

Renegotiation Design by Contract

Richard Holden[†] & Anup Malani^{††}

INTRODUCTION

Professor Ronald Coase famously asked why, if markets are an efficient method of allocating resources, does so much economic activity take place within firms.¹ Coase proposed that there are costs to using the price mechanism: transaction costs. What Coase had in mind was traditional, ex post haggling: arguing over price and quantity and committing the agreement to paper. As the literature on the boundary of the firm developed over the past three-quarters of a century, a more challenging transaction cost was identified: the so-called “hold-up problem.”² If one party to a contract makes a relationship-specific investment, the second party can use the threat not to perform to extract part of the difference between the value of the investment to the relationship and the value in its next best use. This risk of renegotiation will reduce incentives to invest and thus the value of market transactions relative to within-firm transactions.³ The holdup problem has strong empirical support,⁴ is the basis for

[†] Professor of Economics, University of New South Wales; Faculty Research Fellow, National Bureau of Economic Research.

^{††} Lee and Brena Freeman Professor of Law, The University of Chicago Law School.

We are grateful to Philippe Aghion, Douglas Baird, Scott Baker, Omri Ben-Shahar, Rosalind Dixon, Bob Gibbons, Oliver Hart, Todd Henderson, Saul Levmore, Andrei Shleifer, and Alex Stremitzer for helpful discussions and comments.

¹ R.H. Coase, *The Nature of the Firm*, 4 *Economica* 386, 390 (1937).

² For a brief review of the literature discussing the holdup problem, see Bengt Holmström and John Roberts, *The Boundaries of the Firm Revisited*, 12 *J Econ Persp* 73, 73–79 (Fall 1998).

³ See Oliver E. Williamson, *Markets and Hierarchies: Analysis and Antitrust Implications* 9–10, 26–28 (Free Press 1975); Victor P. Goldberg, *Regulation and Administered Contracts*, 7 *Bell J Econ* 426, 432–33 (1976); Benjamin Klein, Robert G. Crawford, and Armen A. Alchian, *Vertical Integration, Appropriate Rents, and the Competitive Contracting Process*, 21 *J L & Econ* 297, 297–98 (1978).

⁴ For a survey of the empirical literature, see Francine Lafontaine and Margaret Slade, *Vertical Integration and Firm Boundaries: The Evidence*, 45 *J Econ Lit* 629, 653–62 (2007).

the modern theory of the firm (or Property Rights Theory),⁵ and was an impetus for Oliver Williamson's Nobel Prize.⁶

Coase's question is fundamental for economists because of the positive and normative implications for economic efficiency. It is also an important question for lawyers: a significant body of law deals with firms, their ownership, and their interactions. It is understandable, then, that enormous attention has been paid to methods for overcoming, or at least ameliorating, the holdup problem. The dominant economic theory of the firm, known as Property Rights Theory (PRT), was pioneered by Grossman and Hart and further developed by Hart and Moore.⁷ PRT equates the firm with asset ownership and asset ownership, in turn, with residual rights of control over those assets.⁸ Although it would be incorrect to think of PRT as a formalization of the holdup problem, it certainly features heavily in the theory.

PRT assumes that contracts are *incomplete*, in the sense that some payoff-relevant states of the world are either unforeseeable by the parties or indescribable to a court.⁹ Since they cannot be (enforceably) contracted on ex ante, they must be bargained over ex post—after the state is revealed. Asset ownership affects this bargaining because it reflects outside options that parties to a transaction have if renegotiation breaks down, and this, in turn, affects incentives for ex ante relationship-specific investments. This is the manner in which asset ownership, and thus the economic notion of a firm, can help alleviate the holdup problem.

If it is possible to write a complete, enforceable contract, then there would be no holdup, and thus asset ownership would be irrelevant in determining the boundary of the firm. It is typically implicitly assumed that the parties cannot foresee all

⁵ See, for example, Oliver Hart and John Moore, *Property Rights and the Nature of the Firm*, 98 J Polit Econ 1119, 1132 (1990).

⁶ Royal Swedish Academy of Sciences, *The Prize in Economic Sciences 2009* (Oct 12, 2009), online at http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2009/press.pdf (visited Mar 2, 2014).

⁷ See generally Sanford J. Grossman and Oliver D. Hart, *The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration*, 94 J Polit Econ 691 (1986); Hart and Moore, 98 J Polit Econ 1119 (cited in note 5). Indeed, the opening sentence of Grossman and Hart is: "What is a firm?" Grossman and Hart, 94 J Polit Econ at 691 (cited in note 7).

⁸ For an overview of PRT and its applications, see generally Philippe Aghion and Richard Holden, *Incomplete Contracts and the Theory of the Firm: What Have We Learned over the Past 25 Years?*, 25 J Econ Persp 181 (Spring 2011).

⁹ See id at 182.

possible future contingencies or that not all contingencies are describable, and therefore a complete contract is not possible. Instead, the bulk of the contract-theory literature has taken a mechanism-design approach to the problem.¹⁰ That is, it asks: How can the environment, including the process under which renegotiation takes place, be structured to avoid or ameliorate holdup?

The difficulty of the problem depends on two factors. One is whether one or both parties must make specific investments and whether those investments are selfish or cooperative. If both parties must make investments, then the problem begins to resemble the moral-hazard-in-teams problem: if you give one party high-powered incentives to take effort (invest), then the other party has fewer incentives to invest.¹¹ The second factor is whether each party makes investments that increase the value of a contract to itself or to the other party. The former we call self-serving investment. The latter, called cooperative investments (or direct externalities), are double trouble. First, the investing party is subject to holdup. Second, by directly benefiting the counterparty, investment increases the counterparty's bargaining power during renegotiation.

Whereas asset ownership can solve only the problem of self-serving, one-sided investment,¹² Professor Tai-Yeong Chung and, importantly, Professor Philippe Aghion, Professor Mathias Dewatripont, and Professor Patrick Rey (ADR) demonstrate that a contract with two simple components (hereafter “renegotiation design” or “RD” contracts) can achieve optimal self-serving, bilateral investment.¹³ The first component (called a “default

¹⁰ See, for example, Eric Maskin and John Moore, *Implementation and Renegotiation*, 66 *Rev Econ Stud* 39, 39 (1999); Ilya Segal and Michael D. Whinston, *The Mirrlees Approach to Mechanism Design with Renegotiation (with Applications to Hold-Up and Risk Sharing)*, 70 *Econometrica* 1, 1 (2002).

¹¹ See Armen A. Alchian and Harold Demsetz, *Production, Information Costs, and Economic Organization*, 62 *Am Econ Rev* 777, 779–81 (1972); Bengt Holmström, *Moral Hazard in Teams*, 13 *Bell J Econ* 324, 325 (1982). Holmström showed that the problem can be solved through the use of a third party—a so-called budget breaker. But this raises its own problems having to do with collusion, sabotage, and renegotiation.

¹² See Grossman and Hart, 94 *J Polit Econ* at 710 (cited in note 7).

¹³ See Tai-Yeong Chung, *Incomplete Contracts, Specific Investments, and Risk Sharing*, 58 *Rev Econ Stud* 1031, 1032–33 (1991); Philippe Aghion, Mathias Dewatripont, and Patrick Rey, *Renegotiation Design with Unverifiable Information*, 62 *Econometrica* 257, 258 (1994). See also Georg Nöldeke and Klaus M. Schmidt, *Option Contracts and Renegotiation: A Solution to the Hold-Up Problem*, 26 *RAND J Econ* 163, 163–64 (1995) (using real options to achieve the same contractual solution). MacLeod and Malcolmson consider a different contract in which parties specify different prices for different

option”) is a trade that guarantees one party a full return to its investment and can be enforced with a specific performance remedy. The second component (called a “take-it-or-leave-it” or “TIOLI” offer) is a provision that gives all bargaining power during a renegotiation to the second party. This ensures that the second party also invests by making it a residual claimant on its investment. Moreover, Professors Aaron Edlin and Stefan Reichelstein show that, in certain circumstances, parties can ensure efficient bilateral investment simply with a default option: there is no need to give the second party all bargaining power during renegotiation.¹⁴

One drawback to all the contracts above, however, is that they cannot solve the bilateral-investment problem when investments are cooperative.¹⁵ In fact, the only context in which simple contracts are theoretically an inadequate substitute for firm boundaries is when investments are both bilateral and cooperative and it is important for parties to be able to adjust to changed circumstances.¹⁶

Of course, theoretical mechanisms are of little use if they cannot be implemented in the real world. We suspect that the reason why asset ownership gets so much attention in the literature on the holdup problem and contracts get so little attention is that economists see variation in asset ownership everywhere, while few see actual contracts that look like the theoretical contracts proposed as solutions to holdup in the mechanism-design

states of nature rather than investment levels, see W. Bentley MacLeod and James M. Malcomson, *Investments, Holdup, and the Form of Market Contracts*, 83 Am Econ Rev 811, 825–26 (1993), but require that the set of events that are verifiable be adequately rich.

¹⁴ See Aaron S. Edlin and Stefan Reichelstein, *Holdups, Standard Breach Remedies, and Optimal Investment*, 86 Am Econ Rev 478, 491–94 (1996). Indeed, it is possible to achieve one-sided investment with a default option enforced only by expectation damages. For two-sided investment, specific performance is a necessary remedy. But see Susanne Ohlendorf, *Expectation Damages, Divisible Contracts, and Bilateral Investment*, 99 Am Econ Rev 1608, 1616–17 (2009) (showing that a contract that employs stochastic per-unit prices can achieve first-best outcomes with two-sided investment even under an expectation-damages remedy). Whether there are real-world contracts that use Ohlendorf’s proposed stochastic price contract is an open question.

¹⁵ See Yeon-Koo Che and Donald B. Hausch, *Cooperative Investments and the Value of Contracting*, 89 Am Econ Rev 125, 142–43 (1999). See also William P. Rogerson, *Contractual Solutions to the Hold-Up Problem*, 59 Rev Econ Stud 777, 788 (1992). Che and Hausch in fact show that the quintessentially incomplete contract—a null contract—can be optimal. Moreover, they show that message games cannot improve upon this.

¹⁶ Even then, contracts may not be worse than firm boundaries if managers are highly specialized or the manager of either of the firms does not have extra time to run the other firm. See Grossman and Hart, 94 J Polit Econ at 706 (cited in note 7).

literature. In this Article we propose to correct that misunderstanding about real-world contracts by describing actual contracts that implement the RD mechanism. In other words, this Article demonstrates that RD contracts are not merely a theoretical novelty but actually feasible and observed in the real world.

To do so, we examine each of the components of the RD mechanism. We begin by taking up contracts that implement the default option with specific performance in the RD mechanism. (Contracts that do this can also implement the Edlin-and-Reichelstein mechanism, which requires only a default option.) Specific performance is critical because it ensures that the first party actually makes an investment (which is only valuable when there is trade rather than payment of damages) and prevents overreliance.¹⁷ The challenge is that contract law disfavors specific performance as a remedy in most Anglo-Saxon countries. In these countries, however, we observe that contract law permits specific performance when items exchanged are unique;¹⁸ this protects investments that involve customization of a product. Contract law also protects certain contracts over noncustomized goods, specifically variable-quantity contracts such as output and requirements contracts, with injunctions.¹⁹ Together, these provisions provide substantial access to specific performance in the United States. Of course, in countries such as France and Spain, parties may have access to specific performance even for simpler exchanges.²⁰

Next, we take up the other component of the RD mechanism: bargaining power allocation. There are no contract-law doctrines that prevent allocation of bargaining power. The difficulty with these provisions is that they, too, may be renegotiated. However, there are methods of preventing this. First, one may contractually give the second party to a transaction control rights over a key asset of the first party. One example is to have the second party purchase and lease back a key piece of machinery. This is merely a contractual implementation of what

¹⁷ See Steven Shavell, *Damage Measures for Breach of Contract*, 11 *Bell J Econ* 466, 472 (1980) (showing that expectation damages yield excessive reliance); William P. Rogerson, *Efficient Reliance and Damage Measures for Breach of Contract*, 15 *RAND J Econ* 39, 41 (1984) (showing that specific performance addresses overreliance).

¹⁸ See UCC § 2-716 (ALI 1995).

¹⁹ See UCC § 2-716, comment 2 (ALI 1995).

²⁰ One prediction we make is that countries with greater access to specific performance should have smaller firms because they enable simple contracts to work as effective substitutes for integration.

Grossman and Hart as well as Hart and Moore call “ownership.”²¹ The example also shows that asset ownership can be combined with contracts to implement the RD mechanism; that is, revelation mechanisms may use contract together with asset ownership to address holdup with selfish, bilateral investment. Second, the parties can write contractual provisions that allow the second party to impose a large penalty on the first party at will. This is related to Williamson’s proposal to take hostages.²² Liquidated-damages clauses, even if they are fully enforceable, are inadequate because they require breach. But other provisions, such as exclusive-dealing commitments, material-adverse-change clauses, and easy-default provisions, which can be waived at will, are sufficient. What makes exclusive-dealing provisions, in particular, interesting is that they typically accompany open-quantity contracts, which function as default options in RD contracts. Third, the contract can impose penalties on one party (but not the other) for any delay in performance. These asymmetric “late fees” can make one party more impatient than the other during renegotiation and, as a result, willing to accept a less generous deal.

There is a small law literature on holdup and the incomplete-contracts problem.²³ The work most closely related to this Article is Professors Benjamin Klein and Lester Saft,²⁴ Professors Henry Hansmann and Reinier Kraakman,²⁵ and Professor Victor Goldberg,²⁶ who appreciate that outputs and requirements

²¹ Grossman and Hart, 94 J Polit Econ at 692–95 (cited in note 7); Hart and Moore, 98 J Polit Econ at 1120–21 (cited in note 5).

²² See Oliver E. Williamson, *Credible Commitments: Using Hostages to Support Exchange*, 73 Am Econ Rev 519, 519–20 (1983).

²³ See, for example, Avery Wiener Katz, *The Option Element in Contracting*, 90 Va L Rev 2187, 2220 (2004), citing Nöldeke and Schmidt, 26 RAND J Econ 163 (cited in note 13); Georg Nöldeke and Klaus M. Schmidt, *Sequential Investments and Options to Own*, 29 RAND J Econ 633, 634 (1998) (recognizing that options embedded in contracts have many uses, including the ability to address holdup problems); Albert Choi and George Triantis, *Strategic Vagueness in Contract Design: The Case of Corporate Acquisitions*, 119 Yale L J 848, 892–97 (2010) (explaining how contract terms designed to ensure that corporate targets make certain noncontractible investments prior to being acquired work better when written using vague language rather than specific language).

²⁴ See generally Benjamin Klein and Lester F. Saft, *The Law and Economics of Franchise Tying Contracts*, 28 J L & Econ 345 (1985).

²⁵ See generally Henry Hansmann and Reinier Kraakman, *Hands-Tying Contracts: Book Publishing, Venture Capital Financing, and Secured Debt*, 8 J L, Econ & Org 628 (1992).

²⁶ See generally Victor P. Goldberg, *Desperately Seeking Consideration: The Unfortunate Impact of UCC Section 2-306 on Contract Interpretation*, 68 Ohio St L J 103 (2007).

contracts can encourage specific investment, though they focus on unilateral investment.²⁷ Importantly, these papers do not connect such contracts with the mechanisms discussed in the theoretical literature on incomplete contracts. Our aim is to connect these contracts with the mechanisms they implement, so that we can identify the broader set of real-world contracts that address problems of bilateral, noncooperative, specific investment.

The remainder of the Article is organized as follows. Part I outlines the connection between the theory of the firm and the bilateral-investment problem. It also explains the contract-theory literature on selfish, bilateral investment. Thereafter the Article focuses solely on selfish rather than cooperative specific investment. Part I provides an example of holdup and how RD contracts solve that problem. Part II provides a numerical example to show how vertical integration and RD contracts work. Part III examines trades that can be enforced by injunction and support RD or Edlin-and-Reichelstein contracts. Part IV shows how contracts can allocate bargaining power as required by RD contracts. The Conclusion highlights areas for future research.

I. FIRM BOUNDARIES AND THE HOLDUP PROBLEM

A simple example illustrates the holdup problem.²⁸ Suppose that an electricity-generating plant could increase the value of a transaction to buy coal from a mine by locating next to the mine to reduce the cost of shipping purchased coal. After the electricity-generating plant was built, the coal mine might successfully demand a higher price for coal.²⁹ Because the generator cannot relocate without great expense, it would have to pay a higher price. Realizing this, the generator might not make value-creating, relationship-specific investments, because it bears the full cost of these, but anticipates that some, perhaps large, portion of those benefits will be extracted by the mine through the higher price *ex post*. If the generator firm vertically integrates

²⁷ See Alexander Stremitzer, *Standard Breach Remedies, Quality Thresholds, and Cooperative Investments*, 28 *J L, Econ & Org* 337, 338–40 (2012); Hansmann and Kraakman, 8 *J L, Econ & Org* at 629 (cited in note 25); Goldberg, 68 *Ohio St L J* at 107 (cited in note 26).

²⁸ The coal-mine–electricity-generator example is a classic one in this literature. It dates, in a formulation much like this one, to Paul L. Joskow, *Vertical Integration and Long-Term Contracts: The Case of Coal-Burning Electric Generating Plants*, 1 *J L, Econ & Org* 33, 46–48 (1985).

²⁹ In this example, it is assumed that the mine’s location is fixed.

with the mine, however, the combined firm can locate the generator next to the mine without fear of holdup.

The holdup problem becomes more severe when both parties to the contract are in a position to make specific investments that can increase the gains from trade. Because each side can be held up by the other, there is the possibility that neither makes the investment and even greater surplus is lost.³⁰ In our example, suppose that the generator needed a particular grade of coal to operate. Once the mine made an investment to reduce its cost of producing that grade of coal, the generator could insist on a discount on the price of coal. This risk of holdup would discourage the mine from making (at least to the efficient level, where the marginal benefit equals the marginal cost) an investment to refine coal particularly for the generator—again because the mine bears the full cost of the investment, but gets only part of the benefit.

In a seminal article by Grossman and Hart, followed by an important and elegant contribution by Hart and Moore (henceforth, collectively referred to as GHM), the role of asset ownership in alleviating holdup was brought to the fore. GHM's main insight is that vertical integration can solve the investment incentives of one party, but not both. For example, if the generator acquired the mine, it could ensure that the mine did not hold it up when it moved the electricity plant next to the mine. However, the manager of the generator could not give the manager of the mine (now an employee) the correct incentives to process the coal.³¹

In the decade following GHM, a number of scholars proposed alternative solutions to the problem of contractual incompleteness.³² Chung and ADR showed, quite remarkably, that it is

³⁰ Or in a setting with continuous rather than discrete investment choices, both parties will underinvest relative to the first best.

³¹ Although GHM speak of asset ownership, their theory is actually more general. They define ownership as having residual rights to joint surplus. To the extent that these rights can be allocated by contract, short of integration, their insights suggest that contracts can also be used to achieve second-best efficiency. Thus, it is sufficient in our example that the generator acquires control of a key asset via contract, for example, the drill bit or a critical computer required by the mine.

³² For example, Professors Eric Maskin and Jean Tirole argued that subgame perfect implementation can lead to truthful revelation of information that is observable to the contracting parties but not verifiable to a court. See Eric Maskin and Jean Tirole, *Unforeseen Contingencies and Incomplete Contracts*, 66 *Rev Econ Stud* 83, 93 (1999), appealing to John Moore and Rafael Repullo, *Subgame Perfect Implementation*, 56 *Econometrica* 1191 (1988). One concern that has been raised with the Maskin and Tirole

possible to write a contract that makes both parties to a transaction the residual claimant vis-à-vis their investment.³³ Such a contract has two parts. One party is given a default option that it can invoke regardless of what happens when parties renegotiate prices. This default option (DO) gives the less important party the right to demand an exchange of a fixed quantity of goods at a fixed price. This DO is constructed precisely so that it is more valuable when the party makes its specific investment and guarantees that party the surplus from that investment. The second party is given all the bargaining power during the renegotiation phase, through the ability to make a take-it-or-leave-it (TIOLI) offer. Since it has all the bargaining power it is the residual claimant on its investment. Together these components constitute an RD contract.

In our example, the RD contract would specify a default price and quantity of coal to be traded if the parties disagree at the renegotiation phase. The contract would then allow the generator to make a take-it-or-leave-it offer to purchase coal after the specific investments have been made (which are assumed to be observable to the mine and generator, but not verifiable to a court). The offer would have to be good enough to ensure that the mine did not invoke the default option. Besides that, however, the generator's offer could extract all the remaining gains from trade. With the DO correctly crafted, both sides would get a surplus that compensates them for their respective investments. Thus, the first best is achieved—just as if the investments themselves could be contracted on.

Edlin and Reichelstein show that the first component of the RD may be sufficient to achieve a first-best level of specific investment. Their insight is that, if a trade is specified before all uncertainty about its value is resolved, then there will be some states of the world in which renegotiation would yield a trade quantity that is higher than that which the parties specified in the contract. Likewise, there will be some states in which renegotiation would yield a trade quantity lower than specified in the contract. In states with a higher renegotiation quantity, both parties would experience ex post a higher return on their

approach is that the procedures they identify are very complicated and have not, as yet, been found in reality. Additionally, Aghion, Dewatripont, and Rey show that such mechanisms are nonrobust to arbitrarily small perturbations from common knowledge. Aghion, Dewatripont, and Rey, 62 *Econometrica* at 260 (cited in note 13).

³³ See note 13 and accompanying text.

investments; in states in which the quantity is lower, they would experience ex post a lower return on their investments.³⁴ If the contractually specified quantity is set at the expected quantity yielded from renegotiation, then, under certain conditions on the surplus generated by investment,³⁵ the specified quantity gives just the right incentives to invest.³⁶ The prospect of renegotiation does not undo those incentives because renegotiation is as likely to raise the ex post return to investment as it is to lower it.

II. SOLUTIONS TO THE HOLDUP PROBLEM

A. Asset Ownership and Property Rights Theory

In this Section we provide a simple numerical example of a hypothetical production process that highlights the economic costs of contractual incompleteness and demonstrates how the logic of the GHM theory applies. This example is based heavily on that in Aghion and Holden.³⁷

There are two parties, a Buyer (B) and a Seller (S) of a widget, and two assets: a widget-making machine (WM) and a final-good-making machine (FM). S uses the WM to make a widget, which B uses along with the FM to produce a single unit of a final good. A consumer values this final good at some amount v .

³⁴ These returns take the allocation of ex post surplus between the parties into account.

³⁵ Specifically, the assumption Edlin and Reichelstein make is that, from either the seller's or buyer's perspective, additional quantity has a constant effect on the private marginal value of investment, and additional investment has a constant effect on the private marginal value of quantity. Edlin and Reichelstein, 86 *Am Econ Rev* at 492 (cited in note 14). (Technically, their assumption is that the function f describing the cost to the seller or the value to the buyer from higher quantity q or investment l satisfies the condition that f_{lq} is independent of both q and the state of nature for both buyer and seller, where subscripts indicate partial derivatives.)

³⁶ The trade specified in the Edlin-Reichelstein model will be different than the one specified in the RD model because the optimal trade has to adjust for the returns the seller will obtain after renegotiation. That return depends on the seller's bargaining power, and the two contracts assume two different levels of bargaining power for the seller. The Edlin-Reichelstein contract takes the seller's bargaining power during renegotiation as given, meaning it can range anywhere between zero bargaining power and full bargaining power (outside the default option). Chung and ADR use a second tool, the take-it-or-leave-it provision, to ensure that the seller has zero bargaining power (outside the default option) in the renegotiation. The two contracts overlap only when the seller's bargaining power in the renegotiation would be zero even without the take-it-or-leave-it provision.

³⁷ See Aghion and Holden, 25 *J Econ Persp* at 183–84 (cited in note 8).

S has human-capital characteristics such that she can make a privately costly investment that makes the widget cheaper to produce by enhancing the WM. This investment costs her \$5. If S makes the investment then the widget costs \$10 to produce. If she does not make the investment the widget costs \$16 to produce. Similarly B has human-capital characteristics such that he can make a privately costly investment that enhances the FM, and affects the amount consumers are willing to pay for the final good. One can think of this as enhancing the quality of the final good. This investment also costs \$5. If B makes the investment then consumer valuation v is \$40; if not, it is \$32. Only B can make the revenue-enhancement investment, and only S can make the cost-reduction investment.

B and S would like to write a contract that specifies that each party make its respective investment. If that were possible then the total surplus in the relationship would be $\$40 - \$10 - \$5 - \$5 = \$20$ —that is, the final good is sold for \$40, it costs \$10 to produce, and B and S each incur an investment cost of \$5.

Suppose now, that contracts are incomplete so that the parties cannot contract to make the investments (or on a cost-sharing rule). This might be justified if, say, B and S know whether the investments have been made—in GHM parlance they are “observable” to B and S—but a third party, like a court, cannot determine whether the investments have been made. They are not “verifiable,” in GHM parlance.³⁸ This might be the case if specific technical knowledge or real-time observations are required to determine whether the investments have been made.

Since B and S cannot contract before investments are made, they will have to bargain about the price that B pays to S for the widget *after* the investment stage. Now suppose that B and S are nonintegrated in the sense that S owns the WM and B owns the FM. At the bargaining stage they will split whatever surplus is generated 50:50.³⁹

Will B invest? No. If he does invest, B will bear a private cost of \$5, but gets half of the increase in surplus of

³⁸ Grossman and Hart, 94 J Polit Econ at 694–95 (cited in note 7); Hart and Moore, 98 J Polit Econ at 1125–26 & n 6 (cited in note 5).

³⁹ See Aghion and Holden, 25 J Econ Persp at 184 n 2 (cited in note 8) (“This split arises in a situation of Nash bargaining because B cannot produce the final good without the widget from S, and S has no use for the widget if it is not sold to B. Therefore, both B and S have zero outside options.”).

$\$40 - \$32 = \$8$, or $\$4$. Similarly, S bears a cost of $\$5$ by investing, but gets an increased payoff of $\$(16 - 10)/2 = \3 in the bargaining. So she too, will not invest. Since neither invests, total surplus is $\$32 - \$16 = \$16$.

Now consider what happens if B and S are vertically integrated, with S owning the FM as well as her WM (forward vertical integration). She no longer needs to bargain with B; she *owns* the FM and can run a widget through it at her will. She still, of course, lacks the human capital to make the value-enhancing investment—only B can do that. So, S will receive *all* the increased surplus from investing in cost reduction (that is $\$16 - \$10 = \$6$), and thus will be prepared to invest. However, B will not invest as he will get none of the benefit of making the final good more valuable. S cannot compel B to invest, nor contract on B making the investment. Total surplus is therefore $\$32 - \$10 - \$5 = \17 . This is larger than under nonintegration, so forward vertical integration is desirable.

In fact, B owning the WM (backward integration) does better still. Now B invests, but S does not, for the same reason as above. This yields a total surplus of $\$40 - \$16 - \$5 = \19 . This is not as good as if contracting were possible (that would yield a surplus of $\$20$), but it is better than the other possible ownership structures.

B ownership of both assets is preferable to S owning both because B's investment is relatively more important than S's. Both cost $\$5$, but B's has a benefit of $\$40 - \$32 = \$8$, whereas S's has a benefit of only $\$16 - \$10 = \$6$.

This is the key message of GHM: asset ownership affects how surplus is shared in renegotiation, and with the parties anticipating this, it affects incentives for *ex ante* investment.

B. Renegotiation Design

Chung and ADR argue that, even without asset ownership, it is possible to improve efficiency in the face of contractual incompleteness.⁴⁰ The key to their idea is the observation that the efficiency loss above arose because it was not possible to make *both* B and S the residual claimants on their investments. Forward integration made S the residual claimant, but not B. Backward integration achieved the converse. ADR's goal is to design the environment in which the renegotiation takes place

⁴⁰ See note 13 and accompanying text.

so that *both* parties are residual claimants. Doing so would achieve an outcome superior to GHM—it would get the first-best outcome (that is, a surplus of \$20 in the above example).

ADR ensure that both parties are residual claimants in the following way. Suppose that, although B and S cannot contract on their investments, they can specify a default: how many widgets will be traded and at what price. Now consider the following contractual mechanism: B makes an offer to S consisting of a quantity of widgets to be traded and a price per widget. If S accepts the offer then trade takes place on those terms, but if she rejects then trade takes place under the terms specified in the default. The trick is to construct the price of trade in the default such that both B and S have the appropriate investment incentives.

In the example of the previous Section, there is use for only one possible widget, so it is convenient to think of setting the default probability that a widget will be traded. The reasoning extends easily to the case of multiple widgets and nonprobabilistic trade. Set the default probability of trade at $5/6$, and set the default price at $\$23 \frac{1}{3}$. These numbers appear to have come out of thin air, but as we will show below, they are designed to do two things: split the ex ante surplus evenly between B and S, and ensure that S has the appropriate incentives to make the cost-reducing investment. We offer an algebraic derivation in the Appendix.

What is the best offer for B to make when he gets that opportunity? Since he is making a take-it-or-leave-it offer he has all the bargaining power and will thus want to trade the efficient number of widgets, which here is 1. Since B has all the bargaining power he will extract S's entire surplus, leaving her indifferent between the default option and accepting B's offer. S will thus anticipate getting her payoff under the default option whatever happens.

Now work backwards, and consider S's decision whether to invest at the earlier stage. If S invests she gets a payoff equal to $\$23 \frac{1}{3} - (5/6) \times \$10 - \$5 = \10 (price minus probability of trade multiplied by the cost of production, minus the investment cost). If she does not invest she gets a payoff equal to $\$23 \frac{1}{3} - (5/6) \times \$16 = \$10$. So she is willing to invest. Note that this indifference could be easily broken by tweaking the $5/6$ default ever so slightly.

What about B? He, of course, is the residual claimant on her investment: obtaining \$8 if he makes it for the cost of \$5. So B also makes the efficient investment. Thus, total surplus is $\$40 - \$10 - \$5 - \$5 = \$20$. Remarkably, this is the first best—the same as if contracts were not incomplete!

It is, at first glance, rather surprising that S finds it optimal to invest despite having *none* of the bargaining power in the renegotiation. The key is her ability to reject B's offer and trigger trade under the terms of the default option. The default option becomes more appealing when the widget is low cost, which happens precisely when S invests. In other words, the presence of the default option makes S's payoff sensitive to her investment and internalizes the externality that would exist under backward vertical integration, which we saw above led to S not investing and hence a total surplus of \$19.

III. SPECIFIC PERFORMANCE

Chung and ADR demonstrate that a contract with two components—a default option and a take-it-or-leave-it provision—can ensure that both parties make selfish, specific investments. Further, Edlin and Reichelstein show that, under certain conditions on the return on specific investment, just the first component is sufficient to protect bilateral specific investment. In this Part we show how the default option can practically be implemented. In the next we turn to implementation of take-it-or-leave-it provisions.

In order to simplify our exposition, we will employ a simple example involving the sale of a widget from a seller to a buyer. We shall assume, as in the numerical example in Part II, that both parties can make an *ex ante* specific investment that increases joint surplus (net of investment) from the transaction, but that the investment by the buyer generates more surplus than the investment by the seller.⁴¹ We focus on the case of bilateral investment rather than unilateral investment because any solution to the bilateral-investment problem can trivially be employed to solve the unilateral-investment problem.⁴²

⁴¹ This example does not compromise the generality of our arguments, as one can easily change the labels of the parties so that the seller makes an equal-valued or even higher-valued investment without compromising the arguments we make herein.

⁴² See, for example, Aghion, Dewatripont, and Rey, 62 *Econometrica* at 276 (cited in note 13) (showing how their bilateral-investment framework may apply to unilateral-investment models for firm-union relationships).

Whether one pursues the default option in an RD contract or the trade in an Edlin-Reichelstein contract, the key to successfully implementing the contract is the seller's ability to obtain specific performance of the contractually specified trade. In the RD proposal, specific performance is required to delineate between a seller that actually makes a specific investment and one that does not. If the seller anticipates that it will be held up and that it will be able to sue for damages, it will have no incentive to invest. The reason is that it gets damages whether or not it invests because, recall, investment is nonverifiable. Since investment has a positive cost, the seller would prefer to obtain damages without making an investment than receiving the same damages after incurring the cost of investment.⁴³ In the Edlin-Reichelstein proposal, specific performance is required because damages lead to overinvestment by the seller. With damages, while the seller is insured against the risk that the optimal *ex post* trade is lower than the contract envisions, the buyer is not.⁴⁴ Thus the seller will tend to overinvest relative to the buyer. This argument is similar to the argument made for why expectation damages induce overreliance, though it generalizes to the bilateral-investment case.⁴⁵

Given the centrality of specific performance to the RD default options and Edlin-Reichelstein contracts, the legal question we take up in this Part is: What sorts of trades are enforceable by injunctions? This is an age-old question in the legal literature, so our contribution is not to provide a new answer. Rather, our aim is to show how the answer fits into the larger debate over the extent to which simple contracts can substitute for ownership and firm boundaries as solutions to the problem of relationship-specific investment. In the process, we think we will make two smaller contributions. First, we provide a novel reason for courts to offer specific performance as a remedy for contract breach. Second, we provide a novel reason for the use of so-called variable-quantity contracts.

In Anglo-Saxon countries, the general rule governing remedies in contract law is that specific performance is an exceptional remedy granted only in specific circumstances, such as if

⁴³ The buyer can observe that an investment was not made, but cannot sue to force the investment or defend against breach with an argument that investment was not made because, again, investment is nonverifiable.

⁴⁴ See Edlin and Reichelstein, 86 *Am Econ Rev* at 941–42 (cited in note 14).

⁴⁵ See Shavell, 11 *Bell J Econ* at 472 (cited in note 17).

courts cannot estimate monetary damages due to breach.⁴⁶ Oliver Wendell Holmes put it succinctly:

The only universal consequence of a legally binding promise is, that the law makes the promisor pay damages if the promised event does not come to pass. In every case it leaves him free from interference until the time for fulfillment has gone by, and therefore free to break his contract if he chooses.⁴⁷

In certain civil law countries⁴⁸ and in China,⁴⁹ however, this priority is reversed and specific performance is favored. This is not accomplished by the use of police to force a breaching party to perform an agreed action or to imprison breaching parties. Rather, it is accomplished by strategic use of monetary payments. In France, for example, courts have historically taken either of two tacks. First, they allow private parties to ensure specific performance by enforcing contracts that include large penalty clauses triggered by nonperformance.⁵⁰ Second, courts themselves impose a monetary judgment for nonperformance—an *astreinte*—that increases with each day of delay in performing a contract.⁵¹

Even in the United States, there are exceptions to the general rule favoring damages.⁵² First, specific performance may be

⁴⁶ See Alan Schwartz, *The Case for Specific Performance*, 89 Yale L J 271, 272 (1979). See also *Klein v Pepsico, Inc.*, 845 F2d 76, 80 (4th Cir 1988) (citing the maxim that “specific performance is inappropriate where damages are recoverable and adequate”).

⁴⁷ O.W. Holmes Jr., *The Common Law* 301 (Little, Brown 1881).

⁴⁸ See, for example, James Beardsley, *Compelling Contract Performance in France*, 1 Hastings Intl & Comp L Rev 93, 111 (1977); Shael Herman, *Specific Performance: A Comparative Analysis (I)*, 7 Edinburgh L Rev 5, 12–13 (2003) (observing that specific performance is preferred in Spain). This is not, however, a universal rule in civil law countries. See John P. Dawson, *Specific Performance in France and Germany*, 57 Mich L Rev 495, 525 (1959) (noting that, unlike France, Germany favors damages over specific performance).

⁴⁹ See Ni Zhu, *A Case Study of Legal Transplant: The Possibility of Efficient Breach in China*, 36 Georgetown J Intl L 1145, 1146 (2005). See also General Principles of the Civil Law of the People’s Republic of China §111 (1999), online at http://www.china.org.cn/china/LegislationsForm2001-2010/2011-02/12/content_21908031.htm (visited Mar 2, 2014).

⁵⁰ See Beardsley, 1 Hastings Intl & Comp L Rev at 102 (cited in note 48).

⁵¹ See *id.* at 95–96. The penal clause was held by French courts to be immutable (and thus free from renegotiation) until the mid-1970s, when courts turned back to the *astreinte* to ensure specific performance. See *id.* at 102–03, citing French Civil Code § 1231.

⁵² One exception we do not dwell upon is that specific performance is a standard remedy for contracts concerning real estate. A buyer of real estate may obtain an injunction

available for goods that are “unique”⁵³ or goods that are identifiable and for which the buyer cannot obtain cover.⁵⁴ The intuition behind the exception is that damages will be inadequate because there is a thin or nonexisting market for a seller’s good or the buyer is unable to reasonably procure a good substitute.⁵⁵ This exception will cover many cases involving specific investment. Often the specific investment made by the seller or buyer is intended to customize a good, making it unique and hard to replace. For example, the buyer may ask the seller to make an input that is custom made for seller’s noncustom final output. Either the seller will not be able to resell the input or the buyer will not be able to procure it from another supplier in a reasonable time.

While the unique-good exception covers many cases of specific investment, it may not cover all cases of specific investment. The reason is that the contract is written over a widget, not an investment. Yet specific investment implies only that the investment is specific, that is, unique—not that the widget is unique. While unique investment and unique goods may be correlated, they are not perfectly so. For example, the seller may decide to adopt inventory-management software recommended by the buyer to track his shipments of a widget to the buyer, but if the widget itself is not unique, the court will not give the seller specific performance on the sale of the widget.⁵⁶

Yet there is a trend in US courts to more liberally allow specific performance that reduces the number of cases of specific performance that might not fit the unique-good exception.⁵⁷ The

requiring the seller to transfer title to land to the buyer. See Restatement (Second) of Contracts § 360, comment e (2008).

⁵³ UCC § 2-716(1) (ALI 1995). See also Restatement (Second) of Contracts § 359 (2008).

⁵⁴ UCC § 2-716(3) (ALI 1995). In this case the buyer is said to have a right of replevin to those goods.

⁵⁵ Restatement (Second) of Contracts § 360 (2008). The good does not have to be perfectly unique. It is sufficient that the good is rare or very limited in supply. See, for example, *Sedmak v Charlie’s Chevrolet, Inc.*, 622 SW2d 694, 699–700 (Mo App 1981) (requiring specific performance of sale of a limited edition Corvette produced to commemorate the selection of the Corvette as a Pace Car for the Indianapolis 500).

⁵⁶ Nor will the court give adequate reliance damages for the investment if it is truly nonverifiable.

⁵⁷ See, for example, Restatement (Second) of Contracts § 359, comment a (2008); UCC § 2-716, comment 1 (ALI 1995). The liberalization is endorsed by James J. White and Robert S. Summers, *Uniform Commercial Code* § 6-6 at 212–21 (West 4th ed 1995), citing Schwartz, 89 Yale L J 271 (cited in note 46).

now-liberal attitude is expressed nicely by the Eighth Circuit in *Laclede Gas Co v Amoco Oil Co*:⁵⁸

Generally the determination of whether or not to order specific performance of a contract lies within the sound discretion of the trial court. However, this discretion is, in fact, quite limited; and it is said that when certain equitable rules have been met and the contract is fair and plain “specific performance goes as a matter of right.”⁵⁹

Such language may, in the general run of cases, swallow up the few cases in which specific investment does not imply unique goods.

A second exception to the preference for damages over specific performance is the use of variable-quantity contracts, such as requirements or output contracts. These are contracts that indicate the price of exchange but omit the quantity of goods that the parties agree to trade.⁶⁰ A requirements contract obligates a seller to provide any quantity of a good that the buyer demands rather than a specific amount of that good. An outputs contract does the opposite, obligating the buyer to purchase any quantity of a good that the seller produces. The UCC § 2-716 comment 2 states that a requirements or outputs contract is the typical setting in which modern courts give specific performance.

A classic case is *Eastern Air Lines, Inc v Gulf Oil Corp*.⁶¹ Eastern signed a requirements contract with Gulf Oil requiring Gulf to provide any jet fuel that Eastern demanded in certain specified cities, so long as Eastern’s demands were reasonable and in good faith. In return, Eastern would agree to purchase jet fuel from only Gulf in those cities.⁶² The price of the jet fuel was set by reference to the average posted price for West Texas sour crude by three oil companies. In response to the oil shocks of the early 1970s, the US government regulated the posted price of sour crude so that it was below the amount that it cost oil companies to procure. Gulf responded by demanding that Eastern pay more for its oil. Eastern sued and won an injunction requiring that Gulf provide oil at the agreed posted price. The court took it as well accepted that an injunction was required, even

⁵⁸ 522 F2d 33 (8th Cir 1975).

⁵⁹ *Id* at 38–39 (citations omitted).

⁶⁰ UCC § 2-306(1), comment 2 (ALI 1995) (approving these contracts).

⁶¹ 415 F Supp 429 (SD Fla 1975).

⁶² *Id* at 434–35.

though the good in question was a commodity rather than a unique good.⁶³

There are two obstacles to the use of variable-quantity contracts as a mechanism to obtain specific performance for default options in RD contracts or trades in Edlin-Reichelstein contracts.⁶⁴ First, the rationale typically given for why courts may grant equitable relief in a variable-quantity contract is that these contracts are long term and that it would be difficult for a nonbreaching party to find a substitute long-term partner or one that would grant it the same level of flexibility.⁶⁵ This logic would seem to bar equitable relief for short-run variable-quantity contracts and thus short-run default options or trades. We do not think this constraint is too binding. For one thing, many contracts with specific investment will be long term in order to amortize the cost of the specific investment. Another thing is that the rationale typically given by courts that award specific performance for breaches of variable-quantity contracts is related to the rationale for parties seeking specific performance after specific investment: it would be hard for the nonbreaching party to find another partner that could utilize the specific investment that the party made.⁶⁶ Thus, the courts' rationale for specific performance in variable-quantity contracts cases does not depend on the contracts being long-term and might support specific performance in short-term contracts.

⁶³ Id at 442–43.

⁶⁴ It is sometimes thought that, in a variable-quantity contract, specific performance is granted only to vindicate an exclusive-dealing obligation rather than the variable-quantity purchase or sale obligation. (The exclusive-dealing requirement is either included explicitly in the contract or is a duty implied by good faith requirements imposed on variable-quantity contracts by courts. See UCC § 2-306 (ALI 1995).) However, there are numerous examples in which courts grant specific performance for the benefit of—rather than against—the party under an exclusive-dealing obligation. See *Laclede Gas*, 522 F2d at 40–41; *Eastern Air Lines*, 415 F Supp at 442–43. Moreover, it is not true that all variable-quantity contracts require the party with the right to demand purchase or supply take on the obligation to deal exclusively with the counterparty to the agreement. See, for example, Allen Blair, “You Don’t Have to Be Ludwig Wittgenstein”: *How Llewellyn’s Concept of Agreement Should Change the Law of Open-Quantity Contracts*, 37 Seton Hall L Rev 67, 73 (2006).

⁶⁵ See Richard E. Speidel, *Court-Imposed Price Adjustments under Long-Term Supply Contracts*, 76 Nw U L Rev 369, 390 (1981).

⁶⁶ The nonbreaching party does not have to quantify the specific investment. Nor does the court have to be able to verify it for this argument to carry the day. The only thing required is that the court can qualitatively see that the nonbreaching party made some specific investment and that it cannot easily find another partner that would fully value that investment.

A second obstacle to the use of variable-quantity contracts to secure specific investment is that RD contracts and Edlin-Reichelstein contracts require a specific amount of trade, not a variable quantity of trade. This problem is more apparent than real. Courts impose a duty of good faith on variable-quantity contracts to ensure that parties do not produce or demand too little or too much quantity.⁶⁷ Too little or too much is judged by the previous behavior of the parties.⁶⁸ Moreover, the parties can limit one another's flexibility by including in the original contract an estimated amount of trade or a minimum or maximum amount of trade. While these protections do not ensure specific performance of a precise amount of trade, they can substantially reduce the degree of variability in a variable-quantity contract.⁶⁹

We should remark that we are not the first to note that variable-quantity contracts can protect a party's interest in an investment. Analyzing the classic outputs-contract case *Feld v Henry S. Levy & Sons, Inc.*,⁷⁰ Goldberg has argued that the court was wrong to deny summary judgment to the defendant, a bakery with the right to put any quantity of output (here: bread crumbs) to the plaintiff. The defendant was not previously in the business of producing bread crumbs. Production of crumbs required the defendant to invest in a custom oven. Goldberg explained that the plaintiff's commitment to buy all the defendant's bread crumbs provided the defendant a measure of security in its investment.⁷¹ Although this case did not strictly involve a nonverifiable investment (the contract could have required purchase of crumbs if the defendant made the oven) or perhaps even a specific investment (anyone may have been able to purchase the bread crumbs), the logic behind Goldberg's argument applies to noncontractible specific investment as well. Our contribution is merely to extend Goldberg's argument to incomplete contracts with specific investment and to explain that, in this context, it is access to specific performance that is critical.

⁶⁷ See Restatement (Second) of Contracts § 205 (2008).

⁶⁸ See UCC § 2-306(1) (ALI 1995).

⁶⁹ Moreover, the variable-quantity contract is compatible with ex post renegotiation. If the parties include an estimate of the amount to be traded, then they can use that estimate as the default option and conduct renegotiation without changing the terms of the variable-quantity contract. While small deviations from the estimate would probably not cause a court to order specific performance, large deviations would. Thus, a variable-quantity contract can prevent more significant cases of holdup.

⁷⁰ 335 NE2d 320 (NY 1975).

⁷¹ See Goldberg, 68 Ohio St L J at 105–07 (cited in note 26).

Beyond the exception for unique goods and for variable-quantity contracts, there may be a third exception to the general rule that US contract law favors damages over specific performance. Actually, it is not quite an exception as it is an opening that parties can employ. It has been suggested that parties may be able to sway courts to enforce obligations with injunctions if parties include a provision in their contract requesting specific performance.⁷² These specific performance clauses can be made even more effective by inclusion of arbitration clauses since courts are generally bound to enforce arbitration agreements.⁷³ These tactics, like the use of the unique-good exception, are strengthened by the trend towards more liberal access to specific performance in US courts.⁷⁴

Finally, there is one last strategy that parties can employ to obtain not specific performance, but the equivalent of specific performance. Common in the natural gas market is the so-called take-or-pay contract.⁷⁵ Such a contract obligates the buyer to either take and pay for a fixed quantity of a good each period or pay for that quantity of the good without taking it.⁷⁶ The seller in a take-or-pay contract does not have access to specific performance for the failure of the buyer to make a given take-or-pay payment because it can be made whole with damages equal to the amount that the buyer was obligated to pay, even though it did not take. The remedy of damages, however, is equivalent to a remedy of specific performance because it does not require the seller to provide any product. This peculiar damage award encourages the buyer to take possession of the seller's product in order to minimize the loss from the damage award: it is better to pay and take the product than to pay without taking the product. Ultimately, take-or-pay provisions induce the buyer to both pay and take possession—the same result as under specific performance.

⁷² Edward Yorio and Steve Thel, *Contract Enforcement: Specific Performance and Injunctions* § 19.3 at 19-12 to -14 (Wolters Kluwer 2d ed 2013).

⁷³ *Id.* at § 19.4 at 19-14 to -18.

⁷⁴ See notes 57–59 and accompanying text.

⁷⁵ See Scott E. Masten and Keith J. Crocker, *Efficient Adaptation in Long-Term Contracts: Take-or-Pay Provisions for Natural Gas*, 75 *Am Econ Rev* 1083, 1083 (1985).

⁷⁶ Typically the contract lasts for several periods. If there is one period in which the buyer does not take the full fixed quantity of the good, the buyer must take that amount in a subsequent period, though the buyer does not have to pay extra for it.

IV. TAKE-IT-OR-LEAVE-IT PROVISIONS

In this penultimate Part we turn to the second component of RD contracts—the take-it-or-leave-it provision—and explore how it can practically be implemented under the constraints of contract law.

Recall that the RD proposal requires that, after a default provision (enforceable by an injunction) guarantees one party (say the seller) a return on its selfish investment, the remaining party (say the buyer) must be allocated all remaining bargaining power so that it becomes a residual claimant on the transaction. This in turn encourages the buyer to make the first-best investment.

A party that can credibly make a take-it-or-leave-it offer has all the bargaining power in a relationship because such an offer can extract all the gains from trade. The seller is indifferent between the remaining options: reject the offer and get zero surplus or accept the offer and get some nominal surplus.⁷⁷ Clearly, the better-choice option is to accept, even though it leaves the buyer with much more ex post surplus. For this reason, we call the second component of the RD contract the take-it-or-leave-it provision.

The challenge with implementing a TIOLI offer is that it has to be credible. If it is not in the interest of the buyer ex post to make the offer, that is, if the buyer would be worse off if the seller rejected than the buyer would be with a more reasonable offer that the seller accepted, then the buyer's demand is not credible.⁷⁸ If it is not credible, it cannot extract all the residual surplus from a transaction and the buyer is left with suboptimal incentives to invest. Thus, the precise question we take up in this Part is how legally to make a credible take-it-or-leave-it offer.

There are three approaches we examine. The first approach is to give the buyer ownership of a critical asset required by the seller. In our mine-generator example, if the mine obtains the default option, the generator can be given ownership of a difficult-to-replace drill bit that is critical to the

⁷⁷ In the case of the RD contract, this surplus is on top of the surplus from the default option.

⁷⁸ In game-theoretic terms, the take-it-or-leave-it offer is said not to be “subgame perfect.” Eric Rasmusen, *Games and Information: An Introduction to Game Theory* 108–10 (Blackwell 4th ed 2007).

mine's operation.⁷⁹ As a result, the mine can control that drill bit and use that control to ensure it gets the upper hand in any bargaining. The generator may use ownership of the drill bit to ensure that the mine is unable to generate any more surplus unless the mine accedes to the generator's demands. In other words, ownership reduces the mine's outside option to zero.

We are cautious in promoting this approach as an example of the RD mechanism, not because it is infeasible, but rather because it requires us to suspend our central thesis that legally valid contracts—as opposed to legal integration or economic ownership—can induce specific investment in more cases than previously thought possible by economists. That said, the use of ownership as an instrument to shift bargaining power as part of an RD contract suggests that ownership and contracting can be complements. This is a useful insight, but it does not give us a full measure of how powerful contracts can be.

The second approach is to give the buyer a hostage:⁸⁰ an asset that is not critical to the buyer's continued function, but has value to the buyer nonetheless. If the buyer has the right to lay waste to that asset, then the buyer can extract from the seller ex post surplus equal in value to that asset. For example, suppose that the seller has some inventory that has market value, but is not specific to its relationship with the buyer. That inventory can be used as a hostage. If the buyer is given the right to prevent sale of that inventory (but not ownership of that inventory), then it can use the threat of preventing sale of that inventory to extract ex post surplus from the seller. The threat is credible because the buyer loses nothing by exercising the right to prevent sale of that inventory.

There are numerous examples of this sort of hostage taking. Variable-quantity contracts, for example, are typically accompanied by exclusive-dealing obligations. Specifically, the party given the right to call or put output also has the obligation not to deal with any other party. To the extent that the party with the option has excess supply or demand,⁸¹ the right to waive that

⁷⁹ Ownership of this sort need not defeat the mine's default option. The ownership provision can be written in a manner that permits the mine to use it to produce whatever quantity of coal is specified in the default option.

⁸⁰ This idea was raised in Williamson, 73 *Am Econ Rev* at 524 (cited in note 22).

⁸¹ This means supply or demand in excess of what a court would judge to be a reasonable amount (under the good faith obligation that accompanies variable-quantity contracts, see UCC § 2-306(1) (ALI 1995)), or a maximum amount set in the requirements contract.

requirement and allow the option holder to buy or sell units on the general market is an important benefit the party that granted the option may confer. The option-granting contract can use this to extract value in renegotiation.

Indeed, any contractual provision that would seem to give one party extreme control over the fate of a counterparty can be employed as part of a hostage-taking strategy. One example is broad event-of-default provisions often found in credit agreements. These provisions give conditions under which the debtor is said to be in default that are very likely to be triggered in the ordinary course of business. For example, the creditor may require onerous audit procedures with which it is difficult to comply.⁸² Another example is a covenant requiring extreme amounts of information to be furnished to the creditor in a short time frame.⁸³ Once the debtor is in default, the creditor has a right to accelerate the loan. This will often force the debtor to shut down due to financial distress unless the creditor releases the debtor from acceleration. The creditor can use this power to shut down the debtor to demand that the debtor accede to its demands in a renegotiation.

Yet another example is a contract that gives the buyer a right to see valuable confidential information (for example, trade secrets) of the seller or to license the seller's intellectual property without conditions on sublicensing. The buyer can use these rights to reduce the value of the seller's intellectual property. The buyer can either reveal the trade secret to a third party or sublicense the intellectual property to undermine the seller's ability to extract full monopoly rents from the property.⁸⁴

The main drawback of this strategy is that it may be difficult to find a hostage whose value is proportional to the residual value of the buyer and seller's transaction. If the hostage has fixed value, then it might be smaller or greater than the value of the transaction if the buyer invests, once you subtract the value

⁸² See R. Wilson Freyermuth, *Enforcement of Acceleration Provisions and the Rhetoric of Good Faith*, 1998 BYU L Rev 1035, 1036–37.

⁸³ For example, a typical clause might read: The debtor shall immediately provide to the creditor any and all financial information that the creditor may, in its sole discretion, deem reasonable. See, for example, *In re HRP Myrtle Beach Holdings, LLC*, 2008 WL 4442606, *68–71 (Bankr D Del).

⁸⁴ It is important that the buyer not be able to profit from the sale of the intellectual property, for example, by licensing it at a low price and sublicensing at a high price. This would make the buyer's threat to sublicense noncredible because the sublicensing, while driving down price for the intellectual property, would also punish the buyer.

of the default option the seller possesses. If it is smaller, then the buyer will have inadequate incentive to invest. If it is greater, then the seller will have inadequate incentive to invest because the hostage taking will reduce the seller's return to investment below the level guaranteed by the default option.⁸⁵ From this perspective, exclusive dealing may be a better hostage-taking device than easy default triggers or intellectual property rights, because its value better correlates with the value of the seller and buyer's own transaction.

A second drawback of the hostage-taking strategy is that the right must be triggered by the buyer's own specific investment. Otherwise the buyer may use the power to extract ex post surplus without itself making any investment. In that case, the hostage taking only exacerbates the holdup problem that discourages seller investment. Even if this problem can be overcome by giving the seller a more generous default option, the hostage-taking power degenerates into the equivalent of an ex ante transfer from the seller to the buyer. While even that risk can be overcome with a separate transfer, the parties are still left with no way to increase the buyer's bargaining power ex post. The trick is to vest the hostage-taking right only if the buyer makes an investment. An example of this approach is, again, exclusive dealing in variable-quantity contracts. If the party that grants the option to call or put output does not honor that option, that is, refuses to supply or purchase output when the option is exercised, it cannot ask a court to enforce the exclusive-dealing provision. Yet the option-granting party, wanting to maximize the value to itself of trade when the option is exercised, will want to invest. Thus the investment takes place before the exclusive-dealing right—the hostage—vests.

A third approach to allocating all bargaining power during renegotiation to the buyer is to impose a penalty on the seller for any delay in the renegotiations. Professor Ariel Rubinstein has demonstrated that, in sequential bargaining in which parties trade offers, if one party has a higher discount rate, the other party can extract value by delaying the negotiating process. The relatively impatient party will be willing to accept a smaller fraction of the renegotiation surplus for a quicker resolution of

⁸⁵ One possible way out is that, given the hostage taking, the seller views the default option as a way to minimize the loss from hostage taking and the default option has greater value if the seller invests than if it does not invest.

the renegotiation.⁸⁶ A higher discount rate, however, can be artificially induced by requiring the seller to pay a penalty for every day that a trade (other than the default option) is delayed. In that case, the seller will want to accede even to an extreme demand by the buyer to avoid the penalty from a delayed trade (under the renegotiation).⁸⁷

These delay penalties are relatively easy to incorporate into contracts. In some cases, however, one party is naturally more impatient than the other. Then, the impatient party is a natural candidate to receive the default option; the patient party will naturally take advantage in any ex post renegotiation. For example, in take-or-pay contracts that govern natural gas sales, the natural gas producer (seller) is often impatient. The producer is typically one of many producers that operate a given natural gas field. That field, in turn, is likely to be regulated by a unitization agreement that gives each producer a right to extract up to a fixed amount of gas each day. If the producer extracts any less than that amount, it loses the remainder of its allotment for that day. This loss is called “drainage.”⁸⁸ Every day that the producer and a pipeline (buyer) cannot agree on a purchase (other than the amount guaranteed under the take-or-pay provision), the producer loses out on its rights to the field and revenue. The pipeline, however, is not penalized for delay as it can procure the gas from another source during the renegotiation process. As a result of this asymmetric loss from delays, pipelines tend to have the advantage in renegotiation.⁸⁹ And, as predicted, the producer is the beneficiary of the take-or-pay option.

CONCLUSION

GHM emphasize the role that asset ownership plays in affecting ex post bargaining, and hence ex ante investment incentives, in states of the world not contracted on. We have argued that contracts and contracts doctrine often play a similar role in affecting ex post bargaining and hence ex ante incentives.

⁸⁶ See Ariel Rubinstein, *Perfect Equilibrium in a Bargaining Model*, 50 *Econometrica* 97, 99 (1982). See also Ingolf Ståhl, *Bargaining Theory* 40 (EFI 1972).

⁸⁷ See Aghion, Dewatripont, and Rey, 62 *Econometrica* at 264–65 (cited in note 13).

⁸⁸ Masten and Crocker, 75 *Am Econ Rev* at 1086 (cited in note 75).

⁸⁹ Indeed, natural gas contracts often impose exclusive-dealing obligations upon the gas producer, further increasing the vulnerability of the producer in renegotiation. *Id.* at 1089 n 18.

Specifically, the DO and TIOLI components that comprise an RD contract do exist in the real world, suggesting that RD contracts are feasible.

Of course, enforcement of these types of contracts is far from universal. Beginning from a hypothesis that asset ownership and RD contracts are substitute solutions to the problem of selfish, specific investment, one might predict that, when courts enforce these contracts properly, one will observe less use of asset ownership and more use of contracts to provide investment incentives. That is, one might observe less integration and more market-based transactions.

Conversely, property rights are not universally enforced either. This is true even in developed economies with strong rule of law. But disparities in the enforcement of property rights are even more evident from comparisons between countries.⁹⁰ The substitutability of asset ownership and RD contracts implies that, as property right enforcement improves, all else equal, one will observe increased use of assets and less use of contracts, that is, more integration and fewer market-based transactions.

Whether this is true is, of course, an empirical question. We leave it to future research to test the two predictions above. In any case, the asset-contracts-substitutability hypothesis suggests that it is the *interplay* between protection of property rights and enforcement of certain kinds of contracts that determines the answer to Coase's famous question: Why do some transactions happen in firms and others in markets?

Finally, we have focused in this Article on holdup and incentives for ex ante investments. Coase himself also spoke of ex post frictions such as haggling and search. Williamson also emphasized ex post frictions.⁹¹ Recently, formal models of ex post frictions have emerged. These models involve "aggrievement" and "shading" whereby a contract acts as a reference point that anchors what the parties feel entitled to,⁹² or involve renegotiation under asymmetric information.⁹³ It would be interesting to

⁹⁰ See Rafael La Porta, et al, *Law and Finance*, 106 J Polit Econ 1113, 1115–16 (1998).

⁹¹ See Oliver E. Williamson, *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting* 20–21 (Free Press 1985).

⁹² Oliver Hart and John Moore, *Contracts as Reference Points*, 123 Q J Econ 1, 8 (2008); Oliver Hart, *Economica Coase Lecture: Reference Points and the Theory of the Firm*, 75 *Economica* 404, 407–08 (2008).

⁹³ See Philippe Aghion, et al, *Subgame-Perfect Implementation under Information Perturbations*, 127 Q J Econ 1843, 1844 (2012).

study what, if any, role different types of real-world contracts can play in ameliorating these frictions.