Review of Economic Studies (1999) **66**, 57–82 0034-6527/99/00040057$02.00

Complexity and Renegotiation: A Foundation for Incomplete Contracts

ILYA SEGAL
University of California at Berkeley

First version received June 1995; *final version accepted June* 1998 (*Eds*.)

The paper considers a hold-up model where only one of n future trading opportunities will prove to be efficient, and where *ex post* renegotiation of the *ex ante* contract cannot be prevented. As the environment becomes more complex ($n\to\infty$), the outcome under any message-contingent long-term contract converges to that of the "incomplete contracting" model where trade is contractible *ex post*, but not *ex ante*. When trades are costly to describe, both *ex ante* and *ex post*, the incomplete contracting result is extended to a broader class of environments.

1. INTRODUCTION

According to the modern paradigm of complete contracting which has grown out of implementation theory and incentive theory, optimal contracts should precisely specify an outcome conditional on all relevant information that is observable by a court (*i.e. verifiable*). Moreover, if some relevant information is observed by some of the contracting parties but not by a court, the optimal contract will be contingent on the parties' announcements of this information (see *e.g*. Laffont and Maskin (1982), Moore (1992)). This implies, in particular, that the parties would never prefer to leave some variables open to future negotiation.

While some contracts do seem to conform to these predictions,[1] most real-life contracts are far less sophisticated. This observation has recently provoked significant interest in so-called *incomplete contracting* models, which assume that certain "transaction costs" prevent some aspects of the future trade from being contracted *ex ante* (while still allowing the parties to contract on these aspects *ex post*). Thus, the parties have to leave future outcomes open to future renegotiation, which compromises their incentives to undertake *ex ante* relationship-specific investments (Williamson (1975)).[2] An important application

1. For example, coal and natural gas contracts often precisely specify quantities and prices for up to fifty years ahead, possibly conditional on such relevant information as the international price of a similar product and verifiable measures of the seller's cost (Joskow (1988), Crocker and Masten (1988)). Some contractual clauses, such as unilateral options, make the outcome indirectly contingent on the parties' non-verifiable valuations for the trade by giving them choices ("messages") within the contractual framework. For example, a party may be given the right to opt out of the contract by paying some contractually specified damages (as under the widely used "take or pay" clause—see Schwartz (1992)), and/or to choose the quantity to trade at a contractually specified price (within contractually specified limits).

2. Fudenberg *et al*. (1990) and Rey and Salanie (1990) offer sets of assumptions which ensure that the outcome of an optimal long-term contract can be replicated with short-term contracts and subsequent renegotiation. These assumptions, in particular, require that at every renegotiation stage a *verifiable* sufficient statistic for the past actions given all the potentially observable information be produced. This requirement is not satisfied in hold-up models, where investments, even though observable at the renegotiation stage, are not verifiable.

of this approach has been to study asset ownership as an incomplete contract which affects *ex ante* investment incentives by determining the parties' bargaining positions in *ex post* renegotiation (Hart, (1995)). Since its emergence, however, the incomplete contracts literature has been under attack for its failure to provide a rigorous foundation for "transaction costs" (see *e.g*. Maskin and Tirole (1999)).[3]

This paper takes a step towards explaining the prevalence of simple contracts. It considers a variation of the hold-up model, which is a staple of the incomplete contracting literature (for early treatments, see Williamson (1975, 1985), Grout (1984), Hart and Moore (1988)). It departs from existing hold-up models by introducing a new variable, complexity of the contracting environment, which is defined as the number of potentially relevant future trade opportunities. As the environment's complexity grows without a bound, the outcome in our hold-up model converges to that of the "incomplete contracting" model where trade is contractible *ex post*, but not *ex ante*. This happens through two channels: by reducing the benefit of writing the optimal complete contract, and/or by raising the cost of writing such a contract.

The vanishing of contracting gains in a complex hold-up environment is due to the fact that inefficient future trades, while not affecting the total surplus available, introduce additional incentive constraints. Each inefficient future trade is a potential hold-up opportunity: each party may claim that it is the efficient trade, in the hope of extracting a greater share of *ex post* surplus. The greater is the number of inefficient trades, the more incentive constraints need to be satisfied to prevent the parties from misrepresenting the efficient trade, and the lower is the value of contracting. This effect is illustrated in a particular environment in which incentive constraints asymptotically rule out any improvement over the no-contract outcome.

In more general environments, incentive constraints by themselves do not necessarily preclude implementation of the first-best. However, improving investment incentives in a complex environment may require a contract in which the parties describe an unbounded number of trades to the court. This highlights the second effect of environmental complexity: the cost of certain complex contracts in such an environment may be prohibitive. To capture this effect formally, it is assumed that (i) trades cannot be described in advance, in the sense of Maskin and Tirole (1999), and (ii) at most a finite number of trades can be described in the *ex post* message game. These assumptions are reasonable as long as describing a trade involves a positive cost or delay. Under these assumptions, asymptotic improvement upon the no-contract outcome is impossible as the environment's complexity grows without a bound.

The rest of the paper is organized as follows. Section 2 presents a model of a complex hold-up environment and its preliminary analysis. Section 3 considers a particular case of the model where incentive constraints asymptotically result in incomplete contracting. Section 4 considers the general case, incorporating a restriction on contractual complexity. Section 5 relates the paper's results to existing justifications of contractual incompleteness and suggests some directions for future research.

2. THE MODEL

Consider a model where a buyer and a seller can trade one (but no more than one) of n widgets, indexed 1 to n. Trade is represented by a number $w \in \{0, 1, \ldots, n\}$, where $w = 0$ stands for "no trade". Let $\boldsymbol{v} = (v_0, v_1, \ldots, v_n)$ denote the vector of the buyer's valuations

3. See Section 5 for a critique of existing justifications for incomplete contracting.

for the trades, and $\boldsymbol{c} = (c_0, c_1, \ldots, c_n)$ denote the vector of the seller's costs, with the convention that $v_0 = c_0 = 0$ (throughout the paper, boldface letters will be used to denote vectors). Assume that $\boldsymbol{v} \in V^{n+1}$, $\boldsymbol{c} \in C^{n+1}$, where V and C are compact subsets of $\Re$.

Both vectors depend stochastically on the parties' *ex ante* relationship-specific investments. Let $\sigma \in \Sigma$ denote the cost of the seller's investment, and $\lambda \in \Lambda$ denote the cost of the buyer's investment, where Λ and Σ are compact subsets of $\Re$. Let $\boldsymbol{v}(\lambda, \zeta)$, $\boldsymbol{c}(\sigma, \zeta)$ denote the parties' valuations as functions of their investments and the realization of uncertainty $\zeta \in Z$, where Z is a measurable set in a Euclidean space. Assume that these valuations are observed by the parties *ex post*.

If the parties' *ex post* valuations are $\boldsymbol{v}$, $\boldsymbol{c}$, the total *ex post* surplus available to the parties is

$$S(\boldsymbol{v}, \boldsymbol{c}) = \max_{w \in \{0,1,\ldots,n\}} [v_w - c_w].$$

Assume that $E_\zeta[S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta))]$ is a continuous function of λ, σ.

Assuming that the maximum *ex post* surplus is always realized, a socially efficient investment pair (λ^*, σ^*) must maximize the expected *ex ante* surplus, *i.e.*

$$(\lambda^*, \sigma^*) \in \arg\max_{(\lambda,\sigma) \in \Lambda \times \Sigma} E_\zeta S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta)) - \lambda - \sigma. \tag{1}$$

Let $\mathscr{E}_n^*$ denote the set of such "first-best" investment pairs (λ^*, σ^*).

The parties' individual investment incentives, on the other hand, will be affected by the *ex ante* contract they write. The contract can specify future trade as a function of verifiable information, and, possibly, the parties' messages about non-verifiable information. It is assumed, however, that the parties cannot commit not to renegotiate contractual inefficiencies:

Assumption 1. Whenever the contractual outcome is not *ex post* efficient, the parties renegotiate to an efficient outcome, splitting the renegotiation surplus equally.[4]

If the contract specifies trading a widget w at a price t, the renegotiation surplus is $S(\boldsymbol{v}, \boldsymbol{c}) - (v_w - c_w)$, and the parties' post-renegotiation utilities U_B, U_S can be calculated as

$$\begin{aligned} U_B &= -t + \frac{v_w + c_w}{2} + \frac{1}{2} S(\boldsymbol{v}, \boldsymbol{c}), \\ U_S &= t - \frac{v_w + c_w}{2} + \frac{1}{2} S(\boldsymbol{v}, \boldsymbol{c}). \end{aligned} \tag{2}$$

For example, suppose that the parties abstain from writing an *ex ante* contract (equivalently, they write a contract specifying $w = 0$). In this case, each party's *ex post* payoff will be $\frac{1}{2} S(\boldsymbol{v}, \boldsymbol{c})$. Therefore, an investment pair (λ^0, σ^0) constitutes a Nash equilibrium of the *ex ante* investment game without a contract if, and only if,

$$\begin{aligned} \lambda^0 &\in \arg\max_{\lambda \in \Lambda} \tfrac{1}{2} E_\zeta S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma^0, \zeta)) - \lambda, \\ \sigma^0 &\in \arg\max_{\sigma \in \Sigma} \tfrac{1}{2} E_\zeta S(\boldsymbol{v}(\lambda^0, \zeta), \boldsymbol{c}(\sigma, \zeta)) - \sigma. \end{aligned} \tag{3}$$

4. Equal split of renegotiation surplus is assumed only to simplify exposition. It is straightforward to verify that all of the paper's results hold if renegotiation splits the surplus in any fixed proportion.

Let $\mathscr{E}_n^0$ denote the set of such "no-contract" investment pairs (λ^0, σ^0).

In general, without a contract one should expect underinvestment relative to the first-best, since each party internalizes only 50% of its investment's contribution to *ex post* surplus. The following result formalizes this intuition using the techniques of monotone comparative statics (Milgrom and Shannon (1994), Edlin and Shannon (1998)).[5]

Proposition 1. *Choose any* $(\lambda^0, \sigma^0) \in \mathscr{E}_n^0$ *and any* $(\lambda^*, \sigma^*) \in \mathscr{E}_n^*$. *Suppose that* $E_\zeta[S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta))]$ *has positive partial derivatives on* $\Lambda \times \Sigma$. *Then*

(i) *If* $\Lambda = \{\bar{\lambda}\}$ (*the buyer has no investment choice*) *and* $\sigma^* \in \operatorname{int} \Sigma$, *then* $\sigma^0 < \sigma^*$.
(ii) *If* $\Sigma = \{\bar{\sigma}\}$ (*the seller has no investment choice*) *and* $\lambda^* \in \operatorname{int} \Lambda$, *then* $\lambda^0 < \lambda^*$.
(iii) *If there exists an ex post efficient trade* $w^*(\zeta) \in \arg\max_w [\boldsymbol{v}_w(\lambda, \zeta) - \boldsymbol{c}_w(\sigma, \zeta)]$ *which is independent of* (λ, σ), *and* $\lambda^* \in \operatorname{int} \Lambda$, $\sigma^* \in \operatorname{int} \Sigma$, *then* $(\lambda^0, \sigma^0) < (\lambda^*, \sigma^*)$.

According to parts (i) and (ii) of the proposition, in the absence of a contract, each party underinvests *given the other party's investment*.[6] Therefore, in the absence of a contract, a socially optimal investment pair cannot be sustained. As it turns out, however, first-best investments can be implemented with a contract which prevents renegotiation in equilibrium by specifying a widget that is *ex post* efficient given these investments, at a price that is independent of the parties' investments:

Proposition 2. *If* $(\lambda^*, \sigma^*) \in \mathscr{E}_n^*$, *then* (λ^*, σ^*) *is an equilibrium investment pair in a contract in which a widget* $w^*(\zeta) \in \arg\max_w [v_w(\lambda^*, \zeta) - c_w(\sigma^*, \zeta)]$ *is traded at a price* $t(\zeta)$.

The proof is based on the following simple argument. Observe that each party's investment does not affect the other party's utility at the contractual outcome, which serves as the disagreement point for renegotiation. Therefore, if (for example) the buyer had all the bargaining power, he would be the residual claimant for *ex post* surplus, so he would find any deviation from a first-best investment level unprofitable. When the buyer has only partial bargaining power, his *ex post* payoff after any deviation is even lower (while his equilibrium payoff is not affected, since in equilibrium there is no renegotiation). Therefore, any deviation would remain unprofitable.[7]

According to Proposition 2, if the *ex post* efficient trade is either known *ex ante* or verifiable *ex post*, the hold-up problem can be easily resolved. To eliminate this possibility, two assumptions are made:

Assumption 2. The parties' investments λ and σ, their valuation vectors $\boldsymbol{v}$ and $\boldsymbol{c}$, and the realization of uncertainty ζ are not verifiable.

5. All proofs are in the Appendix.

6. If investments are substitutes in expected total surplus, one party's underinvestment encourages the other party to invest more. If this indirect effect outweighs the direct effect, one party (but not both) may invest more without a contract than in the first best.

7. The argument fails when a party's investment directly affects the other party's valuation (for example, the seller's investment improves non-verifiable quality of the good). Che and Hausch (1997) show that in this case the parties cannot improve upon the no-contract outcome even when the efficient *quantity* of trade is known *ex ante*. Their model is similar to ours in that contractual incompleteness is due to a court's inability to distinguish between "widgets" of different qualities. (If quality could be contracted in advance, investment would have no direct external effect, and Proposition 2 would apply.)

Assumption 3. For any pair of vectors $\boldsymbol{v}$ and $\boldsymbol{c}$, and a permutation π of the set $\{1, \ldots, n\}$ of widgets, define new valuation vectors $\boldsymbol{v}\pi$ and $\boldsymbol{c}\pi$ such that for all $w \in \{1, \ldots, n\}$,

$$(\boldsymbol{v}\pi)_w = v_{\pi(w)},$$

$$(\boldsymbol{c}\pi)_w = c_{\pi(w)}.$$

Then for any $\lambda \in \Lambda$ and $\sigma \in \Sigma$, the distributions of $\boldsymbol{v}(\lambda, \zeta)\pi$ and $\boldsymbol{c}(\sigma, \zeta)\pi$ do not depend on the permutation π.

Assumption 2 implies, in particular, that the efficient trade is not verifiable *ex post*. Assumption 3 postulates that any permutation of widgets does not affect the parties' priors. Intuitively, this means that any *ex ante* structure on the set of possible trades is lacking, *i.e.* widgets are indistinguishable *ex ante*. In particular, all trades are *ex ante* equally likely to be efficient.[8]

The next assumption will rule out asymptotic improvement by a contract that specifies a noncontingent trade $w \in \{1, \ldots, n\}$ at a fixed price t_0. By Assumption 3, for all $w > 0$ we have $E_\zeta v_w(\lambda, \zeta) = 1/n \sum_{i=1}^n E_\zeta v_i(\lambda, \zeta)$ and $E_\zeta c_w(\sigma, \zeta) = 1/n \sum_{i=1}^n E_\zeta c_i(\sigma, \zeta)$, which implies that specifying a noncontingent trade gives each party the same expected utility as a randomization that picks every widget with probability $1/n$. This expected utility is assumed to be insensitive to investments as $n \to \infty$:

Assumption 4. The average trade is asymptotically non-responsive to investments: for every $\zeta \in Z$, $\lambda, \lambda' \in \Lambda$, and $\sigma, \sigma' \in \Sigma$, as $n \to \infty$,

$$\frac{1}{n} \left| \sum_{w=1}^n [v_w(\lambda', \zeta) - v_w(\lambda, \zeta)] \right| \to 0,$$

$$\frac{1}{n} \left| \sum_{w=1}^n [c_w(\sigma', \zeta) - c_w(\sigma, \zeta)] \right| \to 0.$$

Using expressions (2), this assumption implies that in any noncontingent contract, each party asymptotically receives 50% of its investment's contribution to total surplus, just as it would without a contract.

So far we have not ruled out the possibility that the parties would reveal the efficient trade (or another trade that is responsive to investments) in a message-contingent contract. However, expressions (2) imply that given a fixed price, each party will have an incentive to "hold up" the other party by suggesting a trade that improves its bargaining position: the buyer will suggest a widget w that maximizes $(v_w + c_w)/2$, and the seller will suggest a widget that minimizes $(v_w + c_w)/2$. Assume that there are many such "hold-up" widgets, which are not responsive to investments:

Definition 1. *Widget w is non-responsive to investments in contingency ζ if $v_w(\lambda, \zeta)$ and $c_w(\sigma, \zeta)$ do not depend on λ and σ respectively* (*and otherwise it is said to be responsive to investments in contingency ζ*).

8. Observe that the uncertainty postulated in Assumption 3 is different in nature from the uncertainty about the *ex post* optimal *quantity* of trade, considered *e.g.* in Edlin and Reichelstein (1996). In their model, *ex ante* uncertainty about the optimal trade does not necessarily prevent implementation of first-best investments. For example, under a certain separability assumption, both parties can be made to invest efficiently using a contract that specifies the expected efficient trade given optimal investments. In our model, different trades cannot be ranked *ex ante* as representing higher or lower quantities.

Definition 2. *Widget w is good* (G) [*bad* (B)]*in contingency* ζ *if*
(a) *it is non-responsive to investments in contingency* ζ, *and*
(b) $[v_w(\lambda, \zeta) + c_w(\sigma, \zeta)]/2 > [<][v_i(\lambda, \zeta) + c_i(\sigma, \zeta)]/2$ *for all responsive widgets i an contingency* ζ, *and for all* $(\lambda, \sigma) \in \Lambda \times \Sigma$.

An example of a good (G) widget would be a "goldplated" widget, which is worth a lot to the buyer, but is very costly to produce. A bad (B) widget, on the other hand, could be a "cheap imitation" widget, which costs little to the seller, and is worth little to the buyer. Expression (2) implies that given a fixed price t, the buyer would prefer suggesting a G widget rather than a responsive widget, while the seller would prefer suggesting a B widget rather than a responsive widget.[9] In order to make revelation of responsive widgets difficult, assume that the numbers of G and B widgets are large:

Assumption 5. The numbers of G and B widgets go to infinity in probability[10] as $n \to \infty$.

3. INCENTIVE CONSTRAINTS AND CONTRACTING BENEFITS

This section analyzes a particular model satisfying Assumptions 1–5, in which incentive constraints asymptotically reduce contracting benefits to zero. In this model, in every contingency there is exactly one responsive widget, which will be called widget R (for "regular"). This responsive widget is assumed to always represent the *ex post* efficient trade. All remaining widgets will represent hold-up opportunities (either G or B). The *ex ante* probability of each widget being the R widget is $1/n$, and the types of the remaining widgets are distributed i.i.d., with the probability that a given non-responsive widget is G rather than B being equal to 1/2. Call the vector $N \in \{R, G, B\}^n$ of the realized types of all widgets the *configuration*.

Let $v \in V$ denote the buyer's valuation for the R widget, and $c \in C$ denote the seller's cost for this widget, where V, C are assumed to be compact subsets of $\Re$. Assume that for all B widgets b, for all G widgets g, and for any $v \in V$, $c \in C$, we have

$$v - c > \max\{0, v_b - c_b, v_g - c_g\},$$
$$\underline{a} = \frac{v_b + c_b}{2} < \frac{v + c}{2} < \frac{v_g + c_g}{2} = \bar{a}. \tag{4}$$

9. Truthful revelation of the efficient widget would be easier to achieve if renegotiation could be prevented, or if the split of bargaining power in renegotiation could be affected by the *ex ante* contract. For example, if the contract gave the seller the right to make an *ex post* take-it-or-leave-it offer to the buyer, the seller would become the residual claimant for total surplus. This would give the seller an incentive to suggest the efficient widget *ex post*, and to invest efficiently *ex ante*. More generally, Proposition 5 of Rogerson (1992) establishes that in the absence of renegotiation, first-best investments can be implemented using multi-stage *ex post* message games.

In a related vein, Aghion *et al*. (1994) show that if the renegotiation game uses the original contract as an outside option, rather than as the disagreement point, the contract can make the responder accept an offer leaving it no surplus over the contractual outcome (*e.g*., by imposing per diem penalties or using a financial hostage), which would would again induce the proposer to choose a first-best investment level.

10. Convergence in probability means that for any fixed number m, the probability that the numbers of G and B widgets exceed m goes to one as $n \to \infty$.

The first inequality says that the R widget is always the efficient widget to trade. To understand the second line, recall that, according to (2), specifying the trade of widget w in a contract provides the transfer $(v_w + c_w)/2$ from the seller to the buyer. The second line of (4), therefore, says that all B widgets provide the same transfer $\underline{a}$, all G widgets provide the same transfer a, and that the R widget provides an intermediate transfer (which is consistent with Definition 2).

Since the efficient trade in this model does not depend on investments, Proposition 1(iii) establishes that in the absence of a contract, both parties underinvest. For convenience, assume that the distribution of the parties' valuations for the efficient trade, v and c, does not depend on n. This assumption ensures that neither the set $\mathscr{E}^*$ of first-best investments nor the set $\mathscr{E}^0$ of no-contract investments depend on n. At the same time, the efficiency improvement achieved by writing an *ex ante* contract will depend on n.

As discussed earlier, Assumptions 3 and 4 imply that a contract specifying a noncontingent trade will provide no asymptotic improvement over the no-contract outcome as $n \to \infty$. An improvement may be achieved with message-contingent contracts which provide the parties with the incentives to reveal the efficient trade. When n is small, message-contingent contracts in which the equilibrium price depends on the realized configuration may substantially improve upon noncontingent contracts:

Example 1. Let $n = 2$. Suppose that $\underline{a} = 0 < v + c < \bar{a}$ for all $v \in \bar{V}$, $c \in \bar{C}$ (observe that this requires a to be at least twice as large as stipulated in (4)). Then the following message-contingent contract implements the first-best:

1. The buyer proposes a widget to trade.
2. The seller agrees or challenges.
3. If the seller agrees, the widget proposed by the buyer is traded at t_0. If the seller challenges, he chooses which widget to trade, and the price is $t_0 - a/2$.
4. The contractual outcome can be renegotiated.

If the configuration is BR or RB, the seller will never challenge: according to (2), challenging the R widget and offering the B widget instead saves him $(v + c)/2$, which is less than the penalty $a/2$ he pays for challenging. Therefore, using (2) again, we see that the buyer will propose the R widget.

If the configuration is GR or RG, the seller will always challenge the G widget and offer the R widget instead, since according to (2) this saves him $a - (v + c)/2$, which exceeds the penalty $a/2$ he pays for challenging. The buyer will nevertheless propose the G widget, since this allows him to extract the penalty from the seller.[11]

Therefore, in both cases, the R widget is traded in equilibrium, and the trading price does not depend on v, c. Thus, by Proposition 2 the first-best is achieved.

To characterize the set of incentive-compatible contracts, the model can be cast in the framework of implementation with renegotiation developed by Maskin and Moore (1999). An *ex post* state of the world can be described by a triple $(\boldsymbol{N}, v, c)$, where v is the buyer's valuation for the R widget, c is the seller's cost for this widget,and $\boldsymbol{N}$ is the

11. Formally speaking, this argument derives the subgame-perfect equilibrium of the message game. However, since this is a zero-sum game (renegotiation always restores efficiency), all the Nash equilibrium outcomes are also subgame-perfect equilibrium outcomes. This point also applies to Examples 2 and 3 in Section 4 below.

configuration. Let Θ denote the set of all such states, *i.e.*

$$\Theta = \{\boldsymbol{N} \in \{R, G, B\}^n\text{: for some } i,\ N_i = R \text{ and } \boldsymbol{N}_j \neq \text{R for all } j \neq i\} \times \bar{V} \times \bar{C}.$$

An *ex post* outcome can be described as $(\boldsymbol{x}, t)$, where $\boldsymbol{x} \in \{(x_0, x_1, \ldots, x_n) \in \Re_+^{n+1}\text{: } \sum_{w=0}^{n} x_w = 1\}$ is the randomized trade (x_w is the probability that widget w is traded), and $t \in \Re$ is the monetary transfer from the buyer to the seller.[12,13] Let Y denote the set of all such outcomes. Denote the buyer's and seller's utilities from outcome $y \in Y$ in the state $\theta \in \Theta$ by $u_B(y, \theta)$ and $u_S(y, \theta)$ respectively.

The contractual outcome will be renegotiated whenever it is Pareto inefficient. Maskin and Moore consider an arbitrary exogenously given renegotiation function $h\text{: } \Theta \times Y \to Y$, so that $h(y|\theta)$ specifies the final outcome in state θ when the contractually stipulated outcome in that state is y. Maskin and Moore postulate three properties of the renegotiation function:

(R1) The function $h(\cdot)$ is single-valued (*i.e.* the outcome of renegotiation is uniquely defined).
(R2) The outcome $h(y|\theta)$ is strongly Pareto efficient in state θ for all $\theta \in \Theta$ and all $y \in Y$.
(R3) Renegotiation is individually rational, *i.e.* $u_j(h(y|\theta), \theta) \geqq u_j(y, \theta)$ for $j \in \{B, S\}$, all $\theta \in \Theta$, and all $y \in Y$.

In our model, the renegotiation function $h(\cdot)$ is given by the Nash bargaining solution, which always implements the *ex post* efficient trade, splitting the renegotiation surplus equally between the parties. This renegotiation function clearly satisfies properties (R1)–(R3).

The objective of an *ex ante* contract is to implement a given *choice rule* $f\text{: } \Theta \to Y$ as the post-renegotiation outcome. In general, this can be achieved via a message game. A general message game can be described as $M = \langle M_B, M_S, g \rangle$, where M_B, M_S are the parties' message spaces, and $g\text{: } M_B \times M_S \to Y$ is the outcome function. The message game M is said to implement a choice rule f with renegotiation h if in any state $\theta \in \Theta$ the unique Nash equilibrium outcome of $h \circ g$ is $f(\theta)$. A choice rule is implementable if there exists a message game implementing it. Observe that any social choice rule that is implementable with renegotiation has to be strongly Pareto efficient (any Pareto inefficient outcome would be renegotiated away). The following theorem characterizes social choice rules that are implementable with renegotiation in a two-party environment:

Theorem 0. (Maskin and Moore (1999)) *A strongly Pareto efficient choice rule f is implementable with renegotiation h satisfying assumptions* (R1)–(R3) *if, and only if, for any ordered pair of states* $\theta, \phi \in \Theta$ *there exists an outcome* $y \in Y$ *such that*

$$u_B(f(\theta), \theta) \geqq u_B(h(y|\theta), \theta) \quad \textit{and} \quad u_S(f(\phi), \phi) \geqq u_S(h(y|\phi), \phi).$$

The intuition behind the theorem runs as follows. Since the post-renegotiation outcome is always *ex post* efficient, "contract plus renegotiation" is a game with opposing interests, hence the equilibrium payoffs (and the equilibrium outcome) are unique by a

12. Randomization will play no role in the analysis of this section, but it will play a role in the next section. Allowing for randomized monetary transfers would not change anything because the risk-neutral parties care only about expected transfers.

13. The contract could also specify destruction of wealth when the parties disagree. However, since this inefficiency would be renegotiated in the same way in any state of the world, it would not affect the parties' incentives in any way.

version of the Minimax theorem.[14] Therefore, one should not worry about multiplicity of equilibria, and the Revelation Principle can be used to restrict attention to so-called direct revelation mechanisms. In such mechanisms the parties' messages to the court are complete reports of what they know about the state of the world (*i.e.* $M_B = M_S = \Theta$) and the parties do not lie in equilibrium.

Note that we can restrict attention to contracts that prescribe the choice rule directly when both parties agree on the state of the world. When the parties disagree, on the other hand, we know that one party is lying, but we do not know which one. Therefore, in the case of a disagreement, the mechanism must specify an outcome y that punishes the lying party, whoever it is. Informally, we will say that such an outcome "resolves" the disagreement. The requirement that all potential disagreements be resolved yields the incentive-compatibility constraints described in Theorem 0.

Application of Theorem 0 to our model is not straightforward because of the daunting number of possible disagreements. Fortunately, we will be able to obtain an upper bound on what can be achieved in an incentive-compatible contract using only a small subset of all possible disagreements. Intuitively, we can do this because some disagreements are much easier to resolve than others.

Suppose, for example, that a certain widget is described as bad by the seller and as good by the buyer. Stipulating the trade of this widget would then "call the bluff" by punishing the liar, whoever it is. Indeed, if the seller is lying and the buyer is telling the truth, the seller will have to provide a good widget, which will punish him severely; if instead the buyer is lying and the seller is telling the truth, the buyer will receive a bad widget, which will punish him as well. More generally, if there is a widget that is described as more valuable by the buyer than by the seller, stipulating the trade of this widget would easily resolve the disagreement.

The disagreements that are hard to resolve are those in which no widget is described as more valuable by the buyer than by the seller. We will focus on a set of such disagreements, which will be called *critical disagreements*. An ordered pair of configurations $(\boldsymbol{M}, \boldsymbol{N})$ is related by critical disagreement if there exist widgets i, j such that

$$\begin{aligned} \boldsymbol{M}_i &= R, \boldsymbol{N}_i = B, \\ \boldsymbol{M}_j &= G, \boldsymbol{N}_j = R, \\ \boldsymbol{M}_w &= \boldsymbol{N}_w \quad \text{for } w \notin \{i, j\}. \end{aligned} \tag{5}$$

In words, $\boldsymbol{M}$ is in critical disagreement with $\boldsymbol{N}$ when the widget described as regular in $\boldsymbol{N}$ is described as good in $\boldsymbol{M}$, the widget described as regular in $\boldsymbol{M}$ is described as bad in $\boldsymbol{N}$, and the two configurations agree on the types of all the other widgets.

Let $\mathfrak{D}$ denote the set of ordered configuration pairs related by critical disagreement. The relation of critical disagreement can be represented with a directed graph, whose vertices are configurations, and whose edges are pairs from $\mathfrak{D}$. To draw this graph, let $|\boldsymbol{N}|$ denote the number of B widgets in configuration $\boldsymbol{N}$, and note that $(\boldsymbol{M}, \boldsymbol{N})$ can be related by critical disagreement only if $|\boldsymbol{N}| = |\boldsymbol{M}| + 1$. Therefore, it is convenient to draw the graph of critical disagreements by levels, configurations with $|\boldsymbol{N}| = l$ being located at the lth level, and the edges connecting vertices from neighbouring levels. An example for $n = 3$ is drawn in Figure 1.

The proof of this section's result can now be sketched as follows (a complete proof is in the Appendix). Suppose we want to implement a choice rule $(\boldsymbol{x}(\boldsymbol{N}, v, c),\ t(\boldsymbol{N}, v, c))$,

14. Moreover, in our context of transferable utility "contract plus renegotiation" is a zero-sum game, therefore the equilibrium outcome is unique even when mixed strategies are allowed.

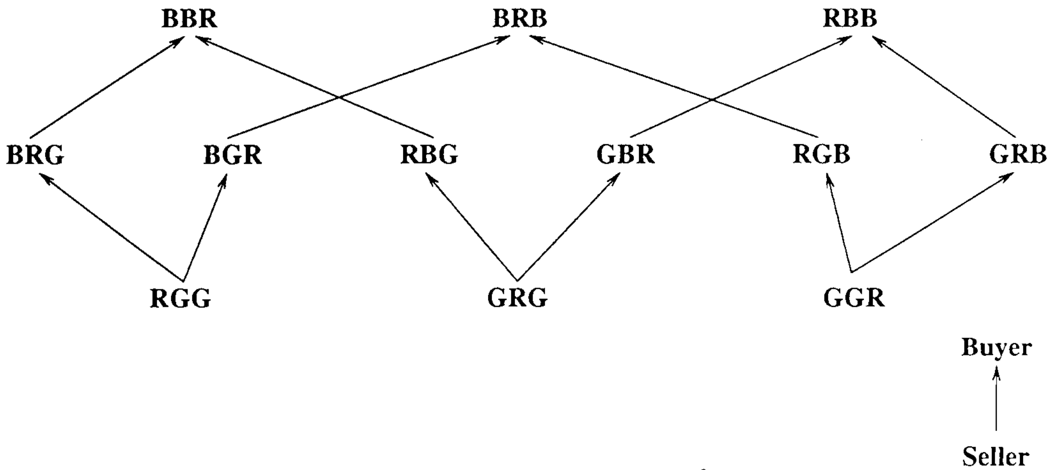

FIGURE 1. The graph of critical disagreements for $n = 3$

where $t(\boldsymbol{N}, v, c)$ is the monetary transfer and $\boldsymbol{x}(\boldsymbol{N}, v, c)$ is the randomized trade prescribed for the state of the world $(\boldsymbol{N}, v, c)$. Pareto efficiency of the choice rule dictates that $x_w(\boldsymbol{N}, \boldsymbol{v}, \boldsymbol{c}) = 1$ for $N_w = R$, but leaves us the freedom to design the transfer function $t(\boldsymbol{N}, v, c)$. Consider the restrictions on the transfer function that are imposed by Maskin–Moore incentive constraints associated with critical disagreements. Namely, suppose that the buyer's announcement is $\theta_B = (\boldsymbol{N}, v_B, c_B)$, while the seller's announcement is $\theta_S = (\boldsymbol{M}, v_s, c_S)$, where $(\boldsymbol{M}, \boldsymbol{N}) \in \mathfrak{D}$, and look for an outcome $(\boldsymbol{x}, t)$ that resolves the disagreement.

Intuitively, trading the widget that is described as R by the seller or the widget that is described as R by the buyer would not be very useful for resolving the critical disagreement, since each widget is ascribed a higher value by the seller than by the buyer. Trading another widget, on whose characteristics the parties agree, can be shown to be equivalent to specifying "no trade" (*i.e.* $x_0 = 1$). In either case, using (2), the incentive constraints of Theorem 0 can be written as

$$-t + \frac{v_S - c_S}{2} \leqq v_S - t(\boldsymbol{M}, v_S, c_S),$$

$$t + \frac{v_B - c_B}{2} \leqq t(\boldsymbol{N}, v_B, c_B) - c_B.$$

Adding the two inequalities, we obtain

$$t(\boldsymbol{N}, v_B, c_B) - t(\boldsymbol{M}, v_S, c_S) \geqq \frac{v_B + c_B}{2} - \frac{v_S + c_S}{2}. \tag{6}$$

Suppose first that the transfer function $t(\boldsymbol{N}, v, c)$ does not depend on the configuration $\boldsymbol{N}$, *i.e.* $t(\boldsymbol{N}, v, c) = t(v, c)$. In that case, by switching (v_B, c_B) and (v_S, c_S) in inequality (6), we can also write

$$t(v_B, c_B) - t(v_S, c_S) \leqq \frac{v_B + c_B}{2} - \frac{v_S + c_S}{2}.$$

Together with (6), this implies that

$$t(v_B, c_B) - t(v_S, c_S) = \frac{v_B + c_B}{2} - \frac{v_S + c_S}{2}.$$

This means that $t(v, c) \equiv (v + c)/2 + \text{const}$; that is, up to a constant, the transfer equals that obtained without an *ex ante* contract. Such a transfer rule, of course, would not improve investment incentives upon the no-contract outcome.

This argument shows that the only hope to improve incentives upon the no-contract outcome lies in having the transfer rule depend on the configuration. Substituting $(v_B, c_B) = (v_S, c_S)$ in inequality (6), we see that in an incentive-compatible contract, the transfer has to be non-decreasing along a critical disagreement. Having the transfer increase along critical disagreements may help improve incentives, as can already be seen from Example 1. In that example, the buyer is able to extract a fine $a/2$ from the seller when the configuration contains no B widgets, but not when the configuration contains a B widgets. The fact that the transfer is increasing in the number of B widgets helps create the incentives for the parties to reveal the R widget. To understand this, observe that when the transfer is increasing along critical disagreements, the buyer is less willing to enter into a in critical disagreement with the seller: when he is believed, he pays a higher transfer than that implied by the seller's report. Symmetrically, the seller in that case is less willing to enter into a in critical disagreement with the buyer: when he is believed, he receives a lower transfer than that implied by the buyer's report.

At the same time, it can be shown that the transfer difference between *any* two configurations $\boldsymbol{N}'$, $\boldsymbol{N}''$ is bounded by $A = \bar{a} - \underline{a}$. Indeed, if we had $t(\boldsymbol{N}', v, c) - t(\boldsymbol{N}'', v, c) > A$, the disagreement in which the seller reports $(\boldsymbol{N}', v, c)$ and the buyer reports $(\boldsymbol{N}'', v, c)$ could not be resolved by specifying any trade. Intuitively, since the stakes involved in trading different widgets are bounded by A, transfers in different states of the world cannot differ by more than that amount.

Finally, observe from Figure 1 that as n goes to infinity, so does the number of critical disagreements squeezed between the top and the bottom of the graph. Since the transfer difference between the top and the bottom of the graph is bounded by A for any n, this implies that the (properly defined) "average" transfer difference between two configurations related by critical disagreement, $t(\boldsymbol{N}, v, c) - t(\boldsymbol{M}, v, c)$, goes to zero as $n \to \infty$. In the limit, the transfer does not change along the "average" critical disagreement, in which case our earlier discussion suggests that investment incentives converge to those without a contract. Formally, we obtain the following result:

Theorem 1. *In an environment with n widgets, choose an incentive-compatible message contingent contract, and let $U_j^n(\boldsymbol{N}, v, c)$ denote the* ex post *payoff of party $j \in \{B, S\}$ in the contract in the state of the world $(\boldsymbol{N}, v, c)$. Then for all $v, v' \in V$, $c, c' \in C$,*

$$\left| E_{\boldsymbol{N}} U_j^n(\boldsymbol{N}, v', c') - E_{\boldsymbol{N}} U_j^n(\boldsymbol{N}, v, c) - \left(\frac{v' - c'}{2} - \frac{v - c}{2} \right) \right| \leq \frac{A}{2^{n-1}} \left[1 + \binom{n-1}{\left\lceil \frac{n-1}{2} \right\rceil} \right],$$

where $A = \bar{a} - \underline{a}$.

Using Stirling's formula ($n! \sim n^n e^{-n} \sqrt{2\pi n}$), we can see that the right-hand side of the above inequality goes to zero as $n \to \infty$. Therefore, asymptotically for each party j we have

$$E_N U_j^n(\mathbf{N}, v, c) \simeq \frac{v-c}{2} + \text{const},$$

i.e. each party shares 50% of the total *ex post* surplus with the other party, just as it would without a contract.

This result can be translated into convergence of investments to the no-contract investments. Since the set of Nash equilibrium investment pairs (λ, σ) (with or without a contract) may not be single-valued, we need to define convergence of sets. We do this using the notion of Hausdorff distance $\rho(X, Y)$ between two sets X, Y in a metric space, defined as

$$\rho(X, Y) = \max\left\{\sup_{x \in X} \text{dist}(x, Y), \sup_{y \in Y} \text{dist}(y, X)\right\},$$

$$\text{where dist}(x, Y) = \inf_{y \in Y} |x - y|,$$

(Hausdorff (1962)).

Let $\mathscr{E}_n \subset \Lambda \times \Sigma$ denote the set of all investment pairs attainable in a Nash equilibrium of a contract with n widgets. Since a "no trade" contract is always feasible, we have $\mathscr{E}^0 \subset \mathscr{E}_n$ for all n. Also, let $W(\lambda, \sigma)$ denote the total expected surplus given the parties' investments (λ, σ), *i.e*.

$$W(\lambda, \sigma) = E_\zeta[S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta))] - \lambda - \sigma.$$

By assumption, $W(\cdot)$ is a continuous function.

Now we can use Theorem 1 to establish convergence of investments and welfare to the incomplete contracting outcomes:

Corollary 1. *As* $n \to \infty$,

(i) $\rho(\mathscr{E}_n, \mathscr{E}^0) \to 0$,
(ii) $\rho(W(\mathscr{E}_n), W(\mathscr{E}^0)) \to 0$.

Part (ii) implies, in particular, that $\sup W(\mathscr{E}_n) \to \sup W(\mathscr{E}^0)$, *i.e*. the maximum total surplus achievable in an incentive-compatible contract with n widgets converges to the maximum total surplus achieved without a contract as $n \to \infty$.

4. THE ROLE OF CONTRACTING COSTS

Consider two perturbations of the last section's model, in which an incentive-compatible message-contingent contract can approach the first-best for a large n:

Example 2. Consider a model whose only difference from that described in Section 3 is that the number of G widgets is known *ex ante* to be $g(n)$. Then the first-best outcome can be implemented for any n using a game in which the seller can veto $g(n)$ widgets, and the buyer chooses one of the remaining widgets to be traded at a fixed price t_0. The seller will veto the widgets that provide him with the lowest post-renegotiation payoff. According to expressions (2), the seller will veto all the G widgets. Out of the remaining widgets, the buyer will choose the one maximizing his post-renegotiation transfer, which according to (2) will be the R widget. Therefore, in equilibrium the R widget is traded at a fixed price, which according to Proposition 2 gives the parties first-best investment incentives.

Example 3. Consider a model whose only difference from that described in Section 3 is that instead of one there are $[\rho n]$ responsive (R) widgets ($0<\rho<1$), *i.e.* the asymptotic proportion of R widgets is ρ. Then the first-best outcome can be approached asymptotically as $n\to\infty$. Indeed, consider a mechanism in which the seller can veto $[n/2]$ widgets, and the buyer chooses one of the remaining widgets to be traded at a fixed price t_0. This game will result in selecting a widget yielding the median post-renegotiation transfer from the seller to the buyer. The median transfer will be provided by an R widget when the numbers of G and B widgets in the configuration do not differ by more than $[\rho n]$, the number of R widgets. By the Law of Large Numbers, the probability of this event goes to one as $n\to\infty$, and therefore the parties' incentives approach the first-best.

Intuitively, the previous section's result does not apply in Example 2 because no critical disagreements exist when the number of G widgets is known. Whenever the parties disagree on which is the R widget, there exists a widget which is ascribed a higher value by the buyer than by the seller. Trading this widget resolves the disagreement. In Example 3, critical disagreements exist, but their number is much smaller than in the previous section. This allows asymptotic implementation of the first-best.

One feature shared by the two examples is that the message game achieving the first-best involves describing a large number of widgets. In Example 2, under Assumption 5 we have $g(n)\to\infty$ as $n\to\infty$, thus the message game requires the seller to describe an unbounded number of widgets. The message game in Example 3 also involves the seller describing an unbounded number of widgets, $[n/2]$. When n is very large, such contracts would be impractical.

To formalize this idea, we will measure the complexity of a message-contingent contract by the number of widgets that are described in the contract. If the parties write no *ex ante* contract, *ex post* they only need to describe one widget to a court—the one they have renegotiated to trade. A natural question to ask is what can be asymptotically achieved by an *ex ante* contract in which the parties only describe a finite number of widgets to a court. In this section we show that under the assumptions presented in Section 2, asymptotically such contracts cannot improve upon the no-contract outcome.

As in the previous section, we apply the Maskin–Moore concept of implementation with renegotiation to characterize incentive-compatibility constraints, the only modification being that an *ex post* state of the world is now described by a pair $(\boldsymbol{v}, \boldsymbol{c})\in\Re^{2(n+1)}$. In addition to the incentive constraints, however, we now introduce constraints on the content and size of the messages the parties send. For example, direct revelation mechanisms, in which each party describes all of the widgets, will be ruled out due to their complexity. Therefore we need to consider more general mechanisms of the form $M=\langle M_B, M_S, g\rangle$, where M_B, M_S are the parties' message spaces, and $g: M_B\times M_S\to Y$ is the outcome function. For the sake of generality, we allow extensive-form message games, in which case the parties' messages $(m_B, m_S)\in M_B\times M_S$ should be understood as contingent strategies. Write the outcome function as $g(m_B, m_S)=(\boldsymbol{x}(m_B, m_S), t(m_B, m_S))$, where $\boldsymbol{x}(m_B, m_S)$ is the randomized trade prescribed for the message pair $(m_B, m_S)\in M_B\times M_S$, and t is the transfer prescribed for this message pair.

In this setup, we introduce two contracting restrictions. The first restriction, first introduced by Maskin and Tirole (1999), is that widgets cannot be described *ex ante*, because at this point the parties cannot foresee them, and cannot give them names. *Ex post*, once widgets' names are known, the parties can communicate by describing widgets to the court, and by reporting their valuations for the widgets that have been described. Therefore, the likelihood of a given widget being chosen in the message game cannot depend on its name. The message games described in Examples 2 and 3 are two examples of such communication. The parties would be able to specify these games in the *ex ante* contract without using widgets' names.

To formalize this restriction on contractual communication, consider two different numberings of widgets: "old" and "new". Suppose that the old numbering could be obtained from the new numbering by means of a permutation π, *i.e.* the old number of a widget whose new number is $w \in \{1, \ldots, n\}$ is $\pi(w) \in \{1, \ldots, n\}$. Suppose that in the old numbering, the parties' messages are $(m_B, m_S) \in M_B \times M_S$, and the contract prescribes the outcome $(\boldsymbol{x}(m_B, m_S), t(m_B, m_S))$. In the new numbering, the same messages will be represented differently. Denote the parties' messages in the new numbering by $(m_B\pi, m_S\pi) \in M_B \times M_S$. For these messages, the mechanism $g(\cdot)$ will prescribe the outcome $(\boldsymbol{x}(m_B\pi, m_S\pi), t(m_B\pi, m_S\pi))$.

If the parties cannot describe widgets in advance, the contractual outcome cannot depend on the numbering of widgets. Therefore, we require that the physical outcome under the new numbering is exactly the same as the physical outcome under the old numbering. This requirement consists of two parts: (i) the new monetary transfer is the same as the old monetary transfer, and (ii) the new randomization over physical trades is the same as the old randomization over physical trades, which in the new numbering can be written as $(\boldsymbol{x}(m_B, m_S))\pi$. Therefore the parties' inability to describe widgets *ex ante* can be formalized with the following contracting restriction:[15]

(C1) For any message pair $(m_B, m_S) \in M_B \times M_S$, and any permutation π of $\{1, \ldots, n\}$,

$$t(m_B\pi, m_S\pi) = t(m_B, m_S),$$

$$\boldsymbol{x}(m_B\pi, m_S\pi) = (\boldsymbol{x}(m_B, m_S))\pi.$$

Theorem 4 of Maskin and Tirole (1999) implies that in our setup, impossibility to describe widgets *ex ante by itself* places no constraint on what can be achieved in a message-contingent contract. Intuitively, even though widgets cannot be described *ex ante*, the parties can always write a message game describing all the widgets *ex post*. In our Examples 2 and 3, widgets need not be described *ex ante*, but a large number of them is described in the *ex post* message game, and the first-best outcome is achieved.

In our view, however, if widgets are hard to describe *ex ante*, it may also be hard to describe a large number of them *ex post*. Therefore, we impose a second contracting restriction, which says that at most a finite number of widgets can be described *ex post*. Formally, say that a set of widgets $\Omega \subset \{1, \ldots, n\}$ is *undescribed* by a message pair (m_B, m_S) if for any permutation π of widgets that only affects Ω,

$$(m_B\pi, m_S\pi) = (m_B, m_S).$$

This means that the parties' messages do not distinguish between the "undescribed" widgets that constitute Ω.

Observe that when the parties communicate sequentially, the sets of undescribed widgets may depend on the parties' messages. For example, which widgets the buyer chooses to describe may, in general, depend on what the seller has previously reported. Let $\Omega(m_B, m_S)$ denote the (largest) set of undescribed widgets given the message pair (m_B, m_S). Our second contracting restriction can now be formalized as follows:

15. Any mechanism satisfying (C1) implements a choice rule that is *welfare-neutral* as defined in Maskin and Tirole (1999), and *neutral* as defined in Segal (1996). See the two papers for further discussion.

(C2) For any message pair $(m_B, m_S) \in M_B \times M_S$, at most k widgets are described, *i.e.* $|\Omega(m_B, m_S)| \geqq n-k$.[16]

To get some intuition for how restrictions (C1) and (C2) affect the set of implementable choice rules, consider the environment of Example 2. Consider a game in which widgets are sequentially described by one or the other party, and the parties announce the types of the described widgets. If the buyer decides to lie, one strategy for him would be to veto every widget described by the seller, by claiming that it is a "bad" widget. Similarly, if the seller decides to lie, one strategy for him would be to veto every widget described by the buyer, by claiming that it is a "good" widget. Of course, since the numbers of bad and good widgets in the configuration are known in advance, there are only so many widgets each party can veto. Thus, after sufficiently many widgets have been described, eventually there will be named a widget for which the buyer will have to concede that it is not "bad", and the seller will have to concede that it is not "good". Trading this widget would "resolve" the disagreement. However, if the configuration contains large numbers of bad and good widgets, a large number of widgets will have to be described before the bluff is called. If attention is restricted to message games in which a bounded number of widgets is described, for a large n the bluff may never be called. Thus, a disagreement which would not be "critical" in the definition of the previous section (it can in principle be resolved by finding a widget ascribed a higher vaue by the buyer than by the seller), becomes critical when at most a finite number of widgets can be described. Formalization of this argument yields the following result:

Theorem 2. *In an environment with n widgets satisfying Assumptions* 1–3, *choose any incentive-compatible message contingent contract which satisfies* (C1) *and* (C2). *Let* $U_j^n(\boldsymbol{v}, \boldsymbol{c})$ *denote the ex post payoff of party* $j \in \{B, S\}$ *in the contract in the state of the world* $(\boldsymbol{v}, \boldsymbol{c})$. *Then for all* $\boldsymbol{v}, \boldsymbol{v}' \in V^{n+1}$, $\boldsymbol{c}, \boldsymbol{c}' \in C^{n+1}$,

$$\left| U_j^n(\boldsymbol{v}', \boldsymbol{c}') - U_j^n(\boldsymbol{v}, \boldsymbol{c}) - \frac{S(\boldsymbol{v}', \boldsymbol{c}') - S(\boldsymbol{v}, \boldsymbol{c})}{2} \right| \leqq \max_{l \leq k} \begin{cases} \left| \dfrac{(\boldsymbol{v}'+\boldsymbol{c}')^{(l)} - (\boldsymbol{v}+\boldsymbol{c})^{(l)}}{2} \right| \\ \left| \dfrac{(\boldsymbol{v}'+\boldsymbol{c}')^{(n-l+1)} - (\boldsymbol{v}+\boldsymbol{c})^{(n-l+1)}}{2} \right| \\ \dfrac{1}{n-k}\left(\left| \sum_{i=1}^{n} \dfrac{(\boldsymbol{v}'+\boldsymbol{c}')_i - (\boldsymbol{v}+\boldsymbol{c})_i}{2} \right| + kA \right) \end{cases},$$

where $\boldsymbol{v}^{(l)}$ *denotes the lth lowest component* (*rank order statistic*) *of a vector* $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N)$, *and* $A = (\max V + \max C)/2 - (\min V + \min C)/2$.

The left-hand side of the inequality is the absolute value of the difference between the change in party j's *ex post* payoff and 50% of the change in total surplus (compare with Theorem 1). The right-hand side is a maximum of several expressions. The first two expressions represent the changes in the valuations of the k best and k worst widgets. The

16. In Segal (1996), this restriction is obtained by applying the computer science concept of communication complexity to the contractual message game. For example, the concept of "tree complexity" defines complexity of an extensive-form message game by the number of its terminal nodes (Karchmer (1989)). Intuitively, if a message game involves describing k widgets, this can be done in $\binom{n}{k}$ possible ways, which gives a lower bound on the number of terminal nodes in the game. If k is bounded regardless of n, as stipulated by (C2), the tree complexity of the message game is polynomial in n, and otherwise it is exponential.

last expression contains the change in the valuations for the average widget. To understand the result intuitively, observe that the buyer will always suggest the best widgets, and the seller will always suggest the worst widgets out of those which have not yet been described. Therefore, in a contract in which at most k widgets are described, only the best k and the worst k widgets have a chance of being described. The contract can either specify trading one of those widgets, or randomize with equal probabilities over the undescribed widgets.

Under Assumption 5, the k best and the k worst widgets are very likely to be non-responsive to investments for a large n. Under Assumption 4, the average widget is also asymptotically non-responsive to investments. Therefore, the expectation of the right-hand side of the above inequality goes to zero, which implies that equilibrium investments with and without a contract converge:

Corollary 2. *In a sequence of environments where Assumptions* 1–5 *are satisfied and* A *is bounded regardless of* n, *as* $n \to \infty$,

(i) $\rho(\mathscr{E}_n, \mathscr{E}_n^0) \to 0$,
(ii) $\rho(W(\mathscr{E}_n), W(\mathscr{E}_n^0)) \to 0$.

Part (ii) implies, in particular, that the maximum contracting gain converges to zero as $n \to \infty$.

5. DISCUSSION

This section relates the paper's results to existing justifications for incomplete contracting, which can be classified along the four types of "transaction costs" identified by Tirole (1994)—unforeseen contingencies, writing costs, enforcement costs, and renegotiation.

Enforcement Costs: When it is difficult to verify parties' *ex post* valuations for different trades, the complete long-term contract specifying the efficient *ex post* trade in all states of the world may be infeasible. It has been argued that whenever the parties' valuations are mutually observable *ex post* but not verifiable, it may be optimal to negotiate the trade after these valuations are observed, rather than rigidly contract the trade *ex ante*.

This argument does not provide a rigorous foundation for incomplete contracting because whenever parties' valuations are mutually observable *ex post*, a long-term contract can be made indirectly contingent on these valuations by allowing the parties to send messages and conditioning the trade on these messages (see Moore (1992) for a survey of the relevant implementation literature). In the context of the holdup problem, Rogerson (1992) shows that using multi-stage message games and the concept of subgame-perfect implementation suggested by Moore and Repullo (1988), first-best investments can be implemented. Therefore, the cost of verifying publicly observable information *by itself* does not impose a serious constraint on contracting.

Unforeseen Contingencies: Some aspects of future trades may not be foreseeable at the contracting date, and therefore have to be left to future negotiation.

However, incomplete contracting theories assume that the contracting parties can foresee their payoffs in each future contingency. Maskin and Tirole (1999) observe that using this minimum foresight consistent with rationality, the parties can design a message game that effectively describes *ex post* all the trades that have not been described *ex ante*. They show that when the parties can use such a game, inability to foresee future contingencies *by itself* does not constrain contracting.[17]

> **Writing Costs:** If specifying contractual contingencies is costly, the optimal contract should trade off the loss from contractual incompleteness against the cost of adding contractual clauses.

A model of contractual incompleteness based on this idea has been offered by Dye (1985). Taken on its face value, this view predicts that when stakes are very large relative to writing costs, contracts should be very close to being complete, and the relative effects of contractual incompleteness should be negligible. Unfortunately, we do not have a good idea of the magnitude and composition of writing costs, and we are not convinced of their significance relative to stakes in important economic situations. Therefore, existence of writing costs *by itself* does not explain why incompleteness of contracts should be an important economic phenomenon.

> **Renegotiation:** The prevailing legal system does not allow contracting parties to prevent consensual renegotiation of their original contract. Since optimal message-contingent contracts often punish lying parties with inefficient outcomes, the possibility of renegotiation constrains the set of feasible contracts.

While a complete survey of the literature on contracting in hold-up models in the presence of renegotiation would be beyond the scope of this paper, it should be pointed out that different papers have adopted different views of the renegotiation procedure. In the models of Hart and Moore (1988), Aghion *et al*. (1994), MacLeod and Malcomson (1993), and Noldeke and Schmidt (1995), renegotiation outcome is given by some version of the Outside Option Principle, as described *e.g*. in Sutton (1986).[18] By setting a high penalty for breach, a contract can make one party's outside option bind, so that the other party receives the entire renegotiation surplus. This allows to implement the first-best in one-sided investment models (see footnote 9), and to achieve a substantial improvement over the no-contract outcome in two-sided models.

In contrast to these papers, Edlin and Reichelstein (1996) assume that the split of renegotiation surplus over the contractual outcome is exogenously given. In particular, they consider the renegotiation function that splits the surplus in fixed proportions, which is the model of renegotiation adopted in this paper. In this case, they find that a noncontingent ("specific performance") contract can achieve or approximate the first-best. In summary, in all the mentioned hold-up models with renegotiation, a substantial improvement upon the no-contract outcome is possible.[19]

17. In the absence of renegotiation, this generally implies that first-best investments can be implemented, in accordance with Rogerson's result. Maskin and Tirole also show that when the parties are risk-averse and unbounded transfers can be used, renegotiation does not constrain the set of implementable outcomes when lotteries can be used to punish both parties for a disagreement. When the parties are risk-neutral (as they are in our model), Maskin and Tirole only establish (in Theorem 4) that undescribability places no *additional* constraint on contracting, but do not rule out renegotiation-proofness constraints.

18. This is not to overlook the substantial differences among the papers. Thus, Hart and Moore and MacLeod and Malcomson assume that the court cannot verify which party has refused to trade, while Aghion *et al*. and Noldeke and Schmidt make the opposite assumption. Also, Aghion *et al*. and MacLeod and Malcomson allow trading to occur at different points in time, while Hart and Moore and Noldeke and Schmidt only allow trading at a single fixed date. In addition, MacLeod and Malcomson, unlike the other three papers, allow the parties to have external trading options, as well as consider external investments (those that directly benefit the other party).

19. A recent paper by Che and Hausch (1997) obtains incomplete contracting in a hold-up model with "external" investments. See footnote 7 for a comparison of their model to our setup.

In contrast to previous explanations of incomplete contracts, this paper emphasizes the role of environmental complexity. It shows that *in a complex environment*, all four aforementioned sources of "transaction costs" may be important ingredients in explaining contractual incompleteness. Theorem 1 demonstrates that in a complex environment, the parties' inability to verify publicly observable information and their inability to prevent renegotiation place severe constraints on contracting and may explain contractual incompleteness. Theorem 2 shows that in a complex environment, the parties' inability to foresee all possible trades *ex ante* and the cost of describing them *ex post* impose additional constraints on contracting and extend the applicability of the incomplete contracting result.

While much has been said about the role of bounded rationality in explaining contractual incompleteness, existing models of bounded rationality have not been able to explain how people could be irrational enough not to be able to describe all the possible contingencies *ex ante*, yet rational enough to foresee their payoffs *ex ante* and to describe any given contingency *ex post* (see *e.g*. Maskin and Tirole 1999). In our view, any attempt to model bounded rationality in a simple environment is doomed to fall into the trap of describing decision makers as either "completely dumb" or "perfectly rational". Neither is an attractive alternative for modeling "transaction costs". It is only in environments reflecting the real world's complexities that an intermediate region of "bounded rationality" emerges. For example, in a complex environment, it may be easy to describe any given potential trade, but prohibitively costly to describe all of them, or any positive fraction of them. Theorem 2 suggests that this difference may be important for understanding incomplete contracts.[20]

An important challenge for a theory of bounded rationality in complex environments is to explain the role of renegotiation in the real world. This paper has followed the traditional view of renegotiation as a constraint on contracting. The parties' inability to legally bind themselves not to renegotiate is an important element of our explanation of incompleteness of contracts. However, theorists have suggested various ways of eliminating renegotiation or tilting its outcome, such as contracting with an outsider (Holmstrom (1982)), using financial hostages (Aghion *et al*. (1994)), and setting up a public registry of contracts (Maskin and Tirole (1999)). In practice, non-renegotiable contracts have sometimes been successfully enforced in the past (*e.g*., non-renegotiable marriage contracts in some countries). Therefore, the absence of legal provisions allowing contracting parties to prevent out-of-contract renegotiation calls for an explanation.

The explanation may come from bounded rationality. In a complex environment in which contracting parties possess private information about trade opportunities, the renegotiation game may be more complex than mere haggling over surplus. Given the considerable cost of administering complicated contractual message games ("formal communication"), the parties may prefer not to write the renegotiation game directly in the *ex ante* contract, instead leaving the option of renegotiation ("informal communication") open. Thus, the "transaction costs" of formal communication may help explain why contracting parties would *strictly* prefer to write an incomplete renegotiable contract, rather than a complete non-renegotiable one.

20. Anderlini and Felli (1994) and MacLeod (1996) also model complex environments in which bounded rationality can result in contractual incompleteness. In the model of MacLeod, the optimal contract is exponentially complex. Anderlini and Felli describe contracting problems in which the process of arriving at the optimal contract is not computable at all.

APPENDIX

Proof of Proposition 1. Observe that both (λ^0, σ^0) and (λ^*, σ^*) solve a maximization program of the form

$$\max_{(\lambda,\sigma)\in\Lambda\times\Sigma} \alpha E_\zeta S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta)) - \lambda - \sigma,$$

where $z = 1/2$ for no-contract investments (λ^0, σ^0) and $z = 1$ for first-best investments (λ^*, σ^*). The assumptions imply that the objective function has increasing marginal returns both in $(\sigma; z)$ and in $(\lambda; z)$, as defined in Edlin and Shannon (1998), and parts (i) and (ii) follows by their Strict Monotonicity Theorem.

In case (iii), the above maximization program can be separated into two:

$$\max_{\lambda\in\Lambda} [zE_\zeta v_{w^*(\zeta)}(\lambda, \zeta) - \lambda],$$

$$\max_{\sigma\in\Sigma} [-zE_\zeta c_{w^*(\zeta)}(\sigma, \zeta) - \sigma].$$

The first objective function has increasing marginal returns in $(\lambda; z)$, and the second objective function has increasing marginal returns in $(\sigma; z)$. The Strict Monotonicity Theorem of Edlin and Shannon (1998) implies that $\lambda^0 < \lambda^*$ and $\sigma^0 < \sigma^*$. ||

Proof of Proposition 2. Suppose the buyer considers deviating to $\lambda' \in \Lambda$. Using (2), the buyer's *ex post* utility in contingency ζ will then be

$$-t(\zeta) + \frac{v_{w^*}(\lambda', \zeta) + c_{w^*}(\sigma^*, \zeta)}{2} + \tfrac{1}{2}S(\boldsymbol{v}(\lambda', \zeta), \boldsymbol{c}(\sigma^*, \zeta))$$

$$= -t(\zeta) + c_{w^*}(\sigma^*, \zeta) + \tfrac{1}{2}[v_{w^*}(\lambda', \zeta) - c_{w^*}(\sigma^*, \zeta)] + \tfrac{1}{2}S(\boldsymbol{v}(\lambda', \zeta), \boldsymbol{c}(\sigma^*, \zeta))$$

$$\leqq -t(\zeta) + c_{w^*}(\sigma^*, \zeta) + S(\boldsymbol{v}(\lambda', \zeta), \boldsymbol{c}(\sigma^*, \zeta)),$$

where the last inequality follows from the definition of $S(\boldsymbol{v}(\lambda', \zeta), \boldsymbol{c}(\sigma^*, \zeta))$. Using in addition the fact that (λ^*, σ^*) solves (1), the buyer's expected *ex ante* utility after a deviation can be bounded as follows

$$E_\zeta[-t(\zeta) + c_{w^*}(\sigma^*, \zeta) + S(\boldsymbol{v}(\lambda', \zeta), \boldsymbol{c}(\sigma^*, \zeta))] - \lambda'$$

$$\leqq E_\zeta[-t(\zeta) + c_{w^*}(\sigma^*, \zeta) + S(\boldsymbol{v}(\lambda^*, \zeta), \boldsymbol{c}(\sigma^*, \zeta))] - \lambda^*$$

$$= E_\zeta[-t(\zeta) - v_{w^*}(\lambda^*, \zeta)] - \lambda^*,$$

which is the buyer's equilibrium payoff. Therefore, the deviation does not make the buyer better off. Similarly for the seller. ||

Proof of Theorem 1. Let $\boldsymbol{v}(\theta)$ and $\boldsymbol{c}(\theta)$ denote the two parties' valuation vectors in the state of the world θ, *i.e.*

$$v_w(\boldsymbol{N}, v, c) = \begin{cases} v \text{ if } \boldsymbol{N}_w = R, \\ v_g \text{ if } \boldsymbol{N}_w = G, \\ v_b \text{ if } \boldsymbol{N}_w = B, \\ 0 \text{ if } w = 0, \end{cases} \qquad c_w(\boldsymbol{N}, v, c) = \begin{cases} c \text{ if } \boldsymbol{N}_w = R, \\ c_g \text{ if } \boldsymbol{N}_w = G, \\ c_b \text{ if } \boldsymbol{N}_w = B, \\ 0 \text{ if } w = 0. \end{cases}$$

Consider the disagreement in which the buyer announces that the state of the world is θ_B and the seller announces that the state of the world is θ_S. Using Theorem 0 and expressions (2), this disagreement can be resolved if, and only if, there exists an outcome $(\boldsymbol{x}, t)$ such that

$$-t + \textstyle\sum_{w=0}^{n} x_w \dfrac{v_w(\theta_S) + c_w(\theta_S)}{2} + \tfrac{1}{2}S(\boldsymbol{v}(\theta_S), \boldsymbol{c}(\theta_S)) \leqq U_B(\theta_S),$$

$$t - \textstyle\sum_{w=0}^{n} x_w \dfrac{v_w(\theta_B) + c_w(\theta_B)}{2} + \tfrac{1}{2}S(\boldsymbol{v}(\theta_B), \boldsymbol{c}(\theta_B)) \leqq U_S(\theta_B).$$

Adding up, and taking into account that $U_B(\theta_S) = S(\boldsymbol{v}(\theta_S), \boldsymbol{c}(\theta_S)) - U_S(\theta_S)$, yields

$$U_S(\theta_B) - U_S(\theta_S) \geqq \frac{S(\boldsymbol{v}(\theta_B), \boldsymbol{c}(\theta_B)) - S(\boldsymbol{v}(\theta_S), \boldsymbol{c}(\theta_S))}{2} + \sum_{w=0}^{n} x_w \left[\frac{v_w(\theta_S) + c_w(\theta_S)}{2} - \frac{v_w(\theta_B) + c_w(\theta_B)}{2}\right]. \tag{7}$$

It is possible to find $\boldsymbol{x} \in \{(x_0, x_1, \ldots, x_n) \in \Re_+^{n+1} : \sum_{w \in \{0,1,\ldots,n\}} x_w = 1\}$ satisfying this inequality if, and only if,

$$U_S(\theta_B) - U_S(\theta_S) \geqq \frac{S(\boldsymbol{v}(\theta_B), \boldsymbol{c}(\theta_B)) - S(\boldsymbol{v}(\theta_S), \boldsymbol{c}(\theta_S))}{2} + \min_{w \in \{0,1,\ldots,n\}} \left[\frac{v_w(\theta_S) + c_w(\theta_S)}{2} - \frac{v_w(\theta_B) + c_w(\theta_B)}{2}\right]. \tag{8}$$

The following three lemmas will apply this incentive constraint to three types of disagreements. Lemma A.1 considers critical disagreements, Lemma A.2 considers disagreements where parties agree on the valuations for the efficient trade and disagree on the configuration, and Lemma A.3 considers disagreements where parties agree on a configuration and disagree on the valuations for the efficient trade.

Lemma A.1. *In an incentive-compatible contract, for all* $(\boldsymbol{M}, \boldsymbol{N}) \in \mathfrak{D}$ *and all* $v_B, v_S \in \bar{V}$, c_B, $c_S \in \bar{C}$ *we must have*

$$U_S(\boldsymbol{N}, v_B, c_B) - U_S(\boldsymbol{M}, v_S, c_S) \geqq \frac{v_B - c_B}{2} - \frac{v_S - c_S}{2}.$$

Proof of Lemma A.1. Let $\theta_B = (\boldsymbol{N}, v_B, c_B)$ and $\theta_S = (\boldsymbol{M}, v_S, c_S)$, where $(\boldsymbol{M}, \boldsymbol{N}) \in \mathfrak{D}$, with $\boldsymbol{M}_i = \boldsymbol{N}_j = R$. Using (4), we can write

$$\frac{v_w(\theta_S) + c_w(\theta_S)}{2} - \frac{v_w(\theta_B) + c_w(\theta_B)}{2} = \begin{cases} (v_S + c_S)/2 - \underline{a} > 0 \text{ when } w = i, \\ \bar{a} - (v_B + c_B)/2 > 0 \text{ when } w = j, \\ 0 \text{ when } \boldsymbol{M}_w = \boldsymbol{N}_w \in \{B, G\} \text{ or } w = 0. \end{cases}$$

Substituting in the incentive constraint (8), we obtain the lemma's statement. ||

Lemma A.2. *In an incentive-compatible contract, for any two configurations* $\boldsymbol{N}$, $\boldsymbol{M}$, *and for all* $v \in \bar{V}$, $c \in \bar{C}$ *we must have*

$$U_S(\boldsymbol{M}, v, c) - U_S(\boldsymbol{N}, v, c) \leqq A.$$

Proof of Lemma A.2. Let $\theta_B = (\boldsymbol{N}, v, c)$ and $\theta_S = (\boldsymbol{M}, v, c)$. Using (4), we can write

$$\min_{w.\{0,1,\ldots,n\}} \left[\frac{v_w(\theta_S) + c_w(\theta_S)}{2} - \frac{v_w(\theta_B) + c_w(\theta_B)}{2}\right] \geqq \min_w \frac{v_w(\theta_S) + c_w(\theta_S)}{2} - \max_w \frac{v_w(\theta_B) + c_w(\theta_B)}{2} \geqq \underline{a} - \bar{a} = -A.$$

Substituting in the incentive constraint (8), we obtain the lemma's statement. ||

Lemma A.3. *In an incentive-compatible contract, for all* $\boldsymbol{N}$ *and all* $v_B, v_S \in \bar{V}$, $c_B, c_S \in \bar{C}$ *we must have*[21]

$$U_S(\boldsymbol{N}, v_B, c_B) - U_S(\boldsymbol{N}, v_S, c_S) \geqq \frac{v_B - c_B}{2} - \frac{v_S - c_S}{2} + \min\left\{0, \frac{v_S + c_S}{2} - \frac{v_B + c_B}{2}\right\}.$$

Proof of Lemma A.3. Let $\theta_B = (\boldsymbol{N}, v_B, c_B)$, and $\theta_S = (\boldsymbol{N}, v_S, c_S)$, where $\boldsymbol{N}_i = R$. Using (4), we can write

$$\frac{v_w(\theta_S) + c_w(\theta_S)}{2} - \frac{v_w(\theta_B) + c_w(\theta_B)}{2} = \begin{cases} \frac{v_S + c_S}{2} - \frac{v_B + c_B}{2} \text{ when } w = i, \\ 0 \text{ when } \boldsymbol{N}_w \in \{B, G\} \text{ or } w = 0. \end{cases}$$

Substituting in the incentive constraint (8), we obtain the lemma's statement. ||

Now we introduce some useful notation. Define the *length* of an edge $(\boldsymbol{M}, \boldsymbol{N}) \in \mathfrak{D}$ as $\tau(\boldsymbol{M}, \boldsymbol{N}, v, c) \equiv U_S(\boldsymbol{N}, v, c) - U_S(\boldsymbol{M}, v, c)$. Substituting $v_B = v_S = v$, $c_B = c_S = c$ in the statement of Lemma A.1, we obtain

$$\tau(\boldsymbol{M}, \boldsymbol{N}, v, c) \geqq 0 \quad \text{for all } (\boldsymbol{M}, \boldsymbol{N}) \in \mathfrak{D}, \tag{9}$$

i.e. the length of an edge is non-negative, which vindicates the term. Using this notation, we can rewrite the statement of Lemma A.1 as follows

$$U_S(\boldsymbol{N}, v_B, c_B) - U_S(\boldsymbol{N}, v_S, c_S) \geqq \frac{v_B - c_B}{2} - \frac{v_S - c_S}{2} - \tau(\boldsymbol{M}, \boldsymbol{N}, v_S, c_S).$$

The inequality holds for every configuration $\boldsymbol{M}$ adjacent to $\boldsymbol{N}$ from below, *i.e.* satisfying $(\boldsymbol{M}, \boldsymbol{N}) \in \mathfrak{D}$. Observe that there are $|\boldsymbol{N}|$ such configurations. To see this, let $\boldsymbol{N}_j = R$ and observe that for each widget i such that $\boldsymbol{N}_i = B$, there is one critical disagreement terminating in $\boldsymbol{N}$, namely the disagreement $(\boldsymbol{M}, \boldsymbol{N})$ satisfying (5). Thus, the number of edges adjacent to $\boldsymbol{N}$ from below equals the number of type B widgets in $\boldsymbol{N}$, *i.e.* $|\boldsymbol{N}|$. (Similarly, it can be shown that the number of edges adjacent to $\boldsymbol{N}$ from above equals the number of type G widgets in $\boldsymbol{N}$, i.e. $n - |\boldsymbol{N}| - 1$.)

After averaging the right-hand side of the above inequality over all the edges adjacent to $\boldsymbol{N}$ from below, we obtain the following bound for $|\boldsymbol{N}| \geqq 1$

$$U_S(\boldsymbol{N}, v_B, c_B) - U_S(\boldsymbol{N}, v_S, c_S) \geqq \frac{v_B - c_B}{2} - \frac{v_S - c_S}{2} - \frac{1}{|\boldsymbol{N}|} \sum_{(\boldsymbol{M}, \boldsymbol{N}) \in \mathfrak{D}} \tau(\boldsymbol{M}, \boldsymbol{N}, v_S, c_S). \tag{10}$$

To see how the seller's expected utility can depend on the parties' valuations, we use (10) for $|\boldsymbol{N}| \geqq 1$, and the statement of Lemma A.3 for $|\boldsymbol{N}| = 0$, taking into account that

$$\min\left\{0, \frac{v_S + c_S}{2} - \frac{v_B + c_B}{2}\right\} \geqq \underline{a} - \bar{a} = -A:$$

$$E_{\boldsymbol{N}} U_S(\boldsymbol{N}, v_B, c_B) - E_{\boldsymbol{N}} U_S(\boldsymbol{N}, v_S, c_S) = \sum_{\boldsymbol{N}} p[U_S(\boldsymbol{N}, v_B, c_B) - U_S(\boldsymbol{N}, v_S, c_S)]$$

$$\geqq \frac{v_B - c_B}{2} - \frac{v_S - c_S}{2} - pnA - p \sum_{|\boldsymbol{N}| \geq 1} \frac{1}{|\boldsymbol{N}|} \sum_{(\boldsymbol{M}, \boldsymbol{N}) \in \mathfrak{D}} \tau(\boldsymbol{M}, \boldsymbol{N}, v_S, c_S), \tag{11}$$

where $p = 1/(n2^{n-1})$ is the probability of any given configuration. This expression can be bounded from below by maximizing the following sum

$$p \sum_{|\boldsymbol{N}| \geq 1} \frac{1}{|\boldsymbol{N}|} \sum_{(\boldsymbol{M}, \boldsymbol{N}) \in)} \tau(\boldsymbol{M}, \boldsymbol{N}, v, c) \tag{12}$$

We are going to bound the sum using Lemma A.2. For this purpose, define a *path* P as a set of $n - 1$ pairwise adjacent edges leading from the bottom to the top of the graph of critical disagreements: $P = \{(\boldsymbol{N}^0, \boldsymbol{N}^1), (\boldsymbol{N}^1, \boldsymbol{N}^2), \cdots, (\boldsymbol{N}^{n-2}, \boldsymbol{N}^{n-1})\} \subset \mathfrak{D}$. Denote the set of all paths by Π. Using this notation, we can use Lemma

21. An interesting corollary of this lemma can be obtained by substituting $v_B = v_S = v$, $c_B = c' < c_S = c$. Then the statement can be rewritten as $U_S(\boldsymbol{N}, v, c') - U_S(\boldsymbol{N}, v, c) \geqq (c - c')/2$. This implies that in any incentive compatible contract, the seller receives at least 50% of his contribution to total surplus. Therefore, for any n, the seller's investment incentive cannot be lower than that without a contract. (The same conclusion can be obtained for the buyer).

A.2 to write

$$\textstyle\sum_{(\boldsymbol{M},\boldsymbol{N})\in P}\tau(\boldsymbol{M},\boldsymbol{N},v,c)=U_S(\boldsymbol{N}^{n-1},v,c)-U_S(\boldsymbol{N}^0,v,c)\leqq A$$

for all $P\in\Pi$.

We will use a weaker inequality obtained by adding up the inequality above for all paths. To do this, first observe that Π contains $n!$ elements. Indeed, there are n ways to choose a configuration from the top, and each configuration at a level $l\in\{1,\cdots,n-1\}$ is connected via a critical disagreement to l configurations at level $l-1$. Thus, a path can be chosen in $n\cdot(n-1)\cdot\ldots\cdot 1=n!$ different ways. Second, using a similar counting argument, we can see that an edge $(\boldsymbol{M},\boldsymbol{N})\in\mathfrak{D}$ belongs to $(n-1-|\boldsymbol{N}|)!(|\boldsymbol{N}|-1)!$ different paths. Therefore, by adding up the above inequalities for all paths, we obtain

$$\textstyle\sum_{(\boldsymbol{M},\boldsymbol{N})\in\mathfrak{D}}(n-1-|\boldsymbol{N}|)!(|\boldsymbol{N}|-1)!\,\tau(\boldsymbol{M},\boldsymbol{N},v,c)\leqq n!\cdot A. \tag{13}$$

Now, consider the linear programming problem of maximizing (12) subject to (13) and (9). The solution to this program is achieved at a vertex, where the length of all edges but one is set to zero. When we put all the weight on an edge $(\boldsymbol{M},\boldsymbol{N})\in\mathfrak{D}$, we obtain $\tau(\boldsymbol{M},\boldsymbol{N},v,c)=n!/((n-1-|\boldsymbol{N}|)!(|\boldsymbol{N}|-1)!)A$, and the value of the objective function is

$$p\frac{n!}{(n-1-|\boldsymbol{N}|)!(|\boldsymbol{N}|-1)!|\boldsymbol{N}|}A=pn\binom{n-1}{|\boldsymbol{N}|}A=\frac{1}{2^{n-1}}\binom{n-1}{|\boldsymbol{N}|}A.$$

The maximum value is achieved at $|\boldsymbol{N}|=[(n-1)/2]$, which gives us the following lower bound on the left-hand side of (11)

$$E_{\boldsymbol{N}}U_S(\boldsymbol{N},v_B,c_B)-E_{\boldsymbol{N}}U_S(\boldsymbol{N},v_S,c_S)-\left(\frac{v_B-c_B}{2}-\frac{v_S-c_S}{2}\right)\geqq-\frac{A}{2^{n-1}}-\frac{1}{2^{n-1}}\binom{n-1}{[(n-1)/2]}A.$$

Substituting $(v_B,c_B)=(v',c')$, $(v_S,c_S)=(v,c)$, we obtain a lower bound on $E_{\boldsymbol{N}}U_S(\boldsymbol{N},v',c')-E_{\boldsymbol{N}}U_S(\boldsymbol{N},v,c)-((v'-c')/2-(v-c)/2)$. Substituting instead $(v_B,c_B)=(v,c)$, $(v_S,c_S)=(v',c')$, we obtain an upper bound on the same expression. Combining the two bounds, we obtain the theorem's statement for $j=S$. Expressing the buyer's utility as $U_B(\boldsymbol{N},v,c)=v-c-U_S(\boldsymbol{N},v,c)$, we obtain the theorem's statement for $j=B$. ||

Proof of Corollary 1. Let $u_j^n(\lambda,\sigma)$ denote party j's expected *ex post* utility in the contract with n widgets as a function of investments (λ,σ). Using the fact that $\boldsymbol{N}$ is distributed independently of (v,c), we can write

$$u_j^n(\lambda,\sigma)=E_\zeta U_j^n(\boldsymbol{N}(\zeta),v(\lambda,\zeta),c(\sigma,\zeta))=E_\zeta[E_{\boldsymbol{N}}U_j^n(\boldsymbol{N},v(\lambda,\zeta),c(\sigma,\zeta))]$$

where $v(\lambda,\zeta)$, $c(\sigma,\zeta)$ are the parties' valuations for the R widget in contingency ζ given investments λ,σ, and $\boldsymbol{N}(\zeta)$ is the configuration in contingency ζ.

Stirling's formula implies that

$$\frac{1}{2^{n-1}}\binom{n-1}{[(n-1)/2]}A\sim\sqrt{\frac{2}{\pi(n-1)}}A\to 0\text{ as }n\to\infty.$$

Therefore, Theorem 1 implies that as $n\to\infty$, for $j\in\{B,S\}$ we have

$$E_{\boldsymbol{N}}U_j^n(\boldsymbol{N},v',c')-E_{\boldsymbol{N}}U_j^n(\boldsymbol{N},v,c)\to\frac{v'-c'}{2}-\frac{v-c}{2},$$

where convergence is uniform on $v,v'\in\bar{V}$, $c,c'\in\bar{C}$. Substituting $v=v(\lambda,\zeta)$, $v'=v(\lambda',\zeta)$, $c=c(\sigma,\zeta)$, $c'=v(\sigma',\zeta)$, and taking the expectation over ζ, we obtain that for all $\lambda,\lambda'\in\Lambda$, $\sigma,\sigma'\in\Sigma$, as $n\to\infty$,

$$u_j^n(\lambda',\sigma')-u_j^n(\lambda,\sigma)\to\tfrac{1}{2}[E_\zeta v(\lambda',\zeta)-E_\zeta v(\lambda,\zeta)]-\tfrac{1}{2}[E_\zeta c(\sigma',\zeta)-E_\zeta c(\sigma,\zeta)]. \tag{14}$$

Now, suppose in negation that part (i) of the Corollary's statement is not true. Since $\mathscr{E}^0\subset\mathscr{E}_n$ for all n, the supposition implies that $\sup_{(\lambda_n,\sigma_n)\in\mathscr{E}_n}\operatorname{dist}((\lambda_n,\sigma_n),\mathscr{E}^0)$ does not go to zero as $n\to\infty$. Therefore, there exists a sequence $\{(\lambda_n,\sigma_n)\}_{n=1}^\infty$ such that $(\lambda_n,\sigma_n)\in\mathscr{E}_n$ for all n and $\operatorname{dist}((\lambda_n,\sigma_n),\mathscr{E}^0)$ does not converge to zero. Since the sequence lies in a compact set $\Lambda\times\Sigma$, we can pick a subsequence $\{(\lambda_k,\sigma_k)\}_{k=1}^\infty\subset\mathscr{E}_k$ that converges to a point $(\lambda,\sigma)\notin\mathscr{E}^0$.

Since (λ_k, σ_k) is a Nash equilibrium investment pair given the contract, for all $\lambda \in \Lambda$ we must have

$$u_B^k(\lambda, \sigma_k) - u_B^k(\lambda_k, \sigma_k) \leqq \lambda - \lambda_k. \tag{15}$$

The left-hand side of this inequality can be rewritten as

$$(u_B^k(\lambda, \sigma_k) - u_B^k(\lambda_k, \sigma_k) - \tfrac{1}{2}[E_\zeta v(\lambda, \zeta) - E_\zeta v(\lambda_k, \zeta)]) + \tfrac{1}{2}[E_\zeta v(\lambda, \zeta) - E_\zeta v(\lambda_k, \zeta)].$$

As $k \to \infty$, the first term converges to zero by the uniform convergence in (14), while the second term converges to $\frac{1}{2}[E_\zeta v(\lambda, \zeta) - E_\zeta v(\hat{\lambda}, \zeta)]$ by the assumption that $E_\zeta v(\lambda, \zeta) = E_\zeta S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta)) - E_\zeta c(\sigma, \zeta)$ is continuous in λ. Thus, taking the limit $k \to \infty$ in inequality (15), we obtain

$$\tfrac{1}{2}E_\zeta v(\lambda, \zeta) - \tfrac{1}{2}E_\zeta v(\hat{\lambda}, \zeta) \leqq \lambda - \hat{\lambda}.$$

Since this should be true for any $\lambda \in \Lambda$, $\hat{\lambda}$ must satisfy the first condition in (3), *i.e.* it must be the buyer's equilibrium investment without a contract. In the same way it can be shown that $\hat{\sigma}$ is the seller's equilibrium investment without a contract. Therefore, $(\hat{\lambda}, \hat{\sigma}) \in \mathscr{E}^0$, which contradicts an earlier statement. This establishes part (i) of the Corollary. Part (ii) follows by the continuity of $W(\cdot)$. ||

Proof of Theorem 2. First we prove two lemmas which establish certain useful consequences of contracting restriction (C1).

Lemma B.1. *For any* $i, j \in \Omega(m_B, m_S)$, *we must have* $x_i(m_B, m_S) = x_j(m_B, m_S)$.

Proof of Lemma B.1. Take $\pi = (ij)$, *i.e.* the permutation that transposes widgets i and j, holding other widgets fixed. This permutation only affects $\Omega(m_B, m_S)$. Using the fact that $\Omega(m_B, m_S)$ is undescribed by (m_B, m_S) and condition (C1), we can write

$$x_i(m_B, m_S) = x_i(m_B\pi, m_S\pi) = x_{\pi(i)}(m_B, m_S) = x_j(m_B, m_S). \quad ||$$

Lemma B.2. *For each party* $j \in \{B, S\}$, *all states of the world* $(\boldsymbol{v}, \boldsymbol{c})$, *and all permutations* π *of widgets, we must have* $U_j^n(\boldsymbol{v}\pi, \boldsymbol{c}\pi) = U_j^n(\boldsymbol{v}, \boldsymbol{c})$.

Proof of Lemma B.2. Follows immediately from (C1). ||

Let $\hat{m}_B(\cdot)$, $\hat{m}_S(\cdot)$ be the parties' equilibrium strategies. Consider disagreements in which the buyer and seller send messages of the form $\hat{m}_B(\boldsymbol{v}_B\pi_B, \boldsymbol{c}_B\pi_B)$ and $\hat{m}_S(\boldsymbol{v}_S\pi_S, \boldsymbol{c}_S\pi_S)$ respectively. Fixing $(\boldsymbol{v}_B, \boldsymbol{c}_B)$ and $(\boldsymbol{v}_S, \boldsymbol{c}_S)$, there exists a pair of permutations (π_B, π_S) such that:

(i) For every widget i described by the seller, there exists $l \leqq k$ such that

$$\begin{aligned}(\boldsymbol{v}_B\pi_B + \boldsymbol{c}_B\pi_B)_i &= (\boldsymbol{v}_B + \boldsymbol{c}_B)^{(l)},\\ (\boldsymbol{v}_S\pi_S + \boldsymbol{c}_S\pi_S)_i &\geqq (\boldsymbol{v}_S + \boldsymbol{c}_S)^{(l)}.\end{aligned} \tag{16}$$

In words, the buyer reports that it is one of the k lowest-ranked widgets, and that its rank l is not higher than what it is according to the seller's report.

(ii) For every widget i described by the buyer, there exists $l \leqq k$ such that

$$\begin{aligned}(\boldsymbol{v}_S\pi_S + \boldsymbol{c}_S\pi_S)_i &= (\boldsymbol{v}_S + \boldsymbol{c}_S)^{(n-l+1)},\\ (\boldsymbol{v}_B\pi_B + \boldsymbol{c}_B\pi_B)_i &\leqq (\boldsymbol{v}_B + \boldsymbol{c}_B)^{(n-l+1)}.\end{aligned} \tag{17}$$

In words, the seller reports that it is one of the k highest-ranked widgets, and that its rank $n-l+1$ is not lower than what it is according to the buyer's report.

The outcome $(\boldsymbol{x}, t)$ for this message pair should satisfy the Maskin–Moore incentive constraint (7) derived in the proof of Theorem 1, which here can be written as

$$U_S(\boldsymbol{v}_B\pi_B, \boldsymbol{c}_B\pi_B) - U_S(\boldsymbol{v}_S\pi_S, \boldsymbol{c}_S\pi_S)$$
$$\geqq \frac{S(\boldsymbol{v}_B, \boldsymbol{c}_B) - S(\boldsymbol{v}_S, \boldsymbol{c}_S)}{2} + \sum_{i=0}^n x_i\left[\frac{(\boldsymbol{v}_S\pi_S + \boldsymbol{c}_S\pi_S)_i}{2} - \frac{(\boldsymbol{v}_B\pi_B + \boldsymbol{c}_B\pi_B)_i}{2}\right].$$

Using Lemmas B.1 and B.2, the constraint can be rewritten as

$$
\begin{aligned}
&U_S(\boldsymbol{v}_B, \boldsymbol{c}_B) - U_S(\boldsymbol{v}_S, \boldsymbol{c}_S) - \frac{S(\boldsymbol{v}_B, \boldsymbol{c}_B) - S(\boldsymbol{v}_S, \boldsymbol{c}_S)}{2} \\
&\geqq \frac{1 - \sum_{i \notin \Omega} x_i}{|\Omega|} \sum_{i \in \Omega} \left[\frac{(\boldsymbol{v}_S \boldsymbol{\pi}_S + \boldsymbol{c}_S \boldsymbol{\pi}_S)_i}{2} - \frac{(\boldsymbol{v}_B \boldsymbol{\pi}_B + \boldsymbol{c}_B \boldsymbol{\pi}_B)_i}{2} \right] \\
&\quad + \sum_{i \in \Omega} x_i \left[\frac{(\boldsymbol{v}_S \boldsymbol{\pi}_S + \boldsymbol{c}_S \boldsymbol{\pi}_S)_i}{2} - \frac{(\boldsymbol{v}_B \boldsymbol{\pi}_B + \boldsymbol{c}_B \boldsymbol{\pi}_B)_i}{2} \right] \\
&\geqq \min_{w \notin \Omega} \left\{ \begin{array}{c} \left\{ \frac{(\boldsymbol{v}_S \boldsymbol{\pi}_S + \boldsymbol{c}_S \boldsymbol{\pi}_S)_w}{2} - \frac{(\boldsymbol{v}_B \boldsymbol{\pi}_B + \boldsymbol{c}_B \boldsymbol{\pi}_B)_w}{2} \right\} \\ \frac{1}{|\Omega|} \sum_{i \in \Omega} \left[\frac{(\boldsymbol{v}_S \boldsymbol{\pi}_S + \boldsymbol{c}_S \boldsymbol{\pi}_S)_i}{2} - \frac{(\boldsymbol{v}_B \boldsymbol{\pi}_B + \boldsymbol{c}_B \boldsymbol{\pi}_B)_i}{2} \right] \end{array} \right\}, \qquad (18)
\end{aligned}
$$

where Ω is the largest set of undesribed widgets for this message pair. To bound the first term in the min in (18), consider the following two cases:

(i) If widget w has been described by the seller, then by (16) we know that for some $l \leqq k$,

$$
\begin{aligned}
\frac{((\boldsymbol{v}_S + \boldsymbol{c}_S)\boldsymbol{\pi}_S)_w}{2} - \frac{((\boldsymbol{v}_B + \boldsymbol{c}_B)\boldsymbol{\pi}_B)_w}{2} &\geqq \frac{(\boldsymbol{v}_S + \boldsymbol{c}_S)^{(l)}}{2} - \frac{(\boldsymbol{v}_B + \boldsymbol{c}_B)^{(l)}}{2} \\
&\geqq - \left| \frac{(\boldsymbol{v}_B + \boldsymbol{c}_B)^{(l)}}{2} - \frac{(\boldsymbol{v}_S + \boldsymbol{c}_S)^{(l)}}{2} \right|
\end{aligned}
$$

(ii) If widget w has been described by the buyer, then by (17) we know that for some $l \leqq k$,

$$
\begin{aligned}
\frac{((\boldsymbol{v}_S + \boldsymbol{c}_S)\boldsymbol{\pi}_S)_w}{2} - \frac{((\boldsymbol{v}_B + \boldsymbol{c}_B)\boldsymbol{\pi}_B)_w}{2} &\geqq \frac{(\boldsymbol{v}_S + \boldsymbol{c}_S)^{(n-l+1)}}{2} - \frac{(\boldsymbol{v}_B + \boldsymbol{c}_B)^{(n-l+1)}}{2} \\
&\geqq - \left| \frac{(\boldsymbol{v}_B + \boldsymbol{c}_B)^{(n-l+1)}}{2} - \frac{(\boldsymbol{v}_S + \boldsymbol{c}_S)^{(n-l+1)}}{2} \right|.
\end{aligned}
$$

To bound the second term in the min in (18), we use the fact that $|\Omega| \geqq n - k$ and the definition of A

$$
\begin{aligned}
&\frac{1}{|\Omega|} \sum_{i \in \Omega} \frac{((\boldsymbol{v}_S + \boldsymbol{c}_S)\boldsymbol{\pi}_S)_i - ((\boldsymbol{v}_B + \boldsymbol{c}_B)\boldsymbol{\pi}_B)_i}{2} \\
&\geqq - \frac{1}{|\Omega|} \left| \sum_{i=1}^{n} \frac{(\boldsymbol{v}_S + \boldsymbol{c}_S)_i - (\boldsymbol{v}_B + \boldsymbol{c}_B)_i}{2} - \sum_{i \notin \Omega} \frac{((\boldsymbol{v}_S + \boldsymbol{c}_S)\boldsymbol{\pi}_S)_i - ((\boldsymbol{v}_B + \boldsymbol{c}_B)\boldsymbol{\pi}_B)_i}{2} \right| \\
&\geqq - \frac{1}{n-k} \left(\left| \sum_{i=1}^{n} \frac{(\boldsymbol{v}_S + \boldsymbol{c}_S)_i - (\boldsymbol{v}_B + \boldsymbol{c}_B)_i}{2} \right| + kA \right).
\end{aligned}
$$

Putting all this together, we can use (18) to write

$$
\begin{aligned}
&U_S(\boldsymbol{v}_B, \boldsymbol{c}_B) - U_S(\boldsymbol{v}_S, \boldsymbol{c}_S) - \frac{S(\boldsymbol{v}_B, \boldsymbol{c}_B) - S(\boldsymbol{v}_S, \boldsymbol{c}_S)}{2} \\
&\geqq - \max_{l \leqq k} \left\{ \begin{array}{c} \left| \frac{(\boldsymbol{v}_B + \boldsymbol{c}_B)^{(l)}}{2} - \frac{(\boldsymbol{v}_S + \boldsymbol{c}_S)^{(l)}}{2} \right| \\ \left| \frac{(\boldsymbol{v}_B + \boldsymbol{c}_B)^{(m-l+1)}}{2} - \frac{(\boldsymbol{v}_S + \boldsymbol{c}_S)^{(m-l+1)}}{2} \right| \\ \frac{1}{n-k} \left(\left| \sum_{i=1}^{n} \frac{(\boldsymbol{v}_B + \boldsymbol{c}_B)_i - (\boldsymbol{v}_S + \boldsymbol{c}_S)_i}{2} \right| + kA \right) \end{array} \right\}.
\end{aligned}
$$

Substituting $(\boldsymbol{v}_B, \boldsymbol{c}_B) = (\boldsymbol{v}', \boldsymbol{c}')$, $(\boldsymbol{v}_S, \boldsymbol{c}_S) = (\boldsymbol{v}, \boldsymbol{c})$, we obtain a lower bound on $U_S(\boldsymbol{v}', \boldsymbol{c}') - U_S(\boldsymbol{v}, \boldsymbol{c}) - (S(\boldsymbol{v}', \boldsymbol{c}') - S(\boldsymbol{v}, \boldsymbol{c}))/2$. Substituting instead $(\boldsymbol{v}_B, \boldsymbol{c}_B) = (\boldsymbol{v}, \boldsymbol{c})$, $(\boldsymbol{v}_S, \boldsymbol{c}_S) = (\boldsymbol{v}', \boldsymbol{c}')$, we obtain an upper bound on the same expression. Combining the two bounds, we obtain the theorem's statement for $j = S$. Expressing the buyer's utility as $U_B(\boldsymbol{v}, \boldsymbol{c}) \equiv S(\boldsymbol{v}, \boldsymbol{c}) - U_S(\boldsymbol{v}, \boldsymbol{c})$, we obtain the theorem's statement for $j = B$. ||

Proof of Corollary 2. Let $u_j^n(\lambda, \sigma)$ denote party j's expected *ex post* utility in the contract with n widgets as a function of investments (λ, σ), *i.e.*

$$u_j^n(\lambda, \sigma) = \int U_j^n(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta))dF(\zeta).$$

Using Theorem 2 and the fact that $|E(R)| \leqq E(|R|)$ for any random variable R, we can write

$$\left| u_j^n(\lambda', \sigma') - u_j^n(\lambda, \sigma) - \frac{E_\zeta S(\boldsymbol{v}(\lambda', \zeta), \boldsymbol{c}(\sigma', \zeta')) - E_\zeta S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta))}{2} \right|$$

$$\leqq E_\zeta \left| U_j^n(\boldsymbol{v}(\lambda', \zeta), \boldsymbol{c}(\sigma', \zeta)) - U_j^n(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta)) - \frac{S(\boldsymbol{v}(\lambda', \zeta), \boldsymbol{c}(\sigma', \zeta)) - S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta))}{2} \right|$$

$$\left| \frac{(\boldsymbol{v}(\lambda', \zeta) + \boldsymbol{c}(\sigma', \zeta))^{(l)} - (\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta))^{(l)}}{2} \right|$$

$$\leqq E_\zeta \max_{l \leqq k} \quad \left| \frac{(\boldsymbol{v}(\lambda', \zeta) + \boldsymbol{c}(\sigma', \zeta))^{(n+l+1)} - (\boldsymbol{v}(\lambda, \zeta) + \boldsymbol{c}(\sigma, \zeta))^{(m-l+1)}}{2} \right|$$

$$\frac{1}{n-k} \left| \sum_{i=1}^n \frac{(\boldsymbol{v}(\lambda', \zeta) + \boldsymbol{c}(\sigma', \zeta))_i - (\boldsymbol{v}(\lambda, \zeta) + \boldsymbol{c}(\sigma, \zeta))_i}{2} \right| + kA \Bigg)$$

$$\leqq \rho_n \frac{1}{n-k} E_\zeta \left(\left| \sum_{i=1}^n \frac{(\boldsymbol{v}(\lambda', \zeta) + \boldsymbol{c}(\sigma', \zeta))_i - (\boldsymbol{v}(\lambda, \zeta) + \boldsymbol{c}(\sigma, \zeta))_i}{2} \right| + kA \right) + (1 - \rho_n)A,$$

where ρ_n is the probability that there are at least k of G and B widgets. By Assumption 5, $\rho_n \to 1$ as $n \to \infty$. Using in addition Assumption 4, we see that for all $\lambda, \lambda' \in \Lambda$, $\sigma, \sigma' \in \Sigma$, the right-hand side of the above inequality goes to zero as $n \to \infty$. Therefore,

$$u_j^n(\lambda', \sigma') - u_j^n(\lambda, \sigma) - \frac{E_\zeta S(\boldsymbol{v}(\lambda', \zeta), \boldsymbol{c}(\sigma', \zeta)) - E_\zeta S(\boldsymbol{v}(\lambda, \zeta), \boldsymbol{c}(\sigma, \zeta))}{2} \to 0 \text{ as } n \to \infty.$$

The remainder of the proof is analogous to that of Corollary 1. ||

Acknowledgement. This paper is a revised and extended version of Chapter 3 of my 1995 Ph.D. thesis at Harvard University. I gratefully acknowledge the financial support of the Alfred P. Sloan Foundation during my dissertation work. I am greatly indebted to Oliver Hart and Eric Maskin, my dissertation advisors, for many invaluable discussions and continuous encouragement. I am also grateful to Bengt Holmstrom, Boaz Moselle, Andrei Shleifer, Steve Tadelis, Jean Tirole, Michael Whinston, and participants of seminars at Berkeley, Boston University, CalTech, Chicago, Harvard, IDEI, LSE, MIT, Northwestern, Penn, Stanford, UCSD, ULB, and Yale for valuable comments. I thank Patrick Bolton (an editor) and two anonymous referees for their thorough reading of the manuscript and suggestions for improvement. I am grateful to Federico Echenique for valuable research assistance. Any errors that still remain are my own.

REFERENCES

AGHION, P., DEWATRIPONT, M., and REY, P. (1994), "Renegotiation Design with Unverifiable Information", *Econometrica*, **62**, 257–282.

ANDERLINI, L. and FELLI, L. (1994), "Incomplete Written Contracts: Undescribable States of Nature", *Quarterly Journal of Economics*, **109**, 1085–1124.

CHE, Y.-K. and HAUSCH, D. (1997), "Cooperative Investments and the Value of Contracting (Social Systems Research Institute Working Paper No. 9714, University of Wisconsin-Madison).

CROCKER, K. J. and MASTEN, S. E. (1988), "Mitigating Contractual Hazards: Unilateral Options and Contract Length", *Rand Journal of Economics*, **19**, 327–43.

DYE, R. (1985), "Costly Contract Contingencies", *International Economic Review*, **26**, 233–250.

EDLIN, A. and REICHELSTEIN, S. (1996), "Holdups, Standard Breach Remedies, and Optimal Investment", *American Economic Review*, **86**, 478–501.

EDLIN, A. and SHANNON, C. (1998), "Strict Monotonicity in Comparative Statics", *Journal of Economic Theory*, **18**, 201–219.

FUDENBERG, D., HOLMSTROM, B. and MILGROM, P. (1990), "Short-term Contracts and Long-term Agency Relationships", *Journal of Economic Theory*, **51**, 1–31.

GROUT, P. (1984), "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach", *Econometrica*, **52**, 449–460.

HART, O. (1995) *Firms, Contracts, and Financial Structure*, (Oxford: Oxford University Press).

HART, O. and MOORE, J. (1988), "Incomplete Contracts and Renegotiation", *Econometrica*, **56**, 755–785.

HAUSDORFF, F. (1962) *Set Theory* (New York: Chelsea).

HOLMSTROM, B. (1982), "Moral Hazard in Teams", *Bell Journal of Economics*, **13**, 324–340.

JOSKOW, P. L. (1988), "Price Adjustment in Long-Term Contracts: The Case of Coal", *Journal of Law and Economics*, **31**, 47–83.

KARCHMER, M. (1989) *Communication Complexity: A New Approach to Circuit Depth* (Cambridge: MIT Press).

LAFFONT, J. J. and MASKIN, E. (1982), "The Theory of Incentives: An Overview", in W. Hildenbrand (ed.) *Advances in Economic Theory*, (Cambridge: Cambridge University Press).

MACLEOD, B. (1996), "Complexity, Contract, and the Employment Relationship" (Mimeo, Universite de Montreal).

MACLEOD, B. and MALCOMSON, J. (1993), "Investments, Holdup, and the Form of Market Contracts", *American Economic Review*, **83**, 811–837.

MASKIN, E. and MOORE, J. (1999), "Implementation and Renegotiation", *Review of Economic Studies*, **66**, 39–56.

MASKIN, E. and TIROLE, J. (1999), "Unforeseen Contingencies, and Incomplete Contracts", *Review of Economic Studies*, **66**, 83–114.

MILGROM, P. and SHANNON, C. (1994), "Monotone Comparative Statics", *Econometrica*, **62**, 157–180.

MOORE, J. (1992), "Implementation in Environments with Complete Information", in J. J. Laffont (ed.) *Advances in Economic Theory*, (Cambridge: Cambridge University Press).

MOORE, J. and REPULLO, R. (1988), "Subgame Perfect Implementation", *Econometrica*, **56**, 1191–1220.

NOLDEKE, G. and SCHMIDT, K. M. (1995), "Option Contracts and Renegotiation: A Solution to the Hold-Up Problem", *Rand Journal of Economics*, **26**, 163–179.

REY, P. and SALANIE, B. (1990), "Long-Term, Short-Term, and Renegotiation: On the Value of Commitment in Contracting", *Econometrica*, **58**, 597–619.

ROGERSON, W. (1992), "Contractual Solutions to the Hold-Up Problem", *Review of Economic Studies*, **59**, 777–793.

SEGAL, I. (1996), "Communication Complexity and Coordination by Authority" (Mimeo, UC Berkeley).

SCHWARTZ, A. (1992), "Legal Theories and Incomplete Contracts", in L. Werin and H. Wijkander (eds.), *Contract Economics*, (Oxford: Blackwell).

SUTTON, J. (1986), "Non-Cooperative Bargaining Theory: An Introduction", *Review of Economic Studies*, **53**, 709–724.

TIROLE, J. (1994), "Incomplete Contracts: Where do we Stand?" (Walras–Bowley lecture at the North American Summer Meetings of the Econometric Society in Quebec City).

WILLIAMSON, O. (1975) *Markets and Hierarchies: Analysis and Antitrust Implications* (New York: Free Press).

WILLIAMSON, O. (1985) *The Economic Institutions of Capitalism* (New York: Free Press).