the aforementioned papers (e.g., "cooperative investments" and/or environmental "complexities") do not disrupt efficiency here even in the absence of *ex ante* contracts!

At the same time, our finding of inefficiencies suggests that the holdup problem may provide a rationale for contractual and organizational remedies, in some circumstances. Even in these circumstances, the required remedies appear to be different from those prescribed by the existing literature. In particular, our finding suggests that individual rationality should be an important consideration in contract design. This insight will likely influence our views on how to allocate asset ownership rights and to design production contracts and how to allocate parties' default rights in disputes through the design of legal rules and institutions. Indeed, our preliminary results indicate that some of the well-known existing prescriptions do not apply in an environment where investment dynamics are allowed.[29] Such an inquiry appears to offer a promising new avenue for developing organization theory.

[29]Specifically, Che and Sákovics (2003) find that the separate ownership of complementary assets can be optimal; exclusivity agreements can promote specific investments even when their values are not transferable to outside parties; and contracts can promote cooperative investments, which contrast with some of the standard prescriptions known in the literature (e.g., Hart (1995), Segal and Whinston (2000), and Che and Hausch (1999)).

References

[1] Asanuma, B. (1989). "Manufacturer-Supplier Relationships in Japan and the Concept of Relation-Specific Skill." *Journal of the Japanese and International Economies*, 3, 1-30.

[2] Acemoglu, D. and R. Shimer (1999): "Holdups and efficiency with search frictions," *International Economic Review*, 40(4), 827-849.

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

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

[5] Aghion, P., Dewatripont, M., and P. Rey (1994): "Renegotiation Design with Unverifiable Information," *Econometrica*, 62, 257-282.

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

[7] Ash, R. (1972): *Real Analysis and Probability*, Academic Press, Inc.

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

[9] Baker, G., Gibbons, R., and K.J. Murphy (2002): "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics,* 117, 39-83.

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

[11] Burt, D. (1989). "Managing Suppliers up to Speed." *Harvard Business Review*, 67, 127-135.

[12] Busch L.-A. and Q. Wen (1995): "Perfect equilibria in a negotiation model," *Econometrica*, 63(3), 545-565.

[13] Che, Y.-K., and I. Gale (2003): "Optimal Design of Research Contests," *American Economic Review*, forthcoming.

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

[15] Che, Y.-K., and J. Sákovics (2001): "A Dynamic Theory of Holdup," *SSRI Working Paper* 2001-25 (http://www.ssc.wisc.edu/econ/archive/wp2001-25.pdf), University of Wisconsin; and *Economics Dicussion Paper* 01-05, University of Edinburgh.

[16] Che, Y.-K., and J. Sákovics (2003): "Contractual Remedies to the Holdup Problem in Dynamic Relationships," mimeo., University of Wisconsin.

[17] Chung, T.-Y. (1991): "Incomplete Contracts, Specific Investment and Risk Sharing," *Review of Economic Studies*, 58, 1031-1042.

[18] Cole, H.L., Mailath G.J. and A. Postlewaite (2001): "Efficient non-contractible investments in large economies," *Journal of Economic Theory*, 101(2), 333-373.

[19] Dewatripont, M., Legros, P., and Matthews, S. (2002): "Moral Hazard and Capital Structure Dynamics," mimeo.

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

[21] Dutta, P. (1995): "A Folk Theorem for Stochastic Games," *Journal of Economic Theory,* 66, 1-32.

[22] Dyer, J. and Ouchi, W. (1993): "Japanese-Style Partnerships: Giving Companies a Competitive Edge." *Sloan Management Review*, 51-63.

[23] Edlin, A.S. and S. Reichelstein (1996): "Holdups, Standard Breach remedies and Optimal Investment," *American Economic Review*, 86, 478-501.

[24] *The Engineer.* (1996): "An Industry that is Good in Parts." October 24.

[25] Felli, L., and K. Roberts (2000): "Does Competition Solve the Hold-up Problem?," mimeo.

[26] Gale, D. (2001): "Monotone Games with Positive Spillovers," *Games and Economic Behavior*, 37, 295-320.

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

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

[29] Gul, F. (2001): "Unobservable investment and the hold-up problem," *Econometrica*, 69, 343-376.

[30] Halonen, M. (2002): "Reputation and the Allocation of Ownership" *Economic Journal,* 112, 539-558.

[31] Hart, O.D. (1995). *Firms, Contracts, and Financial Structure*, Clarendon Press; Oxford.

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

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

[34] Holmstrom, B. (1982). "Moral Hazard in Teams," *Bell Journal of Economics,* 13, 324-340.

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

[36] Lockwood, B., and J.P. Thomas (2002): "Gradualism and Irreversibility," *Review of Economic Studies*, 69(2), 339-356.

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

[38] MacLeod, W.B. and J. Malcomson (1989): "Implicit contracts, incentive compatibility and involuntary unemployment" *Econometrica,* 57, 447-480.

[39] MacLeod, W.B. and J. Malcomson (1993): "Investments, Holdup, and the Form of Market Contracts," *American Economic Review*, 83, 811-837.

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

[41] Maskin, E. and Tirole, J. (1999), Unforeseen Contingencies and Incomplete Contracts," *Review of Economic Studies*, 66, 83-114.

[42] Nöldeke, G., and K. M. Schmidt (1995): "Option Contracts and Renegotiation: A Solution to the Hold-Up Problem," *Rand Journal of Economics,* 26, 163-179.

[43] Nöldeke, G., and K.M. Schmidt (1998): "Sequential Investment and Options to Own," *Rand Journal of Economics,* 29(4), 633-653.

[44] Osborne, M.J., and Rubinstein, A. (1990): *Bargaining and Markets*, Academic Press: London.

[45] Pitchford, R., and C. Snyder (2001): "A Non-Contractual Solution to the Hold-Up Problem Involving Gradual Investment," mimeo, George Washington University.

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

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

[48] Segal, I.R., and Whinston, M., (2000): "Exclusive Contracts and Protection of Investments," *Rand Journal of Economics*, 31, 603-633.

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

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

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

Appendix A: Proof of Proposition 1

A.1 Weakly substitutable investments ($\phi_{bs} \leq 0$)

We first prove the proposition for the case of weakly substitutable investments.

Proposition A1 *Given* (SIR_α) *and* $\phi_{bs} \leq 0$, *there exists a* $\delta^* < 1$ *such that, for all* $\delta \geq \delta^*$, *there exists a MPE in which the parties choose* (b_δ, s_δ) *and trade in the first period.*

Proof. Given (SIR_α), there exists $\hat{\delta}^* < 1$ such that, for all $\delta \geq \hat{\delta}^*$, we have

$$U_0^B(b_\delta, s_\delta) \geq 0 \text{ and } U_0^S(b_\delta, s_\delta) \geq 0. \tag{10}$$

Fix any such δ. Consider now the Markovian investment strategies, $I_s(\cdot,\cdot)$, and the associated bargaining strategies described in the text. We prove that the strategies form a subgame perfect equilibrium starting with any $(b, s) \in \mathcal{X} \times \mathcal{Y}$.

The proof generalizes the argument sketched in the text. If no party deviates, then the buyer and the seller would respectively choose $(x(b, s), y(b, s)) := I_s(b, s)$, which lies in region (iv) (including its boundary). Since $I_s(x(b, s), y(b, s)) = (x(b, s), y(b, s))$ (i.e., no further investment is prescribed following the equilibrium choice), the game becomes a pure bargaining game. Applying the one period deviation principle, the buyer's payoff in equilibrium is

$$U_0^B(x(b, s), y(b, s)) + b \tag{11}$$

(excluding the sunk portion, b). Suppose now that that buyer deviates to $b' > x(b, s)$. Since $B_\delta(\cdot)$ is nonincreasing, such a deviation will leave the parties in region (iv). Consequently, no further investment is prescribed following the deviation. Hence, the buyer's deviation payoff is

$$U_0^B(b', y(b, s)) + b.$$

We have $x(b, s) \geq B_\delta(y(b, s))$ since $(x(b, s), y(b, s))$ is in region (iv). Then,

$$U_0(x(b, s), y(b, s)) \geq U_0(b', y(b, s)),$$

since $b' \geq x(b, s) \geq B_\delta(y(b, s)) \geq B_0(y(b, s))$. Hence, the deviation is not profitable.

Suppose next that the buyer deviates to $b'' \in [b, x(b, s))$. The fact that $b < x(b, s)$ means that (b, s) can only be in region (i) or region (iii). Hence, it can be easily see from Figure 1(a) that $y(b, s) = \max\{s, s_\delta\}$ and $x(b, s) = B_\delta(y(b, s))$. Since $I_s(b'', y(b, s)) = (x(b, s), y(b, s))$ and since $I_s(x(b,s), y(b,s)) = (x(b,s), y(b,s))$, the buyer's deviation payoff is calculated, as explained in footnote 13, as

$$\max\{U_\delta^B(b'', y(b,s)) - \delta(1-\alpha)x(b,s) + b, \delta[U_0^B(x(b,s), y(b,s)) + b''] - (b'' - b)\}. \tag{12}$$

Observe first that

$$U_0^B(x(b,s), y(b,s)) + b'' \geq U_0^B(b_\delta, y(b,s)) + b'' \geq U_0^B(b_\delta, s_\delta) + b'' \geq 0, \tag{13}$$

where the first inequality holds since $x(b,s) = B_\delta(y(b,s)) \leq B_\delta(s_\delta) = b_\delta$, $B_\delta(y(b,s)) \geq B_0(y(b,s))$ and $U_0^B(\tilde{b}, y(b,s))$ is nonincreasing in $\tilde{b}$ for $\tilde{b} \geq B_0(y(b,s))$, the second inequality holds since $y(b,s) \geq s_\delta$ and $U_0^B(b_\delta, \cdot)$ is nondecreasing, and the last inequality follows from (10). Given (13), the second term of (12) can never exceed the equilibrium payoff in (11). Also, note that the first term of (12) attains its maximum at $x(b,s) = B_\delta(y(b,s))$ and equals $U_0^B(x(b,s), y(b,s)) + b$, so it can never exceed the equilibrium payoff.

The proof of the optimality of the seller's strategy is completely symmetric. ▌

A.2 Complementary investments ($\phi_{bs} \geq 0$)

We begin with the construction of the following investment strategies:

$$I_c(b,s) = \begin{cases} (b_\delta, s_\delta) & \text{if } (b,s) \leq (b_\delta, s_\delta), \quad [\textit{region (i)}] \\ (b, \hat{S}_\delta(b)) & \text{if } b > b_\delta \text{ and } s \leq \hat{S}_\delta(b), \quad [\textit{region (ii)}] \\ (\hat{B}_\delta(s), s) & \text{if } s > s_\delta \text{ and } b \leq \hat{B}_\delta(s), \quad [\textit{region (iii)}] \\ (b,s) & \text{if } b \geq \hat{B}_\delta(s) \text{ and } s \geq \hat{S}_\delta(s), \quad [\textit{region (iv)}] \end{cases}$$

where $\hat{B}_\delta : \mathcal{Y}(s_\delta) \mapsto \mathcal{X}$ satisfies

$$U_0^S(\hat{B}_\delta(s), s) = \pi^S(s) := \max_{s' \in \mathcal{Y}(s)} U_0^S(\max\{b_\delta, B_0(s')\}, s'),$$

for each $s \in \mathcal{Y}(s_\delta)$, and $\hat{S}_\delta : \mathcal{X}(b_\delta) \mapsto \mathcal{Y}$ satisfies

$$U_0^B(b, \hat{S}_\delta(b)) = \pi^B(b) := \max_{b' \in \mathcal{X}(b)} U_0^B(b', \max\{s_\delta, S_0(b')\}),$$

for each $b \in \mathcal{X}(b_\delta)$.[30] In words, $\hat{B}_\delta(s)$ is the buyer's investment response that keeps the seller's static payoff $U_0^S(\hat{B}_\delta(s), s)$ from increasing as s rises. Requiring the buyer to choose $\hat{B}_\delta(s)$ will be

[30] These functions are well defined. For instance, since $U_0^B(b, \max\{s_\delta, S_0(\bar{b})\}) \geq \pi^B(b) \geq U_0^B(b, S_0(b))$ and since $U_0^B(b, \cdot)$ is continuous, there exists s' such that $U_0^B(b, s') = \pi^B(b)$.

shown later to control the seller's incentive to overinvest relative to the target, and similarly for the buyer. The phase diagram of I_c is depicted in Figure 2. According to the figure, the functions, $\hat{B}_\delta(\cdot)$ and $\hat{S}_\delta(\cdot)$, pass through (b_δ, s_δ), are nondecreasing and lie between $B_0(\cdot)$ and $B_\delta(\cdot)$, and between $S_0(\cdot)$ and $S_\delta(\cdot)$, respectively, which we establish through the next two lemmas.

Lemma A1 *Assume $\phi_{bs} \geq 0$. There exists $\hat{\delta} < 1$ such that for any $\delta \in [\hat{\delta}, 1)$, $U_0^B(b_\delta, s_\delta) = \pi^B(b_\delta)$ and $U_0^S(b_\delta, s_\delta) = \pi^S(s_\delta)$, and, for some $\nu > 0$, $U_0^B(b, s_\delta) = \pi^B(b)$ for all $b \in [b_\delta, b_\delta + \nu]$ and $U_0^B(b_\delta, s) = \pi^S(s)$ for all $s \in [s_\delta, s_\delta + \nu]$.*

Proof. There exists $\epsilon > 0$ and $\hat{\delta}_0 < 1$ such that $S_0(b_\delta + \epsilon) < s_\delta$ for all $\delta \in (\hat{\delta}_0, 1]$.[31] Fix any such $\epsilon > 0$ and $\delta \geq \hat{\delta}_0$, and let

$$\hat{b}_\delta \in \arg \max_{b' \in \mathcal{X}(b_\delta + \epsilon)} U_0^B(b', \max\{s_\delta, S_0(b')\}).$$

Since $(b_\delta, s_\delta) \to (b_1, s_1)$ as $\delta \to 1$, there exists $\hat{\delta}_1 < 1$ such that, for all $\delta \geq \hat{\delta}_1$, we have

$$\phi(b_\delta, s_\delta) - b_\delta - s_\delta \geq \phi(\hat{b}_\delta, \max\{s_\delta, S_0(\hat{b}_\delta)\}) - \hat{b}_\delta - \max\{s_\delta, S_0(\hat{b}_\delta)\}. \tag{14}$$

Meanwhile, since $\hat{b}_\delta > b_\delta$,

$$U_0^S(b_\delta, s_\delta) < U_0^S(\hat{b}_\delta, s_\delta) \leq U_0^S(\hat{b}_\delta, \max\{s_\delta, S_0(\hat{b}_\delta)\}). \tag{15}$$

It follows that, for any $b' \in \mathcal{X}(b_\delta + \epsilon)$,

$$U_0^B(b_\delta, s_\delta) > U_0^B(\hat{b}_\delta, \max\{s_\delta, S_0(\hat{b}_\delta)\}) \geq U_0^B(b', \max\{s_\delta, S_0(b')\}), \tag{16}$$

where the first inequality is obtained by subtracting (15) from (14), and the second follows from the definition of $\hat{b}_\delta$.

Consider next any $b' \in [b_\delta, b_\delta + \epsilon)$. Since $S_0(b') < s_\delta$ for any such b', $U_0^B(b', \max\{s_\delta, S_0(b')\}) = U_0^B(b', s_\delta)$, which is strictly decreasing in b'. Hence, $U_0^B(b_\delta, s_\delta) \geq U_0^B(b', \max\{s_\delta, S_0(b')\})$ for all $b' \in [b_\delta, b_\delta + \epsilon)$. Combining this with (16), we conclude that $U_0^B(b_\delta, s_\delta) = \pi^B(b_\delta)$ for $\delta \geq \hat{\delta}$ for $\hat{\delta} \equiv \max\{\hat{\delta}_0, \hat{\delta}_1\} < 1$. Furthermore, for such a δ, $U_0^B(b, s_\delta) = \pi^B(b)$ for $b \in [b_\delta, b_\delta + \nu]$, for some $\nu > 0$, because of the continuity of $U_0^B(\cdot, s_\delta)$ and because of (16). Completely symmetric arguments hold for the rest of the lemma. ▌

Lemma A2 *Assume $\phi_{bs} \geq 0$. There exists $\tilde{\delta} < 1$ such that for all $\delta \geq \tilde{\delta}$, $\hat{S}_\delta(b) \in [S_0(b), S_\delta(b)]$ for any $b \geq b_\delta$, $\hat{B}_\delta(s) \in [B_0(b), B_\delta(b)]$ for any $s \geq s_\delta$, and $\hat{S}_\delta(\cdot)$ and $\hat{B}_\delta(\cdot)$ are nondecreasing.*

[31] Clearly, such ϵ is well defined since $S_0(b_\delta) < S_\delta(b_\delta) = s_\delta$ for all $\delta > 0$.

Proof. Without loss of generality, we focus on $\hat{S}_\delta$. First, it easily follows that $\hat{S}_\delta(b) \geq S_0(b)$ for any $b \in \mathcal{X}(b_\delta)$, since

$$U_0^B(b, \hat{S}_\delta(b)) = \pi^B(b) \geq U_0^B(b, \max\{s_\delta, S_0(b)\}).$$

Next, we show that $\hat{S}_\delta(b) \leq S_\delta(b)$ for any $b \in \mathcal{X}(b_\delta)$. First, since $(b_\delta, s_\delta) \to (b_1, s_1)$ as $\delta \to 1$, for any $\epsilon > 0$, there exists $\hat{\delta}(\epsilon) < 1$ such that, for all $\delta \geq \hat{\delta}(\epsilon)$, $\phi(b, S_\delta(b)) - b - S_\delta(b)$ is strictly decreasing in b for $b \in \mathcal{X}(b_\delta + \epsilon)$, and likewise, $\phi(B_\delta(s), s) - B_\delta(s) - s$ is strictly decreasing in s for $s \in \mathcal{Y}(s_\delta + \epsilon)$.

Take $\delta \geq \tilde{\delta} := \max\{\hat{\delta}(\nu), \hat{\delta}\}$, where $\hat{\delta}$ is defined in Lemma A1 with $\nu > 0$ chosen so that $U_0^B(b', s_\delta) = \pi^B(b')$ for $b' \in [b_\delta, b_\delta + \nu]$. Suppose to the contrary that $\hat{S}_\delta(b) > S_\delta(b) \geq \max\{s_\delta, S_0(b)\}$. It must be that $b > b_\delta + \nu$ and that $s_\delta < S_0(b)$. Let $\hat{b} > b$ be such that $\pi^B(b) = U_0^B(\hat{b}, S_0(\hat{b}))$. Clearly, $\hat{S}_\delta(\hat{b}) = S_0(\hat{b}) \leq S_\delta(\hat{b})$. Since $\hat{S}_\delta(\cdot)$ is continuous, there exists $b' \in (b, \hat{b}]$ such that $\hat{S}_\delta(b') = S_\delta(b')$. Since $b' \in (b, \hat{b}]$, we have

$$U_0^B(b', S_\delta(b')) = U_0^B(b', \hat{S}_\delta(b')) = \pi^B(b) = U_0^B(\hat{b}, S_0(\hat{b})). \tag{17}$$

Meanwhile,

$$U_0^S(b', S_\delta(b')) \leq U_0^S(\hat{b}, S_\delta(b')) \leq U_0^S(\hat{b}, S_0(\hat{b})). \tag{18}$$

Summing (17) and (18) yields

$$\phi(b', S_\delta(b')) - b' - S_\delta(b') \leq \phi(\hat{b}, S_0(\hat{b})) - \hat{b} - S_0(\hat{b}). \tag{19}$$

We know, however, that

$$\phi(b', S_\delta(b')) - b' - S_\delta(b') > \phi(\hat{b}, S_\delta(\hat{b})) - \hat{b} - S_\delta(\hat{b}) \geq \phi(\hat{b}, S_0(\hat{b})) - \hat{b} - S_0(\hat{b}),$$

since $\delta \geq \max\{\hat{\delta}(\nu), \hat{\delta}\}$ and $b' \in (b_\delta + \nu, \hat{b})$. Hence, we have a contradiction.

Last, we show that $\hat{S}_\delta(b') \geq S_0(b)$ for any $b, b' \in \mathcal{X}(b_\delta)$ with $b' > b$, if $\delta \geq \tilde{\delta}$. Let $b_* \in \arg\max_{b'' \in \mathcal{X}(b)} U_0^B(b'', \max\{s_\delta, S_0(b'')\})$. There are two possibilities. Suppose first that $b_* \geq b'$. Then,

$$U_0^B(b, \hat{S}_\delta(b)) = U_0^B(b', \hat{S}_\delta(b')) = U_0^B(b_*, \max\{s_\delta, S_0(b_*)\}).$$

Since $\hat{S}_\delta \in [S_0, S_\delta]$ and $b \geq b_\delta$, we have $b \geq B_0(\hat{S}_\delta(b))$. Hence, if $\hat{S}_\delta(b') < \hat{S}_\delta(b)$, $U_0^B(b, \hat{S}_\delta(b)) > U_0^B(b', \hat{S}_\delta(b)) > U_0^B(b', \hat{S}_\delta(b'))$, a contradiction. This proves that $\hat{S}_\delta(b') \geq \hat{S}_\delta(b)$. Suppose next that $b_* \in [b, b')$. The proof follows since

$$\hat{S}_\delta(b) \leq \hat{S}_\delta(b_*) = \max\{s_\delta, S_0(b_*)\} \leq \max\{s_\delta, S_0(b')\} \leq \hat{S}_\delta(b'),$$

where the first inequality follows from the argument in the preceding case and the last inequality follows from the definition of $\hat{S}_\delta$. ▌

Armed with these observations, we are now in a position to prove that I_c forms an subgame perfect equilibrium for sufficiently large δ. We do so first with condition (IR^*): *There exists* $\hat{\delta}_{IR^*} < 1$ *such that for all* $\delta \geq \hat{\delta}_{IR^*}$,

$$U_0^B(\hat{B}_\delta(s), s) \geq 0 \, \forall s \in \mathcal{Y}(s_\delta); \text{ and } U_0^S(b, \hat{S}_\delta(b)) \geq 0 \, \forall b \in \mathcal{X}(b_\delta).$$

Although (SIR_α) does not always imply (IR^*), it is expositionally easier to establish the result first with the latter condition and modify I_c later to accommodate the possibly weaker condition (SIR_α).

Proposition A2 *Given* (IR^*) *and* $\phi_{bs} \geq 0$*, there exists a* $\hat{\delta}^* < 1$ *such that, for all* $\delta \geq \hat{\delta}^*$*, it is MPE behavior for the parties to follow* I_c*, in which case the parties choose* (b_δ, s_δ) *and trade in the first period .*

Proof. Let $\hat{\delta}^* := \max\{\hat{\delta}_{IR^*}, \tilde{\delta}\}$ (where $\hat{\delta}_{IR^*}$ is defined in (IR^*) and $\tilde{\delta}$ is defined in Lemma A2), and fix any $\delta \geq \hat{\delta}^*$. Consider I_c and the associated optimal bargaining strategies, and define $(x(b,s), y(b,s)) := I_c(b,s)$ for any $(b,s) \in \mathcal{X} \times \mathcal{Y}$. We prove that, starting any $(b,s) \in \mathcal{X} \times \mathcal{Y}$, there is no profitable deviation from the equilibrium response, $(x(b,s), y(b,s))$.

Observe that $(x(b,s), y(b,s))$ must be in region (iv) (which includes its boundary). Since $I_c(x(b,s), y(b,s)) = (x(b,s), y(b,s))$ in that region, the buyer earns net payoff of

$$U_0^B(x(b,s), y(b,s)) + b, \tag{20}$$

if he does not deviate.

Suppose first that the buyer deviates to $b'' \in [b, x(b,s))$. As in the proof of Proposition A1, $b < x(b,s)$ means that (b,s) is in either region (i) or region (iii), which in turn implies that $y(b,s) \geq s_\delta$ and that $x(b,s) = \hat{B}_\delta(y(b,s))$. It also follows that $x(b'', y(b,s)) = x(b,s)$ and $y(b'', y(b,s)) = y(b,s)$. Using the one-period deviation principle, the deviation payoff is then

$$\max\{U_\delta^B(b'', y(b,s)) - \delta(1-\alpha)x(b,s) + b, \delta[U_0^B(x(b,s), y(b,s)) + b''] - (b'' - b)\}. \tag{21}$$

Since $x(b,s) = \hat{B}_\delta(y(b,s))$, (IR^*) implies that $U_0^B(x(b,s), y(b,s)) + b \geq 0$. Hence, the second term of (21) cannot exceed the payoff in (20). The first term also cannot exceed the payoff in (20), since $U_\delta^B(b'', y(b,s))$ is increasing in b'' for $b'' \leq x(b,s) = \hat{B}_\delta(y(b,s)) \leq B_\delta(y(b,s))$ (where the last inequality follows from Lemma A2), and attains the payoff of (20) at $b'' = x(b,s)$.

Suppose next that the buyer deviates to $b' > x(b,s)$. There are two possibilities: either $(b', y(b,s))$ lies in region (iv), or it lies in region (ii). Consider the former possibility first. In this case, $b' > \hat{B}_\delta(y(b,s))$ and $I_c(b', y(b,s)) = (b', y(b,s))$, so the deviation payoff is simply

$$U_0^B(b', y(b,s)) + b. \tag{22}$$

Since $b' > \hat{B}_\delta(b,s) \geq B_0(b,s)$ (where the last inequality follows from Lemma A2), the deviation payoff cannot exceed that of (20). Hence, the deviation is not profitable. Now consider the possibility that $(b', y(b,s))$ lies in region (ii). In this case, $x(b', y(b,s)) = b'$ and $y(b', y(b,s)) = \hat{S}_\delta(b') > y(b,s)$. Again using the one-period deviation principle, the deviation payoff is obtained as:

$$\max\{\alpha\phi(b', y(b,s)) + \delta\alpha[\hat{S}_\delta(b') - y(b,s)] + b, \delta[U_0^B(b', \hat{S}_\delta(b')) + b'] - (b' - b)\}. \tag{23}$$

This payoff cannot exceed

$$U_0^B(b', \hat{S}_\delta(b')) + b. \tag{24}$$

By the definition of $\hat{B}_\delta(\cdot)$, $U_0^B(b', \hat{S}_\delta(b')) + b' \geq 0$, so the second term of (23) cannot exceed (24). Likewise, the first term of (23) is increasing in $y(b,s)$ for $y(b,s) \leq \hat{S}_\delta(b')$ (since $\hat{S}_\delta(b') \leq S_\delta(b') \leq S_1(b')$) and equals (24) at $y(b,s) = \hat{S}_\delta(b')$. Now notice that (24) cannot exceed (20) since by definition $U_0^B(b', \hat{S}_\delta(b')) = \pi^B(b')$ is nonincreasing in b'. Hence, the deviation to $b' > x(b,s)$ is unprofitable.

The proof of the optimality of the seller's response is symmetric. ▌

We are now ready to prove our main result.

Proposition A3 *Given (SIR_α) and $\phi_{bs} \geq 0$, there exists a $\delta^{**} < 1$ such that, for all $\delta \geq \delta^{**}$, there exists a MPE in which the parties choose (b_δ, s_δ) and trade in the first period.*

Proof. Given (SIR_α), there exists $\delta_{IR} < 1$ such that $U_0^i(b_\delta, s_\delta) \geq 0$, $i = B, S$, $\forall \delta \geq \delta_{IR}$. Set $\hat{\delta}^{**} := \max\{\delta_{IR}, \tilde{\delta}\}$ (where $\tilde{\delta}$ is defined in Lemma A2), and fix $\delta \geq \hat{\delta}^{**}$. We modify the strategy, I_c, as follows. The new strategy profile, $\hat{I}$, coincides with I_c, except for a subset in region (ii) of (b,s) with $U_0^S(b, \hat{S}_\delta(b)) + s < 0$ and for a subset in region (iii) of (b,s) with $U_0^B(\hat{B}_\delta(s), s) + b < 0$. The new strategies for these cases are described as follows. Fix any (b,s) in (ii) with $U_0^S(b, \hat{S}_\delta(b)) + s < 0$. Take $\tilde{b}(b,s)$ to be the smallest $b'' > b$ such that $U_0^S(b'', \hat{S}_\delta(b'')) + s = 0$.[32] The new strategy then

[32]Existence of $\tilde{b}$ can be seen as follows. First note that $\hat{S}_\delta(b) > \max\{s_\delta, S_0(b)\}$, or else

$$U_0^S(b, \hat{S}_\delta(b)) + s = U_0^S(b, \max\{s_\delta, S_0(b)\}) + s \geq U_0^S(b, \max\{s_\delta, 0\}) + s \geq U_0^S(b_\delta, s_\delta) \geq 0,$$

by (b_δ, s_δ) being individually rational. Hence, there exists $b' > b$ such that $\hat{S}_\delta(b') = \max\{s_\delta, S_0(b')\}$, by the definition of $\hat{S}_\delta(\cdot)$. Since $U_0^S(b', \hat{S}_\delta(b')) + s = U_0^S(b', \max\{s_\delta, S_0(b')\}) + s > 0$ and since $U_0^S(\cdot, \hat{S}_\delta(\cdot)) + s$ is continuous, there exists $\tilde{b}$ with $U_0^S(\tilde{b}, \hat{S}_\delta(\tilde{b})) + s = 0$.

specifies $\hat{I}(b,s) := (\tilde{b}(b,s),s)$. In other words, at (b,s), the buyer invests up to $\tilde{b}$ without the seller making any further investment. (The arguments of $\tilde{b}$ are suppressed whenever no confusion is likely.) Once $(\tilde{b},s)$ is reached, then the original profile, I_c, applies, since $U_0^S(\tilde{b},\hat{S}_\delta(\tilde{b}))+s=0$. That is, $I_c(\tilde{b},s) := (\tilde{b},\hat{S}_\delta(\tilde{b}))$. The $\hat{I}$ for region (iii) is constructed analogously. Since the changes occur only in region (ii) or region (iii), the stated outcome will arise if the parties follow $\hat{I}$. The associated bargaining strategy is specified as above.

We now prove that, for δ large enough, no deviation is profitable. Fix any (b,s) with $b>b_\delta$ and $U_0^S(b,\hat{S}_\delta(b))+s<0$. As stated earlier, $\hat{I}(b,s)=(\tilde{b}(b,s),s)$. We first consider the buyer's incentive. If there is no deviation, the buyer receives

$$\begin{aligned} v^B(\delta) &:= \max\{\alpha\phi(\tilde{b},s)+\alpha\delta(\hat{S}_\delta(\tilde{b})-s)-(\tilde{b}-b), \delta\alpha\phi(\tilde{b},\hat{S}_\delta(\tilde{b}))-(\tilde{b}-b)\} && (25)\\ &\geq \delta\alpha\phi(\tilde{b},\hat{S}_\delta(\tilde{b}))-(\tilde{b}-b)\\ &= \delta[\phi(\tilde{b},\hat{S}_\delta(\tilde{b}))-(\hat{S}_\delta(\tilde{b})-s)]-(\tilde{b}-b), && (26)\end{aligned}$$

where the last equality follows from the fact that $(1-\alpha)\phi(\tilde{b},\hat{S}_\delta(\tilde{b}))-(\hat{S}_\delta(\tilde{b})-s)=0$. By the definition of $\hat{S}_\delta(\cdot)$, $v_B(1)>0$, so there exists $\hat{\delta}(b,s)<1$ such that $v_B(\delta)\geq 0$ for any $\delta\geq\hat{\delta}(b,s)$.

Suppose the buyer deviates to $b'\in[b,\tilde{b})$. Since $\hat{I}(b',s)=(\tilde{b},s)$, the buyer's deviation payoff is no greater than

$$\max\{\phi(b',s)-(b'-b), \delta[v^B(\delta)+b'-b]-(b'-b)\}.$$

For $\delta\geq\hat{\delta}(b,s)$, $v^B(\delta)\geq\delta[v^B(\delta)+b'-b]-(b'-b)$, so the deviation will be unprofitable if

$$\phi(b',s)-(b'-b)\leq\delta[\phi(\tilde{b},\hat{S}_\delta(\tilde{b}))-(\hat{S}_\delta(\tilde{b})-s)]-(\tilde{b}-b). \quad (27)$$

We have

$$\begin{aligned} \phi(b',s)-(b'-b) &< \phi(b',\hat{S}_\delta(b'))-(b'-b)-(\hat{S}_\delta(b')-s)\\ &= U_0^B(b',\hat{S}_\delta(b'))+b+U_0^S(b',\hat{S}_\delta(b'))+s\\ &\leq U_0^B(\tilde{b},\hat{S}_\delta(\tilde{b}))+b+U_0^S(\tilde{b},\hat{S}_\delta(\tilde{b}))+s\\ &= \phi(\tilde{b},\hat{S}_\delta(\tilde{b}))-(\hat{S}_\delta(\tilde{b})-s)-(\tilde{b}-b),\end{aligned}$$

where the first inequality follows since $s<\hat{S}_\delta(b')$[33] and since $\hat{S}_\delta(b')\leq S_\delta(b')$ (by Lemma A2) and the second inequality follows from $U_0^B(b',\hat{S}_\delta(b'))=U_0^B(\tilde{b},\hat{S}_\delta(\tilde{b}))$ (since $\hat{S}_\delta(b'')>\max\{s_\delta,S_0(b'')\}$ for all

[33]Since (b,s) is in region (ii), so is (b',s), which means $s\leq\hat{S}_\delta(b')$. To see $s<\hat{S}_\delta(b')$, suppose to the contrary that $s=\hat{S}_\delta(b')$. Then, $U_0^S(b',\hat{S}_\delta(b'))+s=(1-\alpha)\phi(b',\hat{S}_\delta(b'))>0$, which contradicts the fact that $b'\in[b,\tilde{b}(b,s))$.

$b'' \in [b', \tilde{b}]$, meaning that $U_0^B(b'', \hat{S}_\delta(b''))$ is constant for all $b'' \in [b', \tilde{b}]$) and since $U_0^S(b', \hat{S}_\delta(b')) + s \leq 0 = U_0^S(\tilde{b}, \hat{S}_\delta(\tilde{b})) + s$. The above inequality is rewritten as:

$$\frac{\phi(b', s) - b' + \tilde{b}}{\phi(\tilde{b}, \hat{S}_\delta(\tilde{b})) - (\hat{S}_\delta(\tilde{b}) - s)} < 1, \tag{28}$$

for all b' with $U_0^S(b', s) \leq 0$ and for the corresponding $\tilde{b}(b, s) = \tilde{b}(b', s)$. Hence, if δ exceeds the LHS of (28), then (27) holds, so the deviation will be unprofitable. Now, let $\hat{\delta}^B$ be the value of the following constrained maximization problem:

$$\max_{(b,s) \in \mathcal{X}(b_\delta) \times \mathcal{Y}} \max \left\{ \frac{\phi(b, s) - b + \tilde{b}(b, s)}{\phi(\tilde{b}(b, s), \hat{S}_\delta(\tilde{b}(b, s))) - (\hat{S}_\delta(\tilde{b}(b, s))) - s)}, \hat{\delta}(b, s) \right\}$$
$$\text{s.t.} \qquad U_0^S(b, s) \leq 0.$$

Since the constraint set is compact, the maximum is well defined and must be less than 1. Clearly, for any $\delta \geq \hat{\delta}^B$, the buyer has no incentive to deviate to $b' \in [b, \tilde{b}(b, s))$ from any (b, s) with $U_0^S(b, \hat{S}_\delta(b)) + s < 0$. We can define a similar threshold value, $\hat{\delta}^S < 1$ for (b, s) in region (iv) with $U_0^B(\hat{B}_\delta(s), s) + b < 0$.

Fix any $\delta \geq \delta^{**} := \max\{\hat{\delta}^{**}, \hat{\delta}^B, \hat{\delta}^S\}(< 1)$. Suppose next the buyer deviates to $b' > \tilde{b}(b, s)$, starting from (b, s) with $b > b_\delta$ and $U_0^S(b, \hat{S}_\delta(b)) + s < 0$. The deviation gives the buyer at most

$$\max\{\alpha\phi(b'', s) + \alpha\delta(\hat{S}_\delta(b'') - s) - (b'' - b), \delta\alpha\phi(b'', \hat{S}_\delta(b'')) - (b'' - b)\}, \tag{29}$$

for some $b'' \geq b' > \tilde{b}$.[34] We now compare (25) and (29), term by term. First, note that

$$\begin{aligned} \alpha\phi(\tilde{b}, s) + \alpha\delta(\hat{S}_\delta(\tilde{b}) - s) - \tilde{b} &= \alpha\phi(\tilde{b}, \hat{S}_\delta(\tilde{b})) - \tilde{b} - \int_s^{\hat{S}_\delta(\tilde{b})} \alpha[\phi_s(\tilde{b}, s') - \delta] ds' \\ &\geq \alpha\phi(b'', \hat{S}_\delta(b'')) - b'' - \int_s^{\hat{S}_\delta(b'')} \alpha[\phi_s(b'', s') - \delta] ds' \\ &= \alpha\phi(b'', s) + \alpha\delta(\hat{S}_\delta(b'') - s) - b'', \end{aligned}$$

where the inequality follows since $\alpha\phi(b, \hat{S}_\delta(b)) - b = U_0^B(b, \hat{S}_\delta(b))$ is nonincreasing in b, $\phi_s(b'', s') - \delta \geq \phi_s(\tilde{b}, s') - \delta$, $\phi_s(b'', s') - \delta \geq 0$ for $s' \leq \hat{S}_\delta(b'') \leq S_\delta(b'')$, and since $\hat{S}_\delta(b'') \geq \hat{S}_\delta(\tilde{b})$ (by Lemma A2). Next,

$$\begin{aligned} \delta\alpha\phi(b'', \hat{S}_\delta(b'')) - b'' - [\delta\alpha\phi(\tilde{b}, \hat{S}_\delta(\tilde{b})) - \tilde{b}] &\leq \alpha\phi(b'', \hat{S}_\delta(b'')) - b'' - [\alpha\phi(\tilde{b}, \hat{S}_\delta(\tilde{b})) - \tilde{b}] \\ &= U_0^B(b'', \hat{S}_\delta(b'')) - U_0^B(\tilde{b}, \hat{S}_\delta(\tilde{b})) \leq 0. \end{aligned}$$

[34] Conceivably, $U_0^S(b', \hat{S}_\delta(b')) + s < 0$. But in this case, the argument above shows that, for $\delta \geq \delta'$, it would be more profitable for the buyer to deviate to $b'' \equiv \tilde{b}(b', s)$. Hence, it suffices to show that deviation to b'' is unprofitable in this case.

Combining the two inequalities, we conclude that the equilibrium payoff in (25) dominates the deviation payoff in (29).

We now examine the seller's incentive when starting from (b, s) such that $U_0^S(b, \hat{S}_\delta(b)) + s < 0$. If no deviation occurs, then the seller receives

$$\max\{(1-\alpha)\phi(\tilde{b}, s) - \alpha\delta(\hat{S}_\delta(\tilde{b}) - s), \delta[U_0^S(\tilde{b}, \hat{S}_\delta(\tilde{b})) + s]\} \geq 0,$$

since $U_0^S(\tilde{b}, \hat{S}_\delta(\tilde{b})) + s = 0$. If the seller deviates to $s' \in (s, \hat{S}_\delta(b)]$, she receives

$$\max\{(1-\alpha)\phi(\tilde{b}, s') - \alpha\delta[\hat{S}_\delta(\tilde{b}) - s'] - (s' - s), \delta[U_0^S(\tilde{b}, \hat{S}_\delta(\tilde{b})) + s'] - (s' - s)\}.$$

Note that the first term is nondecreasing in s' (given Lemma A2, since $s' \leq \hat{S}_\delta(\tilde{b}) \leq S_\delta(\tilde{b})$) and equals $U_0^S(\tilde{b}, \hat{S}_\delta(\tilde{b})) + s = 0$ when $s' = \hat{S}_\delta(\tilde{b})$. The second term is nonincreasing in s' and equals $U_0^S(\tilde{b}, \hat{S}_\delta(\tilde{b})) + s = 0$ when $s' = s$. Hence, the deviation payoff is nonpositive, so the deviation is unprofitable. Now suppose the seller deviates to $s' > \hat{S}_\delta(\tilde{b})$. Since (b, s') is now in region (iv), the deviation payoff will be $U_0^S(\max\{\tilde{b}, \hat{B}_\delta(s'')\}, s'') + s$, for some $s'' \geq s' > \hat{S}_\delta(\tilde{b})$.[35] This payoff is no greater than $U_0^S(\tilde{b}, \hat{S}_\delta(\tilde{b})) + s = 0$, since $s'' \geq s' > \hat{S}_\delta(\tilde{b}) \geq S_0(\tilde{b})$ and since $U_0^S(\hat{B}_\delta(s''), s'')$ is nonincreasing in s'' for $s'' \geq s_\delta$. Hence, the deviation is not profitable for the seller.

The proof for the new strategies in region (iii) is completely analogous. Finally, we need to check the deviation incentive, when starting from (b, s) with $\hat{I}(b, s) = I_c(b, s)$. Clearly, the equilibrium payoff will remain the same for each party. The deviation payoff could change since deviation may bring the parties to the region for which the subsequent strategies are different. But as noted above (see footnotes 34 and 35), the deviation into a region with $\hat{I} \neq I_c$ is dominated by a deviation into a region with $\hat{I} = I_c$, so no new incentive for deviation is introduced when starting from (b, s) with $\hat{I}(b, s) = I_c(b, s)$. ▌

Appendix B: Proofs of Section 5

Proof of Lemma 1: Since $(\hat{x}, \hat{y})$ is a limit point of maximizers attaining the value of (7), there is a (sub)sequence $(b_n, s_n) \to (b, s)$ as $n \to \infty$ such that its associated maximizer of (6), $(x^*(b_n, s_n), y^*(b_n, s_n))$, converges to $(\hat{x}, \hat{y})$ and attains the value of (7) as $n \to \infty$. Hence, it follows that

$$\limsup_{(b', s') \to (b, s)} V^S(b', s') = \lim_{n \to \infty} \{\overline{\sigma}(x^*(b_n, s_n), y^*(b_n, s_n)) - (y^*(b_n, s_n) - s_n)\} \leq \overline{\sigma}(\hat{x}, \hat{y}) - (\hat{y} - s), \quad (30)$$

[35] Again, we invoked the fact that if the seller's deviation to s' puts the parties at $(\tilde{b}(b, s), s')$ with $U_0^B(\tilde{b}(b, s), s') < 0$, then it is dominated by her deviation to $s'' := \tilde{s}(\tilde{b}(b, s), s')$.

where the inequality holds since $\overline{\sigma}$ is u.s.c.

Consider next the lowest possible continuation payoff for the buyer when the previous period stock was (b', s') and he now invests up to $\hat{x}(b,s) \vee b'$:

$$\min_{s'' \in \mathcal{Y}(s')} \underline{\beta}(\hat{x} \vee b', s'') - ((\hat{x} \vee b') - b'). \tag{31}$$

Note that

$$\liminf_{(b',s') \to (b,s)} \{ \min_{s'' \in \mathcal{Y}(s')} \underline{\beta}(\hat{x} \vee b', s'') - [(\hat{x} \vee b') - b'] \} \le \min_{s'' \in \mathcal{Y}(s)} \{ \underline{\beta}(\hat{x}, s'') - (\hat{x} - b) \}, \tag{32}$$

since any limit point of the minimizers belongs to $\mathcal{Y}(s)$ (since the set is closed), $\underline{\beta}$ is l.s.c., and $\hat{x} \ge b$.

Now, take any (b', s') in a neighborhood of (b, s). The following Bellman characterization is obtained from our definitions.

$$\overline{w}^S(b', s') \le (1-\alpha) \max\{\phi(b', s') - \delta[\min_{s'' \in \mathcal{Y}(s')} \underline{\beta}(\hat{x} \vee b', s'') - ((\hat{x} \vee b') - b)], \delta V^S(b', s')\} + \alpha \delta V^S(b', s'),$$

where $\min_{s'' \in \mathcal{Y}(s')} \underline{\beta}(\hat{x} \vee b', s'') - ((\hat{x} \vee b') - b)$ is the minimum offer the buyer would accept at (b', s'), as explained in the paragraph following Lemma 1 (see also footnote 21). Taking a limit superior as $(b', s') \to (b, s)$ on both sides, we get

$$\begin{aligned}
& \overline{\sigma}(b,s) \\
= \ & \limsup_{(b',s') \to (b,s)} \overline{w}^S(b', s') \\
\le \ & \limsup_{(b',s') \to (b,s)} \left\{ (1-\alpha) \max\{\phi(b', s') - \delta[\min_{s'' \in \mathcal{Y}(s')} \underline{\beta}(\hat{x} \vee b', s'') - ((\hat{x} \vee b') - b)], \delta V^S(b', s')\} + \alpha \delta V^S(b', s') \right\} \\
\le \ & (1-\alpha) \max \left\{ \phi(b,s) - \delta \left[\liminf_{(b',s') \to (b,s)} \left\{ \min_{s'' \in \mathcal{Y}(s')} \underline{\beta}(\hat{x} \vee b', s'') - ((\hat{x} \vee b') - b) \right\} \right], \delta \limsup_{(b',s') \to (b,s)} V^S(b', s') \right\} \\
& \quad + \alpha \delta \left[\limsup_{(b',s') \to (b,s)} V^S(b', s') \right] \\
\le \ & (1-\alpha) \max \left\{ \phi(b,s) - \delta \left[\min_{s'' \in \mathcal{Y}(s)} \underline{\beta}(\hat{x}, s'') - (\hat{y} - s) \} \right], \delta \left[\overline{\sigma}(\hat{x}, \hat{y}) - (\hat{x} - b) \right] \right\} + \alpha \delta [\overline{\sigma}(\hat{x}, \hat{y}) - (\hat{y} - s)],
\end{aligned}$$

where the first equality follows from the definition, the first inequality follows from the above inequality, and the third inequality follows from (30) and (32).

The first inequality of (9) follows exactly the same line of arguments using (30) and (32), except that the string of inequalities is reversed. The second inequality of (9) is obvious.

The characterizations for $\overline{\beta}$ and $\underline{\sigma}$ are completely symmetric. ▌

Proof of Lemma 2: We first prove the following preliminary lemmas.

Lemma B1 *Assume that the investments are weak substitutes ($\phi_{bs}(b,s) \le 0$). For any (b,s) with $s \ge S_0(b)$ and $b \ge B_0(s)$,*

$$\overline{\sigma}(b,s) \ge (1-\alpha)\phi(b,s) \ge \underline{\sigma}(b,s), \text{ and } \overline{\beta}(b,s) \ge \alpha\phi(b,s) \ge \underline{\beta}(b,s).$$

Proof. Fix any (b,s) with $s \ge S_0(b)$ and $b \ge B_0(s)$. Then, (Markovian) investment strategies, $I_0(b',s') = (b',s')$, for all $(b',s') \ge (b,s)$, form a SPE (along with the associated optimal bargaining strategies). The resulting payoffs at (b,s) are $(1-\alpha)\phi(b,s)$ and $\alpha\phi(b,s)$ for the seller and the buyer, respectively. The inequalities then follow from the definitions of $\overline{\sigma}$, $\underline{\sigma}$, $\overline{\beta}$ and $\underline{\beta}$. ▌

Lemma B2 *Assume that the investments are weak substitutes ($\phi_{bs}(b,s) \le 0$). Fix any $\epsilon > 0$ and $n \in \mathbb{N}$. In any SPE, for any (b,s) with $s \ge S_\delta(b) + \epsilon$ and $b \ge B_\delta(s) + \epsilon$,*

$$\overline{\sigma}(b,s) \le (1-\alpha)\phi(b,s) + O(1/n), \underline{\sigma}(b,s) \ge (1-\alpha)\phi(b,s) + O(1/n)$$

and

$$\overline{\beta}(b,s) \le \alpha\phi(b,s) + O(1/n), \underline{\beta}(b,s) \ge \alpha\phi(b,s) + O(1/n),$$

where $O(1/n)$ is a term that vanishes as $n \to \infty$.

Proof. We first divide $\mathcal{X}$ into n equal-sized closed intervals, $\{B_1, ..., B_n\}$, and $\mathcal{Y}$ into n equal-sized closed intervals, $\{S_1, ..., S_n\}$. For any $1 \le k,l \le n$, let $Z_{kl} := B_k \times S_l$ be the kl-th block, and let $Z_{kl}^+ := \left[\cup_{(k',l') \ge (k,l)} B_{k'} \times S_{l'}\right] \backslash Z_{kl}$ be the set of higher blocks excluding Z_{kl}. Likewise, let $B_k^+ := [\cup_{k' \ge k} B_{k'}] \backslash B_k$ and $S_l^+ := [\cup_{l' \ge l} S_{l'}] \backslash S_l$. Finally, let $\hat{Z} := \{(b,s) \in \mathcal{X} \times \mathcal{Y} \mid s \ge S_\delta(b) + \epsilon \text{ and } b \ge B_\delta(s) + \epsilon\}$.

The proof proceeds inductively relative to the blocks.

Step 1: *The statement holds true for any $(b,s) \in Z_{nn}$.*

Let $(b^+, s^+) \in \arg\max_{(b',s') \in Z_{nn}} \overline{\sigma}(b',s')$ and $(b^-, s^-) \in \arg\min_{(b',s') \in Z_{nn}} \underline{\beta}(b',s')$. Let $(\hat{x}^+, \hat{y}^+) := (\hat{x}(b^+,s^+), \hat{y}(b^+,s^+))$ and $(\hat{x}^-, \hat{y}^-) := (\hat{x}(b^-,s^-), \hat{y}(b^-,s^-))$. We have

$$\begin{aligned}
\overline{\sigma}(b^+,s^+) &\le (1-\alpha)\max\{\phi(b^+,s^+) - \delta[\min_{s' \in \mathcal{Y}(s^+)} \underline{\beta}(\hat{x}^+,s') - (\hat{x}^+ - b^+)], \delta[\overline{\sigma}(\hat{x}^+,\hat{y}^+) - (\hat{y}^+ - s^+)]\} \\
&\quad + \alpha\delta[\overline{\sigma}(\hat{x}^+,\hat{y}^+) - (\hat{y}^+ - s^+)] \\
&\le (1-\alpha)\max\{\phi(b^+,s^+) - \delta\underline{\beta}(b^-,s^-), \delta\overline{\sigma}(b^+,s^+)\} + \alpha\delta\overline{\sigma}(b^+,s^+), \\
&= (1-\alpha)\phi(b^+,s^+) - (1-\alpha)\delta\underline{\beta}(b^-,s^-) + \alpha\delta\overline{\sigma}(b^+,s^+),
\end{aligned}$$

where the first inequality follows from (9), the second inequality follows from the definitions of (b^+,s^+) and of (b^-,s^-), and the equality holds, or else $\phi(b^+,s^+) - \delta\underline{\beta}(b^-,s^-) < \delta\overline{\sigma}(b^+,s^+)$, implying that $\overline{\sigma}(b^+,s^+) \le 0$, a contradiction to Lemma B1.

Now, fix any $(b,s)\in Z_{nn}$. Then, since both (b,s) and (b^+,s^+) are in Z_{nn}, it follows from the above inequality that

$$\overline{\sigma}(b^+,s^+)\leq(1-\alpha)\phi(b,s)-(1-\alpha)\delta\underline{\beta}(b^-,s^-)+\alpha\delta\overline{\sigma}(b^+,s^+)+O(1/n). \tag{33}$$

Following the same line of argument (but utilizing (9)), we have

$$\underline{\beta}(b^-,s^-)\geq\alpha\phi(b,s)-\alpha\delta\overline{\sigma}(b^+,s^+)+(1-\alpha)\delta\underline{\beta}(b^-,s^-)+O(1/n). \tag{34}$$

Combining (33) and (34), we get

$$\overline{\sigma}(b^+,s^+)\leq(1-\alpha)\phi(b,s)+O(1/n) \text{ and } \underline{\beta}(b^-,s^-)\geq\alpha\phi(b,s)+O(1/n).$$

Since $\overline{\sigma}(b,s)\leq\overline{\sigma}(b^+,s^+)$ and $\underline{\beta}(b,s)\geq\underline{\beta}(b^-,s^-)$, it follows that

$$\overline{\sigma}(b,s)\leq(1-\alpha)\phi(b,s)+O(1/n) \text{ and } \underline{\beta}(b,s)\geq\alpha\phi(b,s)+O(1/n),$$

as needed. The proof on $\overline{\beta}$ and $\underline{\sigma}$ is completely symmetric. ▌

Step 2: *If the above statement holds for every $(b',s')\in Z_{kl}^+$, it also holds for any $(b,s)\in Z_{kl}$.*

With a slight abuse of notation, define $(b^+,s^+)\in\arg\max_{(b',s')\in Z_{kl}\cap\hat{Z}}\overline{\sigma}(b',s')$ and $(b^-,s^-)\in\arg\min_{(b',s')\in Z_{kl}\cap\hat{Z}}\underline{\beta}(b',s')$. Let $(\hat{x}^+,\hat{y}^+):=(\hat{x}(b^+,s^+),\hat{y}(b^+,s^+))$ and $(\hat{x}^-,\hat{y}^-):=(\hat{x}(b^-,s^-),\hat{y}(b^-,s^-))$. Since $\phi_{bs}\leq 0$, $(b^j,s^j)\in\hat{Z}$ means $(\hat{x}^j,\hat{y}^j)\in\hat{Z}$ for $j=+,-$. Let $x^*:=\max\{\hat{x}^+,\hat{x}^-\}$.

Suppose first $\hat{x}^+\in B_k^+$. Then, for any $(b,s)\in Z_{kl}\cap\hat{Z}$, we have

$$\begin{aligned}
&\overline{\sigma}(b^+,s^+)\\
\leq\ &(1-\alpha)\max\left\{\phi(b^+,s^+)-\delta\left(\min_{s'\in\mathcal{Y}(s^+)}\underline{\beta}(\hat{x}^+,s')-(\hat{x}^+-b^+)\right),\delta[\overline{\sigma}(\hat{x}^+,\hat{y}^+)-(\hat{y}^+-s^+)]\right\}\\
&\quad+\alpha\delta[\overline{\sigma}(\hat{x}^+,\hat{y}^+)-(\hat{y}^+-s^+)]\\
\leq\ &(1-\alpha)\max\left\{\phi(b^+,s^+)-\delta\left(\min_{y'\in\mathcal{Y}(s^+)}[\alpha\phi(\hat{x}^+,y')-(\hat{x}^+-b^+)]\right),\delta[(1-\alpha)\phi(\hat{x}^+,\hat{y}^+)-(\hat{y}^+-s^+)]\right\}\\
&\qquad+\alpha\delta\{(1-\alpha)\phi(\hat{x}^+,\hat{y}^+)-(\hat{y}^+-s^+)\}+O(1/n)\\
\leq\ &(1-\alpha)\max\left\{\phi(b^+,s^+)-\delta\left(\alpha\phi(\hat{x}^+,s^+)-(\hat{x}^+-b^+)]\right),\delta(1-\alpha)\phi(\hat{x}^+,s^+)\right\}\\
&\quad+\alpha\delta(1-\alpha)\phi(\hat{x}^+,s^+)+O(1/n)\\
=\ &\max\left\{(1-\alpha)\phi(b^+,s^+)+(1-\alpha)\delta(\hat{x}^+-b^+),\delta(1-\alpha)\phi(\hat{x}^+,s^+)\right\}+O(1/n)\\
\leq\ &\max\{(1-\alpha)\phi(b,s)+(1-\alpha)\delta(x^*-b),\delta(1-\alpha)\phi(x^*,s)\}+O(1/n),
\end{aligned}\tag{35}$$

where the first inequality follows from (8); the second inequality follows from the induction hypothesis since $(\hat{x}^+,\hat{y}^+)\in Z_{kl}^+$, the third inequality holds since the minimum is attained at s^+ and the remaining terms are decreasing in $\hat{y}^+$ (since $(\hat{x}, \hat{y}) \in \hat{Z}$), and the last inequality holds since $\hat{x}^+ \leq x^*$, and both (b^+, s^+) and (b, s) are in Z_{kl}.

Suppose next $\hat{x}^+ \in B_k$. Then, for any $(b, s) \in Z_{kl} \cap \hat{Z}$, we have

$$\begin{aligned}
& \overline{\sigma}(b^+, s^+) \\
\leq \quad & (1-\alpha) \max \left\{ \phi(b^+, s^+) - \delta \left(\min_{s' \in \mathcal{Y}(s^+)} [\underline{\beta}(\hat{x}^+, s') - (\hat{x}^+ - b^+)] \right), \delta[\overline{\sigma}(\hat{x}^+, \hat{y}^+) - (\hat{y}^+ - s^+)] \right\} \\
& \quad + \alpha\delta[\overline{\sigma}(\hat{x}^+, \hat{y}^+) - (\hat{y}^+ - s^+)] \\
\leq \quad & (1-\alpha) \max \left\{ \phi(b^+, s^+) - \delta \left[\underline{\beta}(b^-, s^-) \wedge \left(\min_{s' \in \mathcal{Y}(s^+)} [\alpha\phi(\hat{x}^+, s') - (\hat{x}^+ - b^+)] \right) \right], \right. \\
& \quad \left. \delta \Big[\overline{\sigma}(b^+, s^+) \vee \big[(1-\alpha)\phi(\hat{x}^+, \hat{y}^+) - (\hat{y}^+ - s^+) \big] \Big] \right\} \\
& \quad + \alpha\delta \Big[\overline{\sigma}(b^+, s^+) \vee \big[(1-\alpha)\phi(\hat{x}^+, \hat{y}^+) - (\hat{y}^+ - s^+) \big] \Big] + O(1/n) \\
\leq \quad & (1-\alpha) \max \Big\{ \phi(b^+, s^+) - \delta \Big(\underline{\beta}(b^-, s^-) \wedge [\alpha\phi(\hat{x}^+, s^+) - (\hat{x}^+ - b^+)] \Big), \delta[\overline{\sigma}(b^+, s^+) \vee (1-\alpha)\phi(\hat{x}^+, s^+)] \Big\} \\
& \quad + \alpha\delta \Big[\overline{\sigma}(b^+, s^+) \vee (1-\alpha)\phi(\hat{x}^+, s^+) \Big] + O(1/n) \\
\leq \quad & (1-\alpha) \max\{\phi(b, s) - \delta\underline{\beta}(b^-, s^-), \delta\overline{\sigma}(b^+, s^+)\} + \alpha\delta\overline{\sigma}(b^+, s^+)\} + O(1/n) \\
= \quad & (1-\alpha)\phi(b, s) - (1-\alpha)\delta\underline{\beta}(b^-, s^-) + \alpha\delta\overline{\sigma}(b^+, s^+) + O(1/n),
\end{aligned} \tag{36}$$

where the first inequality follows from (8); the second inequality follows since $\overline{\sigma}(\hat{x}^+, \hat{y}^+) - (\hat{y}^+ - s^+) \leq \overline{\sigma}(b^+, s^+)$ if $(\hat{x}^+, \hat{y}^+) \in Z_{kl}$, or else the induction hypothesis applies; the third inequality holds since the minimum is attained at $s' = s^+$; the fourth inequality holds since $(\hat{x}^+, s^+) \in Z_{kl}$, which, along with Lemma B1, implies that $\underline{\beta}(b^-, s^-) \leq \underline{\beta}(\hat{x}^+, s^+) \leq \alpha\phi(\hat{x}^+, s^+)$, and likewise $\overline{\sigma}(b^+, s^+) \geq \overline{\sigma}(\hat{x}^+, s^+) \geq (1-\alpha)\phi(\hat{x}^+, s^+)$, and since (b^+, s^+) and (b, s) are both in Z_{kl}; and the last equality holds since $\phi(b, s) - \delta\underline{\beta}(b^-, s^-) \geq \delta\overline{\sigma}(b^+, s^+)$, or else $\overline{\sigma}(b^+, s^+) = 0$, a contradiction to Lemma B1.

For any $(b, s) \in Z_{kl} \cap \hat{Z}$, (35) and (36) are summarized as,

$$\begin{aligned}
& \overline{\sigma}(b^+, s^+) \\
\leq \quad & \begin{cases} \max\{(1-\alpha)\phi(b, s) + (1-\alpha)\delta(x^* - b), \delta(1-\alpha)\phi(x^*, s)\} + O(1/n) & \text{if } \hat{x}^+ \in B_k^+, \\ (1-\alpha)\phi(b, s) - (1-\alpha)\delta\underline{\beta}(b^-, s^-) + \alpha\delta\overline{\sigma}(b^+, s^+) + O(1/n) & \text{if } \hat{x}^+ \in B_k. \end{cases}
\end{aligned} \tag{37}$$

By a symmetric argument [utilizing (9)], for any $(b, s) \in Z_{kl} \cap \hat{Z}$,

$$\underline{\beta}(b^-, s^-) \geq \begin{cases} \alpha\phi(b, s) - (1-\alpha)\delta(x^* - b) + O(1/n) & \text{if } \hat{x}^- \in B_k^+, \\ \alpha\phi(b, s) - \alpha\delta\overline{\sigma}(b^+, s^+) + (1-\alpha)\delta\underline{\beta}(b^-, s^-) + O(1/n) & \text{if } \hat{x}^- \in B_k. \end{cases} \tag{38}$$

Step 2 then follows from a series of observations.

Claim 1: *Fix any* $(b,s) \in Z_{kl} \cap \hat{Z}$. *If* $x^* \in B_k$, *then* $\overline{\sigma}(b,s) \leq (1-\alpha)\phi(b,s) + O(1/n)$ *and* $\underline{\beta}(b,s) \geq \alpha\phi(b,s) + O(1/n)$. *If* $x^* \in B_k^+$, *then* $\overline{\sigma}(b,s) \leq \max\{(1-\alpha)\phi(b,s) + (1-\alpha)\delta(x^*-b), \delta\phi(x^*,s)\} + O(1/n)$, *and* $\underline{\beta}(b,s) \geq \alpha\phi(b,s) - (1-\alpha)\delta(x^*-b) + O(1/n)$.

Proof. If $x^* \in B_k$, then $\hat{x}^+, \hat{x}^- \in B_k$, so the first statement follows immediately from solving (37) and (38) and the fact that $\overline{\sigma}(b,s) \leq \overline{\sigma}(b^+,s^+)$ and $\underline{\beta}(b,s) \geq \underline{\beta}(b^-,s^-)$. Suppose now $x^* \in B_k^+$. There are three possibilities. If $\hat{x}^+, \hat{x}^- \in B_k^+$, then the result is again immediate from (37) and (38). If $\hat{x}^+ \in B_k^+$ and $\hat{x}^- \in B_k$, then the inequality for $\overline{\sigma}(b,s)$ follows from (37), and the characterization for $\underline{\beta}$ in (38) simplifies to

$$\begin{aligned}
\underline{\beta}(b^-,s^-) &\geq \frac{\alpha\phi(b,s) - \alpha\delta\overline{\sigma}(b^+,s^+)}{1-(1-\alpha)\delta} + O(1/n) \\
&= \alpha\phi(b,s) - (1-\alpha)\delta \max\left\{\frac{\alpha\delta}{1-(1-\alpha)\delta}(x^*-b), \frac{\alpha[\delta\phi(x^*,s) - \phi(b,s)]}{1-(1-\alpha)\delta}\right\} + O(1/n) \\
&\geq \alpha\phi(b,s) - (1-\alpha)\delta(x^*-b) + O(1/n),
\end{aligned}$$

where the equality is obtained by substituting from (37), and the inequality holds since $\frac{\alpha\delta}{1-(1-\alpha)\delta} \leq 1$ and since $\alpha[\phi(x^*,s) - \phi(b,s)] \leq (1-(1-\alpha)\delta)(x^*-b)$ whenever $(b,s) \in \hat{Z}$ and $x^* > b$. Since $\underline{\beta}(b,s) \geq \underline{\beta}(b^-,s^-)$, the claimed inequality for $\underline{\beta}(b,s)$ holds. Finally, if $\hat{x}^+ \in B_k$ and $\hat{x}^- \in B_k^+$, the inequality for $\underline{\beta}(b,s)$ follows from (38), and (37) simplifies to

$$\begin{aligned}
\overline{\sigma}(b^+,s^+) &\leq \frac{(1-\alpha)\phi(b,s) - (1-\alpha)\delta\underline{\beta}(b^-,s^-)}{1-\alpha\delta} + O(1/n) \\
&= (1-\alpha)\phi(b,s) + (1-\alpha)\delta(x^*-b)\frac{(1-\alpha)\delta}{1-\alpha\delta} + O(1/n) \\
&\leq \alpha\phi(b,s) + (1-\alpha)\delta(x^*-b) + O(1/n),
\end{aligned}$$

where the inequality follows since $(1-\alpha)\delta \leq 1-\alpha\delta$. Clearly, the stated condition for $\overline{\sigma}(b,s)$ follows.

Claim 2: *For any* $(b,s) \in Z_{kl} \cap \hat{Z}$, $x^* \leq b + O(1/n)$.

Proof. Let

$$M := \min_{(b',s') \in \hat{Z}} \left| \frac{\partial U_\delta^B(b',s')}{\partial b} \right|.$$

Since the derivatives are all negative in the constraint set and bounded away from 0, we must have $M > 0$.

Consider any $(b,s) \in Z_{kl} \cap \hat{Z}$. Without any loss, assume that $x^* = \hat{x}^+$. (The case with $x^* = \hat{x}^-$ is completely analogous.) The claimed result will hold trivially if $\hat{x}^+ \in B_k$, so assume $\hat{x}^+ \in B_k^+$, implying that $(\hat{x}^+, \hat{y}^+) \in Z_{kl}^+ \cap \hat{Z}$.

Recall that $(\hat{x}^+, \hat{y}^+)$ is a limit point of a sequence of (maximizing) pairs, $(x', y') \in \overline{\mathcal{I}}(b', s')$ with $(b', s') \to (b^+, s^+)$. Take its subsequence, (b'', s''), such that $(x'', y'') \in \overline{\mathcal{I}}(b'', s'')$ and that $(x'', y'') \to (\hat{x}^+, \hat{y}^+)$ as $(b'', s'') \to (b^+, s^+)$.[36] For (b'', s'') sufficiently close to (b^+, s^+), $(x'', y'') \in Z_{kl}^+ \cap \hat{Z}$, since (x'', y'') converges to $(\hat{x}^+, \hat{y}^+) \in Z_{kl}^+ \cap \hat{Z}$. Hence, by the induction hypothesis

$$\overline{\beta}(x'', y'') \leq \alpha\phi(x'', y'') + O(1/n). \tag{39}$$

Meanwhile, (b'', y'') converges to $(b^+, \hat{y}^+)$, which is either in $Z_{kl} \cap \hat{Z}$ or $Z_{kl}^+ \cap \hat{Z}$, depending on whether $\hat{y}^+ \in S_l$ or $\hat{y}^+ \in S_l^+$. In the latter case, for (b'', s'') sufficiently close to (b^+, s^+), the induction hypothesis implies that

$$\underline{\beta}(b'', y') \geq \alpha\phi(b'', y'') + O(1/n) \geq \alpha\phi(b, y'') + O(1/n), \tag{40}$$

where the second inequality holds since b and b'' are both in B_k (for (b'', s'') sufficiently close to (b^+, s^+)) and ϕ is continuous. If $\hat{y}^+ \in S_l$, for (b'', s'') sufficiently close to (b^+, s^+), $(b'', y'') \in Z_{kl} \cap \hat{Z}$. Hence, by Claim 1,

$$\underline{\beta}(b'', y') \geq \alpha\phi(b'', y'') - (1-\alpha)\delta(x'' - b'') + O(1/n) \geq \alpha\phi(b, y'') - (1-\alpha)\delta(x'' - b) + O(1/n), \tag{41}$$

where the second inequality holds, as before, since b and b'' are both in B_k. Clearly, the RHS of (41) provides a lower bound of $\underline{\beta}(b'', y'')$, even when $y'' \in S_l^+$.

Since $(x'', y'') \in \overline{\mathcal{I}}(b'', s'')$, incentive compatibility requires $\overline{\beta}(x'', y'') - (x'' - b'') \geq \underline{\beta}(b'', y'')$, or

$$\alpha\phi(x'', y'') - (x'' - b) + O(1/n) \geq \alpha\phi(b, y'') - (1-\alpha)\delta(x'' - b) + O(1/n), \tag{42}$$

again using the fact that b and b'' are both in B_k.[37] In the limit as $(b'', s'') \to (b^+, s^+)$, we have

$$O(1/n) \geq -\alpha[\phi(\hat{x}^+, \hat{y}^+) - \phi(b, \hat{y}^+)] + (1 - (1-\alpha)\delta)(\hat{x}^+ - b)$$
$$\iff O(1/n) \geq -\int_b^{\hat{x}^+} \frac{\partial U_\delta(\tilde{b}, \hat{y}^+)}{\partial \tilde{b}} d\tilde{b} \geq M(\hat{x}^+ - b).$$

Since $\hat{x}^+ > b$, we must have $x^* = \hat{x}^+ \leq b + O(1/n)$, as needed. ▌

[36] Such a subsequence is well defined since (x', y') lies in a compact set $\mathcal{X} \times \mathcal{Y}$.

[37] Note that (x'', y'') is in $\bar{\mathcal{I}}(b'', s'')$, but it may not be in $\mathcal{I}(b'', s'')$. In any event, since (x'', y'') is in $\bar{\mathcal{I}}(b'', s'')$, there exists a sequence, (x_m, y_m) converging to (x'', y'') as $m \to \infty$, such that $(x_m, y_m) \in \mathcal{I}(b'', s'')$ for each m. Since $(x_m, y_m) \to (x'', y'')$ as $m \to \infty$, for sufficiently large m, $\overline{\beta}(x_m, y_m)$ and $\underline{\beta}(b'', y_m)$ have the same characterizations as (39) and (41), respectively (with (x'', y'') replaced by (x_m, y_m)). Hence, the incentive compatibility condition for $(x_m, y_m) \in \mathcal{I}(b'', s'')$ yields (42) in the limit as $m \to \infty$.

Claim 3: *For any* $(b, s) \in Z_{kl} \cap \hat{Z}$,

$$\overline{\sigma}(b, s) \leq (1 - \alpha)\phi(b, s) + O(1/n) \text{ and } \underline{\beta}(b, s) \geq \alpha\phi(b, s) + O(1/n).$$

Proof. The result immediate from applying Claim 2 to Claim 1. ▮

Since Lemma B2 holds for any $\epsilon > 0$ and for any n, it must hold in the limit as $n \to \infty$ and $\epsilon \to 0$. Hence, Lemma 2 holds for all (b, s) with $b > B_\delta(s)$ and $s > S_\delta(b)$, with respect to $\overline{\sigma}$ and $\underline{\beta}$. The characterizations for $\overline{\beta}$ and $\underline{\sigma}$ are obtained analogously. ▮

Proof of Lemma 3: Fix any $\epsilon > 0$ and $\delta \in [0, 1)$. We then proceed to establish three statements, i), ii) and iii). The proof of each statement resembles almost exactly that of Lemma 2, so we only sketch the proof, with a more fleshed out version available upon request.
i) For any $(b, s) \in \mathcal{X}(b_1) \times [s_1 - \epsilon, s_1]$,

$$\overline{\sigma}_\delta(b, s) \leq (1 - \alpha)\phi(b, s_1) + O(\epsilon) \text{ and } \underline{\beta}_\delta(b, s) \geq \alpha\phi(b, s_1) + O(\epsilon),$$

and

$$\underline{\sigma}_\delta(b, s) \geq (1 - \alpha)\phi(b, s_1) + O(\epsilon) \text{ and } \overline{\beta}_\delta(b, s) \leq \alpha\phi(b, s_1) + O(\epsilon),$$

where $O(\epsilon)$ *is a term that vanishes as* $\epsilon \to 0$.

Proof. As with the proof of Lemma 2, it suffices to show that, for any $(b, s) \in \mathcal{X}(b_1) \times [s_1 - \epsilon, s_1]$ and for any $n \in \mathbb{N}$, $\overline{\sigma}_\delta(b, s) \leq (1 - \alpha)\phi(b, s_1) + O(\epsilon) + O(1/n)$, $\underline{\beta}_\delta(b, s) \geq \alpha\phi(b, s_1) + O(\epsilon) + O(1/n)$, $\underline{\sigma}_\delta(b, s) \geq (1 - \alpha)\phi(b, s_1) + O(\epsilon) + O(1/n)$, and $\overline{\beta}_\delta(b, s) \leq \alpha\phi(b, s_1) + O(\epsilon) + O(1/n)$. The proof for this latter claim is almost the same as that of Lemma B2, so we simply highlight how that proof applies in the current context. We divide $\mathcal{X}(b_1)$ into n equal-sized closed intervals, $\{B_1, ..., B_n\}$, and then proceed inductively. Specifically, we begin with $(b, s) \in B_n \times (s_1 - \epsilon, s_1]$, showing that the inequalities of i) hold with an error of order $O(1/n)$. The proof matches precisely that of Step 1 of the proof of Lemma B2, by noting that, by Proposition 2, any equilibrium investment path starting from this set remains in this set (since $(b, s) \in \Omega_\delta$ for any $s > s_1$ and $b > b_1$).

Next, we show that if the statement holds for $(b, s) \in B_k^+ \times [s_1 - \epsilon, s_1]$, then it holds for any $(b, s) \in B_k \times [s_1 - \epsilon, s_1]$. The proof mirrors almost precisely that of Step 2 of Lemma B2. Specifically, the payoff characterizations for $\overline{\sigma}$ and $\underline{\beta}$ in (37) and (38) hold (with error of order $O(\epsilon) + O(1/n)$ instead of $O(1/n)$, and with the arguments of ϕ replaced with (b, s_1)). Claims 1-3 follow since $s_1 > S_\delta(b)$ for all $b \in \mathcal{X}(b_1)$. The only slight difference arises regarding characterization of $\overline{\beta}$ and $\underline{\beta}$, since the situation is not quite symmetric. But the fact that the investment path can never reach beyond s_1 actually simplifies this case: Only a subcase mirroring that for $x^* \in B_k$ in (37) and (38) arises, which makes the characterization immediate for that case. Other than this difference, all the results follow. As before, the claimed statement then follows since n is arbitrary. ▌

ii) For any $(b, s) \in [b_1 - \epsilon, b_1] \times \mathcal{Y}(s_1)$,

$$\overline{\sigma}_\delta(b, s) \leq (1 - \alpha)\phi(b_1, s) + O(\epsilon) \text{ and } \underline{\beta}_\delta(b, s) \geq \alpha\phi(b_1, s) + O(\epsilon),$$

and

$$\underline{\sigma}_\delta(b, s) \geq (1 - \alpha)\phi(b_1, s) + O(\epsilon) \text{ and } \overline{\beta}_\delta(b, s) \leq \alpha\phi(b_1, s) + O(\epsilon).$$

Proof. The proof for this statement is completely symmetric to that of claim i). ▌

iii) For any $(b, s) \in [b_1 - \epsilon, b_1] \times [s_1 - \epsilon, s_1]$,

$$\overline{\sigma}_\delta(b, s) \leq (1 - \alpha)\phi(b_1, s_1) + O(\epsilon) \text{ and } \underline{\beta}_\delta(b, s) \geq \alpha\phi(b_1, s_1) + O(\epsilon),$$

and

$$\underline{\sigma}_\delta(b, s) \geq (1 - \alpha)\phi(b_1, s_1) + O(\epsilon) \text{ and } \overline{\beta}_\delta(b, s) \leq \alpha\phi(b_1, s_1) + O(\epsilon).$$

Proof. The proof again mirrors Step 2 of Lemma B2. Since no investment path from this region leads to $\mathcal{X}(b_1) \times \mathcal{Y}(s_1)$ by Proposition 2, any investment path into outside of this region leads to one of the regions treated in i) and ii). Using the characterization in these cases, one can proceed precisely as in Step 2 of Lemma B2. ▌

Combining i), ii) and iii) with Lemma 2, we conclude that, for a given $\epsilon > 0$,

$$\overline{\sigma}_\delta(b, s) \leq (1 - \alpha)\phi(b_1, s_1) + O(\epsilon) \text{ and } \overline{\beta}_\delta(b, s) \leq \alpha\phi(b_1, s_1) + O(\epsilon),$$

for all $(b, s) \in (b_1 - \epsilon, b_1 + \epsilon) \times (s_1 - \epsilon, s_1 + \epsilon)$, for any $\delta \in [0, 1)$. In particular, since the upper bounds are independent of δ, we have, for any given $\epsilon > 0$,

$$\sup_{\delta \in [0,1)} \overline{\sigma}_\delta(b, s) \leq (1 - \alpha)\phi(b_1, s_1) + O(\epsilon) \text{ and } \sup_{\delta \in [0,1)} \overline{\beta}_\delta(b, s) \leq \alpha\phi(b_1, s_1) + O(\epsilon).$$

Since $O(\epsilon) \to 0$ as $\epsilon \to 0$, we have

$$\limsup_{(b,s) \to (b_1, s_1)} \sup_{\delta \in [0,1)} \overline{\sigma}_\delta(b, s) \leq (1 - \alpha)\phi(b_1, s_1) \text{ and } \limsup_{(b,s) \to (b_1, s_1)} \sup_{\delta \in [0,1)} \overline{\beta}_\delta(b, s) \leq \alpha\phi(b_1, s_1),$$

as needed. ▌