

ARTICLE

CONTRACTING FOR INNOVATION: VERTICAL DISINTEGRATION AND INTERFIRM COLLABORATION

Ronald J. Gilson*

Charles F. Sabel**

Robert E. Scott***

Rapidly innovating industries are not behaving the way theory expects. Conventional industrial organization theory predicts that, when parties in a supply chain have to make transaction-specific investments, the risk of opportunism will drive them away from contracts and toward vertical integration. Despite the conventional theory, however, contemporary practice is moving in the other direction. Instead of vertical integration, we observe vertical disintegration in a significant number of industries, as producers recognize that they cannot themselves maintain cutting-edge technology in every field required for the success of their products. In doing this, the parties are develop-

* Marc and Eva Stern Professor of Law and Business, Columbia University; Charles J. Meyers Professor of Law and Business, Stanford University; European Corporate Governance Institute.

** Maurice T. Moore Professor of Law, Columbia University.

*** Alfred McCormack Professor of Law and Director, Center for Contract and Economic Organization, Columbia University.

We are especially grateful to Victor Goldberg, who made very substantial contributions to this Article. Michael Klausner, Richard Langlois, Curtis Milhaupt, Richard Nelson, Daniel Raff, Alan Schwartz, Yane Svetiev, Eric Talley, and George Triantis provided helpful comments on an earlier draft, as did participants in workshops at Wharton, MIT, Fordham, and Pennsylvania Law Schools, and at the Columbia Conference on Business Law and Innovation. Gary Herrigel, Josh Whitford, and Jonathan Zeitlin generously kept us abreast of their research on changes in global supply chains. We are indebted to Matthew Jennejohn for his research ingenuity. In the course of our investigations, we encountered, and benefited from exchanges with, others grappling with related themes. See Margaret M. Blair & Erin O'Hara, *Outsourcing, Modularity and the Theory of the Firm* (Oct. 30, 2008) (unpublished manuscript, on file with the *Columbia Law Review*), available at <http://law.vanderbilt.edu/academics/scholarly-events/symposia—conferences/supplychaingovernanceconf/download.aspx?id=3267>; George S. Geis, *The Space Between Markets and Hierarchies*, 95 Va. L. Rev. (forthcoming Mar. 2009), available at <http://ssrn.com/abstract=1272697> (on file with the *Columbia Law Review*) (discussing growth of business outsourcing and reasons that firms enter into outsourcing agreements). For a useful essay in the spirit of this one see Matthew C. Jennejohn, *Collaboration, Innovation, and Contract Design* (Columbia Law & Econ. Working Paper No. 319, 2008), available at http://works.bepress.com/cgi/viewcontent.cgi?article=1000&context=matthew_jennejohn (on file with the *Columbia Law Review*). We also appreciate the support of the Kauffman Foundation.

ing forms of contracting beyond the reach of contract theory models. In this Article, we connect the emerging contract practice to theory, learning from what has happened in the real world to frame a theoretical explanation of this cross-organizational innovation and to reconceptualize the boundaries of the firm accordingly. We argue that the vertical disintegration of the supply chain in many industries is mediated neither by fully specified technical interfaces that allow suppliers to produce a modular piece of the ultimate product, nor by entirely implicit relational contracts supported only by norms of reciprocity and the expectation of future dealings. Rather, we suggest that the change in the boundary of the firm has given rise to a new form of contracting between firms—what we call “contracting for innovation.” This pattern braids explicit and implicit contracting to support iterative collaborative innovation by raising switching costs. These costs, represented by the parties’ parallel transaction-specific investments in knowledge about their collaborators’ capacities, deter opportunism under circumstances where explicit contracting, renegotiation, and the anticipation of future dealings cannot.

| | |
|--|-----|
| INTRODUCTION | 433 |
| I. VERTICAL DISINTEGRATION AND COLLABORATION AMONG FIRMS: INDUSTRIAL ORGANIZATION IN AN UNCERTAIN WORLD | 436 |
| A. Vertical Integration and the Chandlerian Firm | 437 |
| B. The Vertical Disintegration of the Chandlerian Firm: The Shift from Risk to Uncertainty | 438 |
| C. Alternative Characterizations of the Emergent Institutional Framework: Modular, Relational, and Iterative Collaboration | 444 |
| 1. Modular Collaboration | 444 |
| 2. Relational Collaboration | 445 |
| 3. Iterative Collaboration | 446 |
| II. THE CONTRACTING RESPONSE: COPING WITH CONTINUOUS UNCERTAINTY | 448 |
| A. Elements of Contractual Governance Under Continuous Uncertainty | 450 |
| B. The Technology of Contracts: The Problem of Incompleteness | 451 |
| C. Contracting Under Continuous Uncertainty: The Limits of Contract Theory | 455 |
| III. THREE CONTRACT EXEMPLARS | 458 |
| A. The Deere–Stanadyne Agreement | 459 |
| B. The Apple–SCI Agreement | 463 |
| C. The Warner–Lambert–Ligand Agreement | 467 |
| IV. A THEORY OF CONTRACTUAL COLLABORATION AND CO- DESIGN | 472 |
| A. Transactional Variations in Collaborative Contract Design | 473 |

| | | |
|----|---|-----|
| B. | Coordination Cascades and the New Governance Mechanisms | 476 |
| 1. | The Function of the Collaborative/Learning Phase | 476 |
| 2. | Mechanisms for Deterring Ongoing Opportunism | 481 |
| 3. | Deterring Opportunism at the End Stage: Dividing the Pie with Options | 489 |
| C. | Risks and Rewards of Braiding Explicit and Implicit Contracts | 492 |
| V. | COMING FULL CIRCLE: CONTRACTING FOR INNOVATION AND THE THEORY OF THE FIRM..... | 494 |
| A. | From One-to-One to Many-to-Many: Current Understandings of the Boundaries of the Firm | 494 |
| B. | Contracting for Innovation and the Boundaries of the Firm..... | 499 |
| | CONCLUSION | 502 |

INTRODUCTION

Rapidly innovating industries are not behaving the way theory expects. Conventional industrial organization theory predicts that when parties in a supply chain have to make transaction-specific investments, the heightened risk of opportunistic behavior by their counterparties will drive them away from contractual relationships and toward vertical integration.¹ The pressure toward vertical integration will be especially powerful in rapidly innovating industries where swift technological change produces uncertainty in supply-chain relationships—that is, where the future states of those relationships cannot be predicted probabilistically.²

1. Much of this transaction-cost literature is an extension of the work of Oliver Williamson. See generally Oliver E. Williamson, *The Economic Institutions of Capitalism* (1985) [hereinafter Williamson, *Economic Institutions*] (examining economic organization through lens of transaction-cost economics); Oliver E. Williamson, *Markets and Hierarchies: Analysis and Antitrust Implications* (1975) [hereinafter Williamson, *Markets and Hierarchies*] (discussing organization of economic activity within and between markets and hierarchies); Oliver E. Williamson, *Transaction-Cost Economics: The Governance of Contractual Relations*, 22 *J.L. & Econ.* 233 (1979) [hereinafter Williamson, *Transaction-Cost Economics*] (describing main governance structures of transactions and how they are matched with institutions).

2. The reference here is to uncertainty in Knightian terms. See generally Frank H. Knight, *Risk, Uncertainty and Profit* (1921). In Knight's usage, uncertainty is distinguished from risk. Risk exists when alternative future states of the world occur with quantifiable probability: The future can be expressed as a probability distribution. Uncertainty exists when alternative future states of the world do not occur with quantifiable probability. Since we can mitigate or even eliminate (through insurance and other means) the effects of unfavorable states that occur with known probability, a risky world is one in which we can, with near certainty, live in the conditions we choose. *Id.* at 197–232. The same is not true of an uncertain world. The Knightian distinction between risk and uncertainty is a useful way to illustrate the manner in which accelerating

In the presence of uncertainty, contemporary contract theory offers no general solution to the problem of assuring both efficient levels of transaction-specific investment *ex ante* and adjustment to an efficient outcome *ex post*, after uncertainty is resolved.³ When contract cannot address opportunism successfully, firms should dominate markets as a means to organize supply relationships.

Despite conventional industrial organization theory, however, contemporary practice is moving away from vertical integration. Producers today recognize that they cannot themselves maintain cutting-edge technology in every field required for the success of their products. Accordingly, companies are increasingly electing to acquire by contract components that in the past they would have made themselves. Put otherwise, instead of vertical integration, we observe vertical *disintegration* in a significant number of industries. Moreover, in the process of vertically disintegrating, firms are developing forms of contracting beyond the reach of existing contract theory models.

To explain the incursions of contract on the domain of the vertically integrated firm, some theorists have emphasized “modularity.” Their claim is that new production tools permit parties to more or less standardize the interface between separate stages (or modules) of production.⁴ Each module can serve many purposes and therefore fit a variety of different products. The result is a moderation in the intensity of firm-specific investments and corresponding reductions in the risk of opportunism and thus in the need for vertical integration.

But the modularity hypothesis does not fit many of the new supply-chain relationships. Despite its apparent benefits, modularity is a double-edged sword: It may trap a firm in a no-longer-competitive technology.⁵ To avoid the “modularity trap,” firms instead are engaging each other in a process of iterative collaboration and co-design of both the interface and the components it joins.

Other theorists have suggested, therefore, that the collaborative process that replaces vertical integration is governed by relational con-

technology and global competition have created unique circumstances that resist probabilistic classification.

3. See *infra* Part II.B. See generally Francine Lafontaine & Margaret Slade, Vertical Integration and Firm Boundaries: The Evidence, 45 J. Econ. Literature 629, 649 (2007) (noting that hold-up risk “clearly pose[s] problems for long-term contracting, and those problems are exacerbated in volatile environments”).

4. See Richard N. Langlois, The Vanishing Hand: The Changing Dynamics of Industrial Capitalism, 12 Indus. & Corp. Change 351, 370–76 (2003) [hereinafter Langlois, Vanishing Hand].

5. Henry Chesbrough, Towards a Dynamics of Modularity: A Cyclical Model of Technical Advance, in *The Business of Systems Integration* 174, 181 (Andrea Prencipe et al. eds., 2003).

tracting rather than modularity.⁶ Here, the focus is on the dominance of noncontractual social and network bonds and informal cooperation as the mechanisms that support collaboration by constraining opportunism. Pure relational contracting, however, does not explain the complex contracting forms that disintegrated firms have devised to cope with the continuing uncertainty caused by rapid technological change.

What we see instead is a rich braiding of explicit (i.e., legally enforceable) obligations and implicit (i.e., legally unenforceable) obligations. The explicit and implicit obligations interact within a formal governance structure that regulates the exchange of highly revealing information but does not necessarily impose legally enforceable obligations to buy or sell anything. This braiding creates an interactive process that constrains opportunism as the parties' investments in detailed knowledge of each other's character and capabilities raise switching costs—the costs one party to a contract must incur in order to replace the other party to the contract.⁷

In this Article, we connect this emerging contract practice to theory, learning from what has happened in the real world to frame a theoretical explanation of this cross-organizational innovation and to reconceptualize the boundaries of the firm accordingly. We argue that the vertical disintegration of the supply chain observed in many industries is mediated neither by fully specified explicit contracts that allow suppliers to produce a modular component of the ultimate product nor by entirely implicit relational contracts supported only by norms of reciprocity and the expectation of future dealings. Rather, we suggest that the changes in firm boundaries are mediated by a new form of contracting—what we call “contracting for innovation.”⁸ Contracting for innovation supports iterative collaboration between firms by interweaving explicit and implicit terms that respond to the uncertainty inherent in the innovation process: The inability of the parties to specify *ex ante* the nature of the product to be produced or its performance characteristics means that the terms of performance will be determined by the very governance process the contract creates.

6. See Naomi R. Lamoreaux et al., *Beyond Markets and Hierarchies: Toward a New Synthesis of American Business History*, 108 *Am. Hist. Rev.* 404, 408–09 (2003) [hereinafter Lamoreaux et al., *Beyond Markets*] (outlining relational-contracting model).

7. See *infra* Part IV.B.2.a.

8. Lafontaine and Slade provide a useful survey of the theory and empirical studies of the vertical integration decision and the consequent boundary of the firm. See Lafontaine & Slade, *supra* note 3. However, the authors limit their attention to polar definitions of contract and vertical integration, explicitly excluding consideration of the intermediate collaborative case that we argue is central to understanding the new transactional structure that supports collaborative innovation. See *id.* at 631 (“[We] do not question the definition of vertical integration and markets that is used in the empirical studies. In most cases, this implies that we equate contracts with arms length transactions and contrast firms’ decisions to rely on such transactions versus vertical integration.” (citations omitted)).

As with the conventional account of the forces pushing toward vertical integration, opportunism plays a central role in explaining the organization of disintegrated innovation in the supply chain. However, here the opportunism is of a character somewhat different than that discussed in the conventional vertical integration literature. Moreover, it is addressed in a radically different fashion—the process of collaboration itself erects a barrier to taking advantage of the other party's specific investments. More precisely, as the parties invest in developing information about their respective capabilities, the cost of switching partners, and therefore the constraint on opportunism, goes up in tandem.

This Article proceeds in five Parts. We begin our discussion in Part I with an account of the ongoing vertical disintegration of the supply chain in innovative industries and consider various theoretical accounts of the process before offering our own assessment. Part II moves to a review of contract theory and its inability to offer a general solution to the canonical contracting problem of ensuring both the efficient level of specific investment in the face of uncertainty and efficient production once the uncertainty is resolved. Part III then describes three real world exemplars illustrating a continuum of contracts that support collaborative innovation. We use these transactional exemplars in Part IV to frame, but not to prove, our theoretical account of contracting for innovation. In Part V, we return to the theory of the firm and suggest that there is no theory of *the* firm. Following Bengt Holmström and John Roberts,⁹ we argue that the organizational boundaries of production, and the techniques that govern conduct within and across those boundaries, represent a variety of mechanisms that evolve in response to changes in the firm's real activities and the problems the firm must address. We conclude that future work investigating parties' efforts to contract for innovation requires both qualitative and quantitative data sufficient either to give confidence that our account captures current practice or to instruct us on what we have missed.

I. VERTICAL DISINTEGRATION AND COLLABORATION AMONG FIRMS: INDUSTRIAL ORGANIZATION IN AN UNCERTAIN WORLD

In the last two decades, the organization of firms in the United States and other advanced economies has changed in two fundamental ways. First, many transactions that used to take place within firms—for example, research and development, the manufacture of key components, and the assembly of final products—are now organized by agreements between firms. There is thus a decrease in the proportion of economic activity coordinated within firms and a corresponding increase in the pro-

9. Bengt Holmström & John Roberts, *The Boundaries of the Firm Revisited*, 12 J. Econ. Persp. 73 (1998).

portion of economic activity conducted through contract in the market.¹⁰ Second, market transactions between firms increasingly involve novel forms of collaboration—particularly rich and carefully organized exchanges of information designed to identify and utilize possibilities for innovation. This novel collaborative form constitutes contracting for innovation.¹¹

Collaborative innovation is not just a shift from hierarchy—the organization of transactions within firms—to contract. Rather, the unavoidable mutual vulnerabilities among collaborators motivate corresponding innovations in contractual governance to support the new transactional structure. In this part, we briefly canvass the most salient evidence for the shift from hierarchically organized transactions to interfirm collaborative innovation; speculate on the reasons for it; and present two contrary, but similarly incomplete, accounts of current economic organization to underscore the novelty of the simultaneously formal and informal mechanisms by which firms learn to innovate together.

A. Vertical Integration and the Chandlerian Firm

For much of the twentieth century, the dominant firms in industries such as steel, automobiles, electric machinery, and food processing, both in the United States and abroad, used the technologies of the second industrial revolution to achieve dramatic economies of scale through the mass production of standard goods with single-purpose or dedicated machinery.¹² The most conspicuous organizational feature of firms in these industries was vertical integration: The manufacturer of the final good was likely to own upstream producers of key inputs, or downstream distributors, or both.

Vertical integration was a response to the threat of production-process disruption. Because achieving economies of scale entailed large spe-

10. The existence of vertical disintegration is not controversial. Thus, while Langlois disagrees with Lamoreaux and colleagues about both the causes of, and the response to, vertical disintegration, both treat the phenomenon's existence as fact. Compare Richard N. Langlois, *Chandler in a Larger Frame: Markets, Transaction Costs, and Organizational Form in History*, 5 *Enterprise & Soc'y* 355, 355 (2004) [hereinafter Langlois, *Chandler in a Larger Frame*] ("In 1977, . . . the large, vertically integrated 'Chandlerian' corporation had dominated the organizational landscape for nearly a century. . . . A quarter century later, however, the Chandlerian firm no longer dominates the landscape."), with Naomi R. Lamoreaux et al., *Against Whig History*, 5 *Enterprise & Soc'y* 376, 376–77 (2004) [hereinafter Lamoreaux et al., *Against Whig History*] ("By the end of the twentieth century, it had become clear that . . . the acme of capitalist economic organization—the large, vertically integrated, horizontally diversified, managerially directed corporation—was clearly in retreat.").

11. The discussion that follows draws on Charles F. Sabel & Jonathan Zeitlin, *Neither Modularity nor Relational Contracting: Inter-Firm Collaboration in the New Economy*, 5 *Enterprise & Soc'y* 388, 388–89 (2004) (critiquing theories of Langlois, Lamoreaux, and others and presenting alternative interpretations of the post-Chandlerian economy).

12. Michael J. Piore & Charles F. Sabel, *The Second Industrial Divide: Possibilities for Prosperity* 19–20 (1984).

cific investments in production equipment that had little or no value unless used for the purpose to which it was dedicated, interruptions in the flow of production could be ruinous. In the familiar phrase of Alfred Chandler (the preeminent historian of the mass-production firm) the “visible hand” of managerial hierarchy supplanted market exchange as the organizing agent when potential costs of disruption grew.¹³

Economists, following Oliver Williamson, interpreted the vertical integration of the Chandlerian firm as a response to a particular class of potential disruption: the threat of hold-ups inherent in co-specialized or specific investments.¹⁴ When the values of two independently controlled investments are mutually dependent, each investor tries to induce the other to invest first in order to extract more favorable terms once an irrevocable commitment has been made.¹⁵ Placing both assets under the control of a single owner—vertical integration—unblocks this logjam. Much of the most interesting work in the theory of the firm since the 1980s explores the conditions under which parties to investments in interdependent assets can allocate initial contractual rights so that the one best able to maximize the joint value of the investment is in a position to bargain for exclusive control once it becomes clear what conditions actually prevail.¹⁶

B. *The Vertical Disintegration of the Chandlerian Firm: The Shift from Risk to Uncertainty*

Current developments in industrial organization make the historical dominance of the Chandlerian firm and the view that the resolution of the hold-up problem is decisive to the structure of the firm rather anachronistic. Instead, a current observer sees something radically different: the disintegration of vertical combinations in sectors where they once seemed irrevocably established and the exploration of collaborative alternatives to full integration in domains unencumbered by the legacy of vertical integration. Put otherwise, in certain economically significant sectors, fear of hold-ups, at least in their traditional form, no longer compels firms to vertically integrate.

For example, U.S. automobile makers, such as General Motors, whose acquisitions of suppliers in the 1920s were often invoked to illustrate the imperatives of vertical integration,¹⁷ have divested many of their

13. Alfred D. Chandler, Jr., *The Visible Hand: The Managerial Revolution in American Business* 1, 6–7 (1977).

14. See sources cited *supra* note 1 (introducing Williamson’s relevant work).

15. Williamson, *Economic Institutions*, *supra* note 1, at 85–102; Williamson, *Markets and Hierarchies*, *supra* note 1, at 106–16.

16. See *infra* note 65 and accompanying text.

17. See, e.g., Ronald Coase, *The Conduct of Economics: The Example of Fisher Body and General Motors*, 15 J. Econ. & Mgmt. Strategy 255, 255–63 (2006) [hereinafter Coase, *Conduct of Economics*] (tracing how GM-Fisher Body contract developed in economics literature into frequently cited example of vertical-integration response to hold-up problem); Benjamin Klein, *Fisher-General Motors and the Nature of the Firm*, 43 J.L. &

internal component makers and emulated more competitive Japanese firms whose success has depended on close, continuing collaboration with a wide range of sophisticated outside suppliers. Similarly, pioneers of the mainframe computer industry, such as IBM, that initially modeled themselves on vertically integrated industrial firms, have sold internal makers of key components and now routinely purchase from outsiders devices they long insisted on making themselves. Many of the more recently founded firms in the personal computer industry make none of the key components themselves and organize final assembly by agreement with specialized “contract manufacturers” (who also play a role in product design).¹⁸ A sprawling literature on the modularization of production and the globalization of supply chains investigates this decentralization and the organizational disintegration that accompanies it.¹⁹

The disintegration of production, moreover, is not limited to the manufacture of physical goods. The production of knowledge needed to define and realize new generations of products is also illustrated by the decreasing importance of the large, centralized research laboratory in industries, such as pharmaceuticals and telecommunications, where it was pioneered. Today, research is likely to be conducted by an ad hoc consortium of firms of very different sizes, often including publicly funded laboratories as well, all contributing highly specialized, complementary expertise.²⁰ It is now routine for a large pharmaceutical company with, say, expertise in the metabolic pathways that produce a particular pathology, to search for therapeutic compounds in concert with a small start-up company that has developed tools for identifying, among billions of pos-

Econ. 105, 105–07 (2000) (describing how “Fisher Body-GM case . . . illuminates the economic advantages of vertical integration”); Benjamin Klein, Vertical Integration as Organizational Ownership: The Fisher Body-General Motors Relationship Revisited, 4 J.L. Econ. & Org. 199, 200–11 (1988) [hereinafter Klein, Vertical Integration] (discussing how GM-Fisher Body contract exemplifies vertical-integration response to hold-up problem).

18. See Dave Nelson et al., *The Purchasing Machine: How the Top Ten Companies Use Best Practices to Manage Their Supply Chains*, at xi–xii (2001) (describing IBM’s shift to higher percentage of outside suppliers and its creation of “multi-company supply chain”); Joel M. Podolny & Karen L. Page, Network Forms of Organization, 24 Ann. Rev. Soc. 57, 68 (1998) (contrasting older, vertically integrated computer firms with younger disintegrated ones).

19. See generally Gary Gereffi et al., The Governance of Global Value Chains, 12 Rev. Int’l Pol. Econ. 78, 79–96 (2005) (noting trend toward value-chain disintegration and outlining five value-chain governance types, four contract-based and one hierarchical); Timothy J. Sturgeon, Modular Production Networks: A New American Model of Industrial Organization, 11 Indus. & Corp. Change 451, 451–64 (2002) (noting emerging forms of decentralized production and outlining “modular production network model” in explanation).

20. David A. Hounshell, The Evolution of Industrial Research in the United States, *in* *Engines of Innovation: U.S. Industrial Research at the End of an Era* 13, 51–56 (Richard S. Rosenbloom & William J. Spencer eds., 1996); Mark B. Myers & Richard S. Rosenbloom, Rethinking the Role of Industrial Research, *in* *Engines of Innovation: U.S. Industrial Research at the End of an Era*, supra, at 209, 210–11.

sibilities, the classes of molecules most likely to correct the metabolic defect without producing toxic side effects.²¹

At the intersection of both these developments—disintegration of physical production into chains of component suppliers and the collaborative networking of research—is the “platform” organization of production. Consider a computer operating system, a current-model cell phone, or an airliner like the Boeing 787. In each case, the performance of the product depends on the performance of a series of independently produced and rapidly developing subsystems²²—microprocessor, web browser, media players, and other applications in the case of an operating system; digital signal processor, radios, and antennas for various frequencies in the case of a cell phone; wings, engines, and fuselage in the case of a plane. The performance of each of these subsystems depends correlatively on the performance of the others, as transmitted through the architecture—the platform—linking them all.

In each case, the platform owner—the operating system developer, the cell phone maker, or the airframe producer—knows that it could not possibly produce all or even most of the components or applications whose interplay creates the platform. In particular, the producer could not develop or sustain the capacity for cutting-edge innovation in all the necessary areas for the various components. As a result, collaboration with groups of key technology suppliers, involving continuing mutual adjustment and exchange of quintessentially proprietary knowledge, becomes the norm.²³ The recent emergence of a literature on platform

21. See Walter W. Powell et al., *Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology*, 41 *Admin. Sci. Q.* 116, 122–24 (1996) (noting that, because “full range of required skills” for successful biotech product development “could not be easily assembled under one roof[,] . . . the various participants in biotech have turned to . . . various kinds of partnerships to make up for their lack of internal capabilities and resources”); Weijan Shan et al., *Interfirm Cooperation and Startup Innovation in the Biotechnology Industry*, 15 *Strategic Mgmt. J.* 387, 387–88 (1994) (noting frequency and importance to successful biotech product development of collaboration between large, established pharmaceutical companies and small biotech start-ups).

22. See Annabelle Gawer & Michael A. Cusumano, *Platform Leadership: How Intel, Microsoft and Cisco Drive Industry Leadership* 2–3 (2002) (describing “the character of the modern platform” as “an evolving system made of interdependent pieces that can each be innovated upon” (emphasis omitted)). See generally Yves Doz & Mikko Kosonen, *Fast Strategy: How Strategic Agility Will Help You Stay Ahead of the Game* 3–15 (2008) (tying success of mobile-telephony firm Nokia to “strategic agility”: the ability to rapidly adapt business strategy to shifting market and competitive conditions).

23. See Joseph Farrell & Philip J. Weiser, *Modularity, Vertical Integration, and Open Access Policies: Towards a Convergence of Antitrust and Regulation in the Internet Age*, 17 *Harv. J.L. & Tech.* 85, 90–96 (2003) (describing evolution toward open-access modular platform technology in three industries and noting that modularity facilitates innovation by organizing independent, complementary specialized producers); cf. David S. Evans, *The Antitrust Economics of Multi-Sided Platform Markets*, 20 *Yale J. on Reg.* 325, 331–36 (2003) (describing economics and conditions of platform markets, which productively coordinate needs and abilities of complementary participants).

industries and the forms of predation that tempt platform owners²⁴ attests to the diffusion of enduring (though sometimes fraught) cooperation across firm boundaries just as the literature on modules and the global value chain attests to the decline of vertical integration.

The replacement across a wide range of industries of vertically integrated firms by interfirm collaboration poses a critical question: What accounts for the decline of vertical integration and the nature of the collaboration that is displacing it?

Just as stability is the precondition of mass production and vertical integration, instability and the volatility of markets have been their bane. The connection between stability, mass production, and vertical integration was suggested above: The high fixed costs associated with specific investments can only be amortized over long production runs.²⁵ The larger the expected demand, the greater the volume of investment that can be financed and the larger the economies of scale that can be achieved. External shocks to markets, or systemic sources of instability, conversely, deter investments in the tightly linked, dedicated equipment that makes mass production possible even as it creates the potential for hold-ups and so induces vertical integration. As Adam Smith, thinking of the economies of scale attained in the pin factories of his day, put it: “[T]he division of labour is limited by the extent of the market.”²⁶

Future generations of economic historians will no doubt clarify the circumstances that first encouraged the progression of the vertically integrated mass-production firm and that then cut the ground from under it. Nevertheless, for present purposes, one set of developments is especially relevant. Beginning roughly in the 1980s and continuing today, the profusion of new technological possibilities associated with what is loosely called “the information revolution” operated to intensify the systemic uncertainty faced by producers. Innovations cascaded, often leading to improvement cycles that became self-perpetuating and ultimately transformative in the possibilities for new applications they afforded. The computer itself is a prime example: Increases in computational power

24. See generally David S. Evans & Richard Schmalensee, Some Economic Aspects of Antitrust Analysis in Dynamically Competitive Industries, in 2 *Innovation Policy and the Economy* 1 (Adam B. Jaffe et al. eds., 2002) (describing changing economics of antitrust in industries with recent technological change and heavy investments in intellectual property); David S. Evans et al., A Survey of the Economic Role of Software Platforms in Computer-Based Industries (CESifo, Working Paper No. 1314, 2004), available at <http://ssrn.com/abstract=618982> (on file with the *Columbia Law Review*) (analyzing case studies of computer system-based industries and examining organization and evolution of these industries); Jean-Charles Rochet & Jean Tirole, Two-Sided Markets: An Overview (Mar. 12, 2004) (unpublished manuscript, on file with the *Columbia Law Review*), available at http://www.frbatlanta.org/filelegacydocs/ep_rochetover.pdf (offering an “introduction and a road map to the burgeoning literature on two-sided markets”).

25. See *supra* notes 12–13 and accompanying text.

26. Adam Smith, *The Wealth of Nations* 19 (Edwin Cannan ed., Random House, Inc. 2000) (1776).

led to improved tools for the design of microprocessors, more sophisticated materials, and more exacting manufacturing techniques for realizing the new designs. These improvements then led to further increases in the power of computers, and the cycle replayed.

This increasing unpredictability is manifest as the pervasive fear of what Clayton Christensen calls “disruptive” technologies.²⁷ A disruptive technology is a superior alternative to the currently dominant know-how in a particular domain that devalues the skills of incumbent industry leaders. But because the disruptive technology reflects a starkly different approach rather than a linear improvement of the dominant method, the best producers and most sophisticated consumers of the dominant method initially are blind to the disruptive technology’s potential and to the threat it represents. Precisely because their experience teaches them how to improve on what they already know and how to provide what their similarly focused customers believe they need, dominant producers do not see a threat coming from an entirely different direction. Disruptive technologies therefore typically get footholds in secondary or peripheral markets of no interest to the dominant players or their customers. They are then generalized to core domains of application, dislodging the incumbent producers.²⁸ In the heyday of vertical integration, incumbency was the goal, allowing firms to see over the horizon of technical development and providing, through economies of scale, the means to realize the possibilities they saw. Now incumbency is seen as a burden, proficiency with current technologies obstructing the view of future directions. Precisely the organizational capabilities that underlie success in the industry as currently understood blind the firm to threats from outside of the dominant conception.²⁹

If increasing uncertainty is inimical to integrated forms of industrial organization based on specific investments, we should find disintegrated forms in uncertain environments. We do. Two such responses to uncertainty—industrial districts (or “clusters”) and systems engineering—are especially interesting here. Each provides guidance in understanding the form of interfirm cooperation that has developed in the post-Chandlerian economy and that is our focus in this Article.

27. Clayton M. Christensen, *The Innovator’s Dilemma*, at xviii–xix (1st Collins Bus. Essentials ed., HarperCollins Publishers 2006) (1997) (introducing and explaining concept of “disruptive technologies”).

28. Examples of this phenomenon include electric-arc or mini-mill steel producers, hydraulically activated earth-moving equipment, and—in the realm of general production technologies—Japanese or lean production methods. See *id.* at xxi, 172–76. Christensen argues—unchallenged, so far as we know—that all established technologies are in principle disruptable in this way.

29. See *id.* at 3–4, 15–26. The replacement of the centralized research laboratory, where stable project groups could pursue a line of research for a decade or more, by ad hoc research consortia that connect expertise from disjoint domains reflects this transformation.

First, consider clusters. Clusters are geographically compact agglomerations of small- and medium-sized firms in industries characterized by volatile or rapidly shifting demand, each firm specializing in a particular phase of production or in a particular production process. Finished goods are produced by groups of firms collaborating in rapidly shifting constellations.³⁰ By recombining and thereby augmenting fragmented, specialized, and mostly tacit knowledge, a multiplicity of cooperative firms in a cluster adapts rapidly to changes in the economic environment.

Agglomerations of this kind played an important role in the industrialization of parts of Europe and the United States from the late eighteenth century onward.³¹ Variants are common in more recent industrializers ranging from Japan to Taiwan to Brazil to Kenya, and in the development of Silicon Valley.³² Since the turbulence in the markets for mass-produced goods in the mid-1980s made valuable the ease with which clustered firms could recombine as conditions changed, clusters are a microcosm of the “new” economy, able to prosper in much more volatile conditions than the vertically integrated large corporation.

Second, consider systems engineering. Systems engineering also facilitates cooperation, but, in contrast to free-form combinations of clusters, it supports formalized cooperation among very large firms to produce complex products in very uncertain technological environments. Systems engineering emerged in the United States after World War II to develop weapons systems that were then at the frontier of technological capability.³³ Since no single firm could produce, say, both the inertial guidance system and the rocket motor needed for a missile, coordination

30. Up to some limit, the more firms in a cluster, the easier it is for each firm to find the partners it needs and the lower its costs of production. Up to the size limit, therefore, firms in a cluster constitute positive externalities for each other. The attraction of these positive externalities is (part of) what draws firms to the cluster in the first place, causing agglomeration. See Edgar M. Hoover & Raymond Vernon, *Anatomy of a Metropolis: The Changing Distribution of People and Jobs Within the New York Metropolitan Region* 49–55 (1959); Paul Krugman, *Increasing Returns and Economic Geography*, 99 *J. Pol. Econ.* 483, 484–85 (1991).

31. See Charles F. Sabel & Jonathan Zeitlin, *World of Possibilities: Flexibility and Mass Production in Western Industrialization* 47–48, 155–56 (1997) (discussing French silk industry and German cutlery industry).

32. On Japan, see David Friedman, *The Misunderstood Miracle: Industrial Development and Political Change in Japan 177–200* (1988). On Silicon Valley, see AnnaLee Saxenian, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*, at 29–58 (1994); Ronald J. Gilson, *The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete*, 74 *N.Y.U. L. Rev.* 575 (1999). On Taiwan, see AnnaLee Saxenian, *The New Argonauts: Regional Advantage in a Global Economy* 122–96 (2006). On Brazil and Latin America, see Elisa Giuliani et al., *Upgrading in Global Value Chains: Lessons from Latin American Clusters*, 33 *World Dev.* 549 (2005). And on Italy, see Roberta Rabellotti et al., *Italian Industrial Districts on the Move: Where Are They Going?*, 17 *Eur. Plan. Stud.* 19 (2009).

33. These weapon systems included the Polaris submarine-launched intercontinental ballistic missile, its land-based counterpart, the Minuteman, and the DEW line early-warning radar system. See generally Harvey Sapolsky, *The Polaris System Development:*

of specialist “subsystem” suppliers was necessary. Responsibility for the elaboration of the initial design and its refinement in collaboration with the specialist suppliers was entrusted to a prime contractor. Related methods of systems integration developed in response to current conditions are common today at the (vastly extended) technical frontier.³⁴

*C. Alternative Characterizations of the Emergent Institutional Framework:
Modular, Relational, and Iterative Collaboration*

The preceding examples—clusters and systems engineering—buttress the claim that disintegrated industrial organization is a creature of context—and in particular a response to uncertain environments. However, they suggest quite different and inconsistent interpretations of how disintegrated firms cooperate. Cooperation in clusters is extremely fluid; the creation of new firms, the re-contracting among existing ones, and the circulation of skilled workers from firm to firm is continuous. Organization seems highly informal, indeed nearly spontaneous. Mutual trust, born of long and close observation of actual behavior, underpins complex transactions memorialized in a handshake. In contrast, the organization of cooperation in systems engineering is highly formalized. Design parameters are specified in great detail and translated into precise formal contractual obligations with the intent of rendering the transaction as explicit as possible. In systems engineering, informality is treated as an unruly threat to exchange, not its foundation. Each of these examples has helped inspire a current but quite distinct interpretation of the disintegrated industrial organization emerging today.

1. *Modular Collaboration.* — One interpretation, advanced in the work of Richard Langlois, elaborates and generalizes the experience of systems engineering.³⁵ Its central claim is that the availability of new tools of design and production allows the development of technical standards or design rules that standardize the interfaces between organizationally separate stages of production. This standardization of interfaces is thought to so reduce the volume of information required for interfirm coordination that products can be decomposed into distinct modules, each of which can be produced in virtual isolation from the others.³⁶ Each producer need know only the interface that connects its contribution to the product; it need know little about the other components or interfaces. At the extreme, one can think of a “Lego-like” manufacturing process, with different firms producing differently shaped, sized, and colored pieces, all linked by a common form of connection. In Langlois’s

Bureaucratic and Programmatic Success in Government (1972); Harvey Sapolsky, *Inventing Systems Integration*, in *The Business of Systems Integration*, *supra* note 5, at 15.

34. See, e.g., Michael H. Best, *The Geography of Systems Integration*, in *The Business of Systems Integration*, *supra* note 5, at 201, 209–10 (describing relation between Intel, the microprocessor manufacturer, and suppliers of its chip-making equipment).

35. See Langlois, *Vanishing Hand*, *supra* note 4.

36. *Id.* at 374.

view, the formulation of standards and modules, periodically refined, now allows firms to achieve economies of scale and scope through the market rather than through the Chandlerian firm.³⁷ Langlois focuses, however, only on the opportunity for market substitutes to vertical integration; he does not address the form those arrangements actually take.

2. *Relational Collaboration.* — A contrary interpretation, advanced by Naomi Lamoreaux, Daniel Raff, and Peter Temin is, in turn, a variant of the stylized experience of industrial districts or clusters.³⁸ Lamoreaux, Raff, and Temin see the new economy as a shift away from coordination by managerial hierarchies in vertically integrated firms toward coordination through long-term relationships, based on “informal restraints on self-interested behavior,” among networks of formally separate firms.³⁹ Where Langlois emphasizes the resurgence of the market over the visible hand of management, they emphasize the resurgence of social bonds as the underpinnings of economic relations. In the pre-Chandlerian era, they write, “[B]usinesspeople in . . . industrial communities interacted socially as well as economically, and the resulting multidimensional relationships facilitated cooperation for purposes besides production.”⁴⁰ Relational coordination is possible again today, they argue, because cost reductions in transportation and communications allow multidimensional relationships to develop at a distance, all of which gives decentralized networks of firms the flexibility to respond to the increasingly differentiated demands of consumers. These flexible relationships are of particular value “where there is a great deal of uncertainty about the direction of technological change and both parties can benefit from the pooling of information and resources that trust makes possible.”⁴¹ Put another way, the chief advantage of formal disintegration of the firm is to create the possibility for sustained informal cooperation between independent pro-

37. See Langlois, Chandler in a Larger Frame, *supra* note 10, at 365–71 (finding that historical movement from market to hierarchical organization is actually bidirectional, that markets, when conditions allow for appropriate mediation by standards or modules, can be equal or superior to hierarchies). While Langlois acknowledges that many of the market substitutes for vertical integration cannot be reduced to “hard modularity,” he is explicit that he is “not really attempting to pronounce on which specific kinds of contractual arrangements constitute the New Economy. [His] claim is only that they are ‘market’ arrangements in the broad sense” Richard N. Langlois, Rejoinder 5 (July 2004) (unpublished manuscript, on file with the *Columbia Law Review*), available at <http://web.uconn.edu/ciom/Rejoinder.pdf>. As he puts it in Langlois, Vanishing Hand, *supra* note 4, at 376, “[a]s a central tendency, however, the buffering functions of management are devolving to the mechanisms of modularity and the market—informational decomposition, flexibility, and risk spreading.”

38. See Lamoreaux et al., Beyond Markets, *supra* note 6.

39. *Id.* at 430.

40. *Id.* at 417.

41. *Id.* at 409.

ducers, a different vector from the highly formalized exchanges stressed by Langlois.⁴²

We share Lamoreaux, Raff, and Temin's recognition that uncertainty has driven the disintegration process, and agree that relationships between firms substitute for vertical integration. However, again, the critical point from our perspective is to understand these relationships: How is collaborative innovation organized?⁴³

3. *Iterative Collaboration.* — Examining the limits of modularity, and how firms address those limits, calls attention to a type of cooperative institution—more formal than “relationships” but designed to facilitate learning among collaborating peers by means much less formal than hierarchical ordering.⁴⁴ This mid-range institutional form, neither purely relational nor based on a hierarchical specification of modular interfaces, is indispensable to cooperation among firms in platform settings, modern supply chains, or collaborative research and development. Given the relentless innovation of the modern economy, it is unsurprising that there are costs as well as benefits to fixing standards for technical interfaces for components and modules. In the short term, the cost is a sacrifice in performance of the product as a whole. Because the standard setter itself cannot observe the technological cutting edge across every field, the standards specified will be less ambitious than the outcome of a collaborative process among parties who are at the cutting edge. In a careful study of the hard disk drive industry, for example, Clayton Christensen found that only at the low-performance end of the market could the performance of finished products be completely predicted from the performance of their component modules.⁴⁵ Thus, modularization is not a stable strategy because it leaves open too many possibilities for competitive improvements

42. The informal foundations of market adjustment in general and innovation in particular are developed in an important strand of network sociology that begins with Mark S. Granovetter, *The Strength of Weak Ties*, 78 *Am. J. Soc.* 1360 (1973), and proceeds to Brian Uzzi, *Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness*, 42 *Admin. Sci. Q.* 35 (1997) and Ronald S. Burt, *The Network Structure of Social Capital*, 22 *Res. Organizational Behav.* 345 (2000).

43. Lamoreaux, Raff, and Temin describe the new transaction patterns that have arisen to substitute for vertical integration in the late twentieth century as “repeat interactions in which the parties involved made decisions about price and about the quality and quantity of output through a process of negotiation.” Lamoreaux et al., *Against Whig History*, *supra* note 10, at 384. However, this characterization does not advance the matter. That the resolution is through negotiations is either tautological—in the absence of a fully state-contingent contract, how else could the response to uncertainty be resolved—or simply undertheorized because current contract theory provides no guide as to how that renegotiation would take place. See *infra* Part II.C.

44. The following draws on Charles F. Sabel, *A Real-Time Revolution in Routines*, in *The Firm as a Collaborative Community* 106 (Charles Heckscher & Paul S. Adler eds., 2006) [hereinafter Sabel, *Real-Time Revolution*].

45. See Clayton M. Christensen et al., *Disruption, Disintegration and the Dissipation of Differentiability*, 11 *Indus. & Corp. Change* 955, 966–69 (2002).

that would cumulatively undermine the initial interfaces.⁴⁶ Therefore, in contrast to modular interfaces specified by the standard setter, an interesting set of firms engages in a process of iterative co-design, in which suppliers contribute to the redefinition of interface specifications for new products by building on their experience in manufacturing existing models. These disciplines of iterated co-design are neglected both by those who stress modularization and by those who find those connections in informal relationships; in the former they are ignored, and in the latter they are underspecified.

Iterated co-design establishes a first idea of what and how to produce through benchmarking: an exacting survey of current products and processes, augmented by assessments of promising new techniques. From this provisional starting point, each party responsible for a constituent component proposes modifications of the initial plan, having taken into account the implications of like proposals by the other subunits for its own activities. (This process—radically decentralized compared to the coordination provided by the prime contractor in systems engineering—is often called simultaneous engineering.) Provisional designs are then evaluated and refined. Once production begins, systems of error detection and correction focus on breakdowns in the new routines to trigger a search for weaknesses of the design or production process that escaped earlier examination. This root-cause analysis traces disruption back to its original source, presumed to be distant from the proximate cause of the breakdown. Participant firms must routinely question the suitability of their current routines and continuously readjust their approach in light of the contributions of their collaborators.⁴⁷

Taken together, these iterative, cooperative techniques play an important part in shaping the links that connect firms in the vertically disintegrated economy. As each collaborating party monitors and learns from the others' participation in the process, observation renders tacit knowledge at least partly explicit, easing long-range collaboration (by reducing the chances that the parties take incompatible things for granted) and reducing the chance that all the parties cling to limiting assumptions held by any single party.⁴⁸ Moreover, these methods also address the governance problems arising from the mutual vulnerability inherent in such

46. There are, moreover, long-term costs to modularity in that a commitment to particular interface standards can lock component and end-product manufacturers into obsolescent product architecture. This is what Chesbrough calls a "modularity trap." Chesbrough, *supra* note 5, at 181. In the firm, "the focus on developing products to compete within the standard eventually erodes the amount of system-level knowledge." *Id.*

47. For a fuller discussion on which this presentation draws, see generally Susan Helper et al., *Pragmatic Collaborations: Advancing Knowledge While Controlling Opportunism*, 9 *Indus. & Corp. Change* 443 (2000); Sabel, *Real-Time Revolution*, *supra* note 44.

48. On the high level of general skills needed by "supply-chain" integrators to coordinate product development and other functions across firm boundaries, see Geoffrey G. Parker & Edward G. Anderson, Jr., *From Buyer to Integrator: The Transformation of*

open-ended collaboration among different entities. The exchanges of information required for benchmarking, simultaneous engineering, and error detection and correction increase the mutual transparency of the actors to each other, revealing to each how rigorously and cooperatively the others scan for solutions in addressing joint problems of design or quality. By such monitoring, each party learns its counterparty's capabilities to operate in this transactional structure and to perform the substantive tasks required for the particular product. This framework establishes the position to which we now turn.

In the next Part, we address the contractual structure of collaborative innovation: how learning by monitoring is institutionalized in forms of contractual governance that allow the parties to rapidly establish confidence in each other's intentions and in their joint capacity to accomplish the tasks they set for themselves.

II. THE CONTRACTING RESPONSE: COPING WITH CONTINUOUS UNCERTAINTY

In Part I, we described the effects of continuing technological change on contemporary economic organization. In some markets and for some products, increases in the complexity of the technology and in the rate of change have made it difficult for a single firm to sustain state-of-the-art capacity across all the technologies necessary for successful product development. The response has been collaborative innovation across organizational boundaries with, for example, upstream and downstream participants in the supply chain specializing in particular technologies and the ultimate product resulting from cooperation among different organizations, each having contributed its special expertise. This ongoing process of vertical disintegration has stimulated the development of networks of explicit contracts among collaborating firms. In this Part, we argue that conventional contract theory cannot explain this process.

The emerging contractual networks are incompatible with the models of economic organization that purport to describe the modern production process. On the one hand, the explicit contracts that govern these new collaborative relationships do not fit the central tendency toward modularity described by Langlois, whose model has, as an end point, (relatively) complete contingent contracting in which parties specify the relevant interfaces. On the other hand, the new supply patterns are not regulated simply by noncontractual continuing relations and tit-for-tat enforcement—as suggested by the relational models of

the Supply-Chain Manager in the Vertically Disintegrating Firm, 11 *Production & Operations Mgmt.* 75, 80–88 (2002).

Lamoreaux, Raff, and Temin.⁴⁹ Rather, what we see emerging are organizational networks linked by explicit, formal contracts that rely on collaboration and co-design to stimulate continuous improvement in product development and engineering. Because the collaborative process is continuous, the parties operate in an ongoing state of uncertainty, one in which operational decisions must be continually updated and refined. This phenomenon of continuous uncertain change poses a unique challenge for contract design.

In this environment, we observe contracts in which parties create elaborate governance mechanisms in lieu of the more familiar risk-allocation provisions of conventional contracts.⁵⁰ The contracts can be arrayed along a continuum, ranging from contractual relationships that impose no formally enforceable obligations on the parties but that contemplate ongoing relationships of unlimited duration, to collaborative research agreements that look to the development of a particular product and a consequent end game. In each case, there is an iterated process of continuous collaboration and innovation that functionally substitutes for ex ante specification of the desired product. In each case, the parties make relation-specific investments in learning about their collaborator's capabilities, and these investments erect barriers to either party's taking advantage of their mutual dependencies. Thus, even in relationships that one might traditionally describe as stable, the spillover effects of continuous innovation create "coordination cascades"—innovation by one party requires coordination with a second party, whose response then requires adjustment by and further coordination with the first party.

None of the familiar mechanisms for coping with the problem of contractual incompleteness adequately responds to the challenge posed by structuring transactions in the face of continuous uncertainty. Nevertheless, while theory and conventional legal practice have lagged behind the conditions in the marketplace, transactional lawyers in a number of industries, apparently responding to their clients' need to structure new relationships in light of the constraints that uncertainty imposed, have begun creating the novel contracting patterns whose characteristics we now address.

49. While we are confident that they would include contracting for innovation within their general framework as set out *supra* Part I.C.2, the hard task is understanding the structures that we actually observe.

50. The distinction we draw between risk and uncertainty does not imply that the "conventional" contracts that we distinguish from those we study here are free from uncertainty in Knightian terms. For example, both "types" of contracts must account for moral hazard and other kinds of endogenous uncertainty. Rather, the distinction is between those contracts characterized by continuous uncertainty and those where relevant uncertainties such as product performance are resolved during the life of the contract.

A. Elements of Contractual Governance Under Continuous Uncertainty

The location of the innovative activity distinguishes the contracts of interest to us from more traditional relational contracting. In the new arrangements, innovation is the product of a joint effort by two or more organizations; it is metaphorically situated between them and is dependent on both. The development of the Boeing 787 aircraft is a good example. Innovation in the design and manufacture of the wing, the province of one supplier (or group of suppliers), is dependent on the design and manufacture of the fuselage, the province of a different supplier (or group of suppliers), and vice versa.⁵¹ Innovation in one structure must mesh with innovation in the other in order for either to be successful. The design of the wing must not only be compatible with the design of the fuselage on all relevant dimensions; the two must physically fit together. Innovation is thus a collaborative and iterative process rather than a discrete product supplied by a party upstream in the supply chain according to specifications set by a downstream customer.⁵²

Precisely how have parties to these new collaborative relationships structured their contracts? We set out to answer this question in two stages. We began our research for this Article with a small group of twelve contracts, each of which committed the parties to a collaborative process of design and production. From that initial group we selected three exemplars, described in detail in Part III, that reflect distinct patterns of collaborative production and supply. The Deere–Stanadyne contract addresses collaboration but without any product/sale obligation;⁵³ the Apple–SCI contract couples collaboration with production for a fixed period, while contemplating joint efforts for a longer term to which, however, neither party was obligated;⁵⁴ and the Warner–Lambert–Ligand contract covers the collaborative search for a product and the noncollaborative commercialization of it.⁵⁵

51. For a description of collaborative innovation in the production of commercial aircraft, see Alan O'Sullivan, *Why Tense, Unstable, and Diverse Relations Are Inherent in Co-designing with Suppliers: An Aerospace Case Study*, 15 *Indus. & Corp. Change* 221, 228–44 (2006).

52. See *id.* at 222 (describing “complex coupling” involved in airplane co-design, which “leads to iteration across the many design tasks and requires the designers to proceed in parallel”).

53. The Deere–Stanadyne contract obligates neither party to supply or purchase anything. Instead, it establishes only the terms of future purchases, should they occur, and includes Stanadyne in already established formal programs covering Deere's evaluation of a supplier's characteristics and performance. See *infra* text accompanying notes 76–82.

54. The Apple–SCI contract is a turnkey arrangement for production of a substantial fraction of Apple's personal computers. The arrangement necessarily involves collaborative, iterative innovation to coordinate changing technology and demand with changes in the manufacturing and assembly process but does not bind either party after the first three years. See *infra* text accompanying notes 90–93.

55. The Warner–Lambert–Ligand contract concerns a collaborative effort to discover and commercialize pharmaceutical products over a specified term and specifies the

As the foregoing suggests, the transactions governed by these contracts share a number of characteristics. First, the primary output is an innovative “product,”⁵⁶ one whose characteristics, costs, and manufacture, because of uncertainty, cannot be specified *ex ante*. Second, neither party alone has the capacity to specify and develop the product’s characteristics, costs, and methods of manufacture; hence, there must be collaboration among companies with different capabilities. Third, the process of specification and development will be iterative: Individual design elements will depend on the recurrent input from those working upstream or downstream and from those working on other design elements. Thus, central to these transactions are communication and cooperation across the two (or more) firms—the design, specification, and determination of manufacturing characteristics will be the result of repeated interactive collaborative efforts by employees of separate firms with distinct capabilities.

These commonalities highlight the conceptual questions that any explanatory theory must resolve. How do the parties deal with the problems of opportunism and the risk of hold-up that seem endemic in such interactive collaborative relationships? In particular, how do the parties constrain the temptation to exploit for private purposes information that is developed collaboratively? And how do the parties divide the eventual gains from the collaborative relationship when uncertainty precludes specifying the division *ex ante* and specific investment makes *ex post* allocation subject to hold-up? Is this temptation to use jointly produced information opportunistically and to hold up the counterparty when dividing gains adequately deterred by the elaborate set of formal and informal governance mechanisms that are a defining characteristic of these collaborative contracts?

B. *The Technology of Contracts: The Problem of Incompleteness*⁵⁷

To begin to answer these questions, recall first the principal reasons that transacting parties seek to write explicit contracts and the limits to such efforts. Explicit contracts can protect, and thereby encourage, spe-

options at the end of collaboration if the parties’ efforts result in a marketable product. See *infra* text accompanying notes 101–103.

56. We use the term “product” in this part to describe a range of innovative outputs: It might refer to a single product, but it also might refer to a stream of innovations (for example, the Deere–Stanadyne, Apple–SCI, and General Motors–Fisher Body contracts discussed, respectively, *infra* Parts III.A, III.B, and notes 120, 169–170 and accompanying text) or intellectual property that results in a single patent or patentable things (for example, the Warner–Lambert–Ligand contract discussed *infra* Part III.C).

57. This discussion draws on Robert E. Scott & George G. Triantis, *Anticipating Litigation in Contract Design*, 115 Yale L.J. 814 (2006), and Robert E. Scott & George G. Triantis, *Incomplete Contracts and the Theory of Contract Design*, 56 Case W. Res. L. Rev. 187 (2005).

cific investments,⁵⁸ which are often critical to transactions that contemplate more than a single simultaneous exchange. Yet contractual terms that encourage both parties to make efficient *ex ante* investments in the subject matter of the contract may undermine the *ex post* efficiency of the transaction if completion is compelled whenever one party still benefits, even when circumstances have so changed that the result is a net loss for the parties jointly.⁵⁹ Thus, the goal of efficient specific investment *ex ante* and of efficient *ex post* trade will often be at loggerheads when parties contract under uncertainty. The commitment necessary to motivate specific investments that maximize the contractual surplus will typically conflict with the flexibility needed to halt transactions (even when one party will still benefit) that have insufficient net value when uncertainty is resolved.

To see why, consider a benchmark solution to the dual objective of *ex ante* and *ex post* efficiency: a complete, legally enforceable, state-contingent contract. Such a contract specifies *ex ante* the parties' obligations in each possible *ex post* state of the world and is enforceable according to its terms, thereby assuring that performance occurs when, but only when, it is efficient. But while complete state-contingent contracts theoretically can address the tension between efficient *ex ante* investment and efficient *ex post* performance, the transaction costs of contracting frustrate this outcome. Of particular importance are the information barriers that prevent parties from controlling moral hazard when the future states of the world depend on their own actions. As a result, when the level of uncertainty is high, contracts will be incomplete because it simply costs too much (or may be impossible) for contracting parties to foresee and then describe appropriately the contractual outcomes for all (or even most) of the possible future states of the world that might materialize.⁶⁰

The information costs of contracting are incurred in two stages. *Ex ante* contracting costs are those of anticipating contingencies that may affect efficient performance and therefore efficient investment, and of writing a contract that specifies an outcome for each. *Ex post* enforcement costs are those of observing and proving any fact relevant to determining the actual state of the world (given that the parties have an incentive to misrepresent reality). It is costly to specify what should happen in different future states, and it is costly to prove what actually did happen.

58. Cf. *supra* notes 1–3, 13 and accompanying text (describing vertical integration as the more efficient means of mitigating hold-up risk of specific investment).

59. An *ex post* efficient contract should seek to ensure that exchange proceeds in all circumstances in which it produces value, but not otherwise. Trade is inefficient when the realized cost of performance to the promisor turns out to exceed the value of performance to the promisee.

60. See, e.g., Oliver Hart & John Moore, *Foundations of Incomplete Contracts*, 66 *Rev. Econ. Stud.* 115, 124–26 (1999); Oliver Hart & John Moore, *Incomplete Contracts and Renegotiation*, 56 *Econometrica* 755, 756 (1988); Oliver Hart, *Incomplete Contracts and the Theory of the Firm*, 4 *J.L. Econ. & Org.* 119, 123 (1988).

Both *ex ante* and *ex post* contracting costs prevent parties from writing complete state-contingent contracts.

Facing uncertainty and information costs, how should parties formalize their contracts? One option is to write an intentionally incomplete contract with precise, unchanging terms—i.e., determinate outcomes that apply across the board regardless of the eventual state of the world. For example, Buyer might contract with Seller at a fixed price for the manufacture of a precisely specified, customized machine, where Seller promises to deliver and Buyer promises to pay even if subsequent events increase Seller's costs or reduce Buyer's value. Such "hard" terms bind the parties to their respective commitments, which motivates each party as promisee to undertake relation-specific investments and encourages each party as promisor to take cost-effective steps to reduce anticipated risk-bearing costs.

But, because the hard terms of such an intentionally incomplete contract do not change based on what actually happens, they may be inefficient *ex post*, when the passage of time replaces uncertainty with fact. As suggested above, the actual cost to Seller of manufacturing the customized machine precisely as specified in the contract may exceed its value to Buyer.⁶¹ Both parties would prefer to design their contract *ex ante* so as to avoid the possibility of inefficient production *ex post*.

In theory, one solution to the inflexibility of hard terms is for the parties to renegotiate the contract once uncertainties are resolved. But if parties have made specific investments in the contract, later renegotiation raises the risk of a hold-up. Increased risk of hold-up, in turn, undermines the incentive to make those investments in the first place.⁶²

Alternatively, if information costs are high because neither the likelihood nor character of *ex post* change can be anticipated, the parties may emphasize *ex post* rather than *ex ante* efficiency in seeking to balance the two. In that case, they could draft a formal contract with vague standards, i.e., "soft" terms that invite subsequent adjustment to reflect what actually happened. Thus, for example, Seller might agree to adjust in good faith the specifications for the customized machine if the cost of providing the machine as originally specified later proved greater than its value to Buyer. By agreeing to "good faith adjustment," the parties seek to ensure that their contract is efficient both *ex ante* (by constraining *ex post* hold-up) and *ex post* (by providing for a mechanism that assures that the machine is produced if and only if it is efficient to do so).

But a contract that uses soft terms to address both *ex ante* uncertainty and the risk of *ex post* hold-up raises a moral hazard problem of its own through the actual operation of the soft terms. Here, moral hazard

61. Charles J. Goetz & Robert E. Scott, *Principles of Relational Contracts*, 67 Va. L. Rev. 1089, 1099 (1981).

62. See Oliver Hart & John Moore, *Property Rights and the Nature of the Firm*, 98 J. Pol. Econ. 1119, 1151 (1990) [hereinafter Hart & Moore, *Property Rights*]; Klein, *Vertical Integration*, *supra* note 17, at 202–04 & n.5.

results from a promisor with the discretion to adjust performance as conditions change always choosing the best alternative for himself rather than the “good faith” adjustment required by the soft terms, even though the self-interested choice is unlikely to be best for the promisee or to maximize the parties’ joint welfare.⁶³ Nor can the moral hazard problem necessarily be solved by delegating authority to determine the proper adjustment to a court. Soft terms such as “good-faith adjustment” remain as intractably ambiguous to judges as to the parties themselves, especially since the latter can act in bad faith in establishing the facts and in persuading the former what “good faith” should entail. Given, therefore, that a judge or other third party verifying contract performance under a broad standard of “good-faith adjustment” may mistakenly permit the promisor to substitute a lower-cost proxy for the agreed performance (say, by tendering an inferior machine), the promisor will be tempted to do so, even when this reduces joint welfare.

In short, neither hard nor soft contract terms can, standing alone, solve the problem of incomplete contracts. Under conditions of uncertainty, therefore, parties predictably seek to optimize total contracting costs by trading off the respective benefits and costs of commitment and flexibility. They can do this by shifting costs between the front and back end—the two stages—of the contracting process. As the preceding discussion illustrates, a core feature of contract design is the allocation of resources between drafting and enforcement. When the parties agree, for instance, to use their best efforts or to behave in a commercially reasonable manner, the subsequent adjudication of contractual disputes concerning their efforts or behavior requires a court to give precise meaning to those vague phrases. Thus, by using soft terms, parties delegate the specification of performance requirements to a court at the back end of the contracting process. The parties must bear the expected costs of litigation (including the costs of moral hazard in their conduct). But because a court has the benefit of some information unavailable to the parties at the time of formation, adjudication potentially allows them to benefit from more efficient performance standards than they could have specified *ex ante*. Alternatively, when the parties agree to precise (hard) terms, such as the obligation to supply a precisely specified, customized machine at a fixed price, they withdraw authority from courts to determine their particular performance obligations and instead direct enforcement of the obligations specified in advance. As noted above, this strategy requires the parties to fix performance obligations that rely on mere estimates of the likelihood of various future events rather than the actual occurrence of those events that is available to a court at a later date. The parties thus trade off the benefits of *ex ante* precision (with resulting *ex*

63. See Alan Schwartz & Robert E. Scott, *Contract Theory and the Limits of Contract Law*, 113 *Yale L.J.* 541, 601–05 (2003).

post inefficiency) against the hindsight advantage of the court in later litigation tempered by the moral hazard costs inherent in the process.

C. Contracting Under Continuous Uncertainty: The Limits of Contract Theory

The preceding discussion highlights the problem that contracting for collaborative innovation must confront. As discussed in Part II.A, the transactional structure must provide mechanisms for the sharing of information between the parties. In particular, the parties need credible information about each other's technical capacity, ability to manage a collaborative effort, capability for cooperative interaction, and especially each party's capacity to deal productively with disagreements that necessarily will arise when the characteristics of the desired innovation cannot be specified in advance. Moreover, this sharing of information is a continuous, collaborative process, one that requires asymmetric investments by each party as the collaboration proceeds along the critical path. Under these conditions, the contract design problem is particularly acute, because the collaborative process generates continuous uncertainty. As we have just seen, uncertainty creates problems for contract design, forcing parties to balance ex ante and ex post efficiencies. But, under conditions of continuous uncertainty, the problem is even more significant. When uncertainty is continuous, the parties cannot simply agree on the optimal trade off between ex ante and ex post informational advantages. There is no ex post period in which hindsight can be used to optimize a contractual relationship; the parties are continually cycling between different combinations of ex ante and ex post states. The crucial question thus becomes whether one's counterparty acts opportunistically—that is, takes advantage of the collaborative process to capture a larger share of the jointly created surplus (say, by using jointly produced information for its private benefit).⁶⁴ And the key challenge for transactional design is correspondingly to support the cooperative effort by constraining the strategic behavior made possible by ex ante specific investments in the collaborative project.

As discussed above, renegotiation of the contract ex post can, in theory, assure both ex ante and ex post efficiency in the face of uncertainty. Once the uncertainty is resolved, parties can in effect write a new contract specifying the decision the party in control should take—whether to per-

64. Cf. *Static Control Components v. Mitsubishi Kagaku Imaging Corp.*, No. 1:06CV00154, 2007 WL 586710, at *1 (M.D.N.C. Feb. 21, 2007) (outlining plaintiff's claim that defendant breached confidentiality provision by selling co-developed products to third parties and favored other competitors over plaintiff); *Eli Lilly & Co. v. Emisphere Techs., Inc.*, 408 F. Supp. 2d 668, 693 (S.D. Ill. 2006) (finding Emisphere entitled to terminate contract for collaborative research where Lilly created a secret research team that used the jointly produced information for purposes outside the collaboration); *Pharmacia & Upjohn Co. v. Elan Pharmaceuticals*, 781 N.Y.S.2d 95, 96–97 (App. Div. 2004) (describing breakdown in contractual relationship that occurred when defendant allegedly used information transferred pursuant to contract at issue to support “separate research program”).

form, alter the terms of performance, abandon the transaction, or make a side payment. This renegotiation can achieve ex post efficiency through Coasian bargaining: If a contract remains profitable to the promisor and yet is inefficient, the promisee will “bribe” the promisor not to perform.

As we have seen, however, renegotiation addresses only half the problem. It creates the flexibility to achieve ex post efficiency. But the prospect of renegotiation itself creates the possibility of hold-ups, which in turn undermines ex ante efficient investment. Contract and property-rights theorists have proposed solutions to the hold-up problem that rely on regulating renegotiation to constrain the eventual sharing of the surplus. For example, property rights theorists propose favoring the “efficient” owner, the one whose human capital is most complementary to the physical assets deployed in the project and whose propensity to invest in those assets is therefore most sensitive to the assurance of continuing control of them. Contract theorists, in turn, have proposed several alternative mechanisms to increase one party’s bargaining power in the future negotiations, for example, by allocating to that party the rights to control key decisions or property rights in assets specific to the exchange.⁶⁵ From these perspectives, a contract sets the field for future renegotiation of the terms of exchange after uncertainty has been resolved.

But efforts to constrain hold-ups by ex ante assignment of ex post decision rights fail for the contracts that concern us here. If there is no clear separation between the ex ante contract that supports transaction-specific investment and the resolution of uncertainty ex post, the identity of the party to whom decision or property rights should be allocated will continually shift, if it can be detected at all. Assignment of decision rights to this ephemeral owner will thus be meaningless. The discussion in Part I described a new pattern of collaborative innovation in the supply chain, one characterized by multiple information flows, iterative design, and mutual adaptation, all between separate firms. This network production responds to a technological and commercial environment where change is constant, adaptation must take place quickly and continuously, and the technology necessary to produce a cutting-edge product is not found in a single firm. The parties are not contracting over a temporary state whose resolution can be anticipated with enough precision to choose the efficient structure of post-resolution negotiation.⁶⁶ Rather, they are contracting over the creation of something whose features—and the comple-

65. See Sanford J. Grossman & Oliver Hart, *The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration*, 94 *J. Pol. Econ.* 691, 697–700 (1986); Hart & Moore, *Property Rights*, *supra* note 62, at 1121–24.

66. The renegotiation models that theorists have developed have the following key assumptions: Parties’ efforts (i.e., investment decisions) are noncontractible ex ante, but after uncertainty is resolved efforts are contractible ex post. That means that ex post there will be a renegotiation and (following Coase) the parties will allocate decision rights efficiently. The theorists’ answer to the ex post hold-up problem is to allocate decision rights and control ex ante so as to give the bargaining power in the ex post renegotiation to the investing party. See *supra* note 65 and accompanying text.

mentarities between those features and their own (changing) interests—emerge only through many iterations between them. When it is unclear at the time of the formation how large the contractual investments should be, and which party should make them and how gains should be shared, it is plainly impossible to mitigate the risk of ex post strategic behavior by regulating renegotiation in the familiar ways.⁶⁷

In response to these limitations, George Baker, Robert Gibbons, and Kevin Murphy have developed a model that realistically assumes that decision rights often are not contractible ex post, so that neither renegotiation to the efficient outcome nor the allocation of decision rights through options is possible. In this environment, the formal contract dictates a governance structure that motivates self-enforcing informal adjustments.⁶⁸ The optimal governance structure is achieved in this model by the ex ante contractual allocation—often in the form of an option—of ex post decision rights to the party who, because of informal constraints, has the least incentive to behave opportunistically. As one of us put it some time ago, “the goal is to shift the discretion to the party whose misuse of it can be most easily constrained,”⁶⁹ rather than to specify the appropriate adjustment. For example, in venture capital contracting, the decision whether to continue a project is shifted to the venture capitalist through staged financing because his decision, unlike that of the entrepreneur, is policed by an effective reputation market.⁷⁰

But while ex ante assignment of ex post decision rights via an option can address the governance problem where the option holder is constrained by informal mechanisms (as well as by contractually determined “prices”), an option approach has important limits. When the parties must adapt continuously, uncertainty about *which* party’s opportunism needs to be constrained, and a consequent inability to predict the decisions that actually will have to be made, imply that options are not a feasi-

67. The problem faced by parties to collaborative contracting is similar to the problem faced by parties to preliminary agreements: They also function in a complex environment in which a profitable project can take a number of forms and just what form will work, if any, is unknown at the start. In the preliminary agreement context, simultaneous investment by both parties makes a project sufficiently tangible to support a complete contract. But during the investment period, there is a perverse incentive to behave opportunistically by delaying a promised investment. Contemporary law can best solve this problem by characterizing that defection as a breach and protecting the promisee’s reliance expenditures. See Alan Schwartz & Robert E. Scott, Precontractual Liability and Preliminary Agreements, 120 Harv. L. Rev. 661, 685–91 (2007).

68. George Baker et al., Contracting for Control 2–7 (Mar. 21, 2006) (unpublished manuscript, on file with the *Columbia Law Review*), available at http://www.law.columbia.edu/null/Contracts+Conf++April+7-8,+2006++Paper++Gibbons?exclusive=filemgr.download&file_id=961193&showthumb=0.

69. Ronald J. Gilson, Engineering a Venture Capital Market: Lessons from the American Experience, 55 Stan. L. Rev. 1067, 1081 (2003) [hereinafter Gilson, Venture Capital].

70. Id. at 1086.

ble technique for assuring efficient adaptation.⁷¹ This setting, which describes the transactional environment in many of the new collaborative arrangements, requires instead a formal governance mechanism that stimulates the development of stable cooperative equilibria to support informal, relational contracting.

In Part IV, we show that these insights offer valuable tools for explaining the contractual patterns that we observe in a small number of co-design contracts. Before doing so, however, we turn in Part III to a more detailed description of three contracts for innovation that guide our analysis.

III. THREE CONTRACT EXEMPLARS

Our development of a theory of contractual collaboration and co-design in Part IV uses three real-life transactional exemplars.⁷² The contracts were chosen to illustrate a continuum of circumstances involving collaborative innovation across organizational boundaries. Despite variations owing to the particular transaction, our exemplars demonstrate surprising consistency of core features across industry settings. One (the Deere–Stanadyne contract) helps establish and maintain a long-term supply arrangement but does not obligate either party to supply or to purchase anything; the parties perform under the contract without any formally enforceable obligations. A second contract (the Apple–SCI contract) involves two clearly specified and legally enforceable obligations—to continue to manufacture and supply a product for a specified term (following a purchase of the manufacturing plant from the seller of the product)—and two unspecified and legally unenforceable obligations—to collaborate on continual improvements in the product and to supply it beyond the specified term. The third (the Warner–Lambert–Ligand contract) involves collaboration in the creation of a single class of pharma-

71. Baker, Gibbons, and Murphy solve part of this problem by assuming that there is no transaction-specific investment, see Baker et al., *supra* note 68, at 17, which eliminates the ex ante-ex post tension. As we will see, collaborative innovation does require specific investment, though of a different kind than usually assumed in the contract theory literature and, we think, in the circumstances that Baker, Gibbons, and Murphy actually have in mind. Our analysis here generalizes the Baker, Gibbons, and Murphy approach to encompass the type of transaction-specific investment inherent in collaborative innovation settings.

72. The contracts were obtained from one of two sources of contracts available on the Internet: onecle.com (<http://www.onecle.com>) and the Contracting and Organizations Research Institute (<http://cori.missouri.edu>). These organizations, in turn, obtain most of their contracts from SEC filings in the SEC's EDGAR database. Firms with a class of security registered under the Securities Exchange Act of 1934 are required to file with the Securities and Exchange Commission copies of their material contracts as exhibits to their periodic reports. These exhibits are accessible over the Internet through the SEC's EDGAR database. It is commonplace for certain provisions in a contract to be redacted prior to being added to the public database pursuant to a company's request for confidentiality.

ceutical products, which would then be guided through the process of regulatory approval and commercialized by only one of the parties.

We use the analysis of these three contracts to develop, in Part IV, a working theory (or extended hypothesis) to explain the contractual governance mechanisms we observe. In subsequent work, we will test our theoretical predictions against a larger group of contracts that support collaborative innovation. We stress that we use these three contracts principally as exemplars in developing our theory; we simply do not know yet the extent to which they generalize to other collaborative ventures. What we do know is that the key features of the contracts we highlight are not captured by the literature discussed above. If contract is substituting for organization in the vertical disintegration of the supply chain, current contract theory does not explain the resulting arrangement.

A final point of caution about generalizing from a very small number of contracts: We have no reason to believe that individual lawyers and clients negotiating and writing individual contracts will craft the efficient structure every time—even in competitive markets, efficiency constraints are simply not that binding.⁷³ Each contract will contain singularities; some may include mistakes. Mindful of this, we examine only central or exemplary features in the documents, and we connect these to suggest how, taken together, they respond to the contracting problem under discussion. Put differently, the aim is to stylize what appear to us to be the parties' intuitions and experience about what works, so that we can later assemble a larger group of contracts to assess whether we have correctly identified the mechanisms underpinning their success.

A. *The Deere–Stanadyne Agreement*

The Deere–Stanadyne agreement covers two very different functions: the supply and purchase of parts for Deere's existing products, and the collaborative process of developing new products as technology and the market for Deere's products evolve. The bulk of the formal contract concerns parts for Deere's current products. However, what is in many ways the commercially more important part of the arrangement concerns as-yet-unidentified future products and is established by indirection rather than by explicit provisions. This part of the agreement provides the key to understanding the structure of the relationship between the parties.

Deere manufactures and sells "machinery and equipment used in agriculture, construction, and commercial-residential lawn-garden care."⁷⁴ Stanadyne "owns and operates design and manufacturing facilities for

73. As Goldberg and Erickson observed in another context, "The contract terms meet only the weak test of birth, not the stronger test of survival." Victor P. Goldberg & John R. Erickson, *Quantity and Price Adjustment in Long-Term Contracts: A Case Study of Petroleum Coke*, 30 J.L. & Econ. 369, 371 (1987).

74. Deere & Co. & Stanadyne Corp., Long Term Agreement art. I(A) (Nov. 1, 2001) (on file with the *Columbia Law Review*) [hereinafter Deere–Stanadyne Agreement].

precision engine components including injection equipment.”⁷⁵ Most of the agreement addresses Stanadyne’s provision of parts listed in an appendix to the agreement (that was not available to us); other parts could be added by mutual written consent. Since Deere’s product line would change over the five years, it was inevitable that its parts requirements would change as well. The new parts would be co-developed by the parties, although the contract says nothing of that.

Interestingly, the bulk of the formal contract, while focusing on the provision of specified parts for Deere’s existing products, is not legally enforceable because it did not actually require the parties to do anything. Although the contract does refer to anticipated levels of Deere purchases, Stanadyne did not have to produce any parts, and, if it did produce them, Deere was under no obligation to take them (unless Deere issued a purchase order). The parties could easily have written a plainly enforceable supply contract. The parsimonious conclusion is that they chose to avoid legally enforceable commitments, not that they were creating an opportunity to convince a court to disregard the contract’s language. Conversely, there is no question that orders actually placed by Deere were enforceable at specified prices and subject to a sharing arrangement for cost savings achieved with respect to designated products.⁷⁶

This brings us to the matter of central concern. A substantial fraction of the parts specified in 2001 would be supplanted during the five-year life of the contract and thereafter.⁷⁷ The written agreement gives Stanadyne very little comfort with regard to new parts. Stanadyne can coordinate with Deere in devising the parts and controlling their costs. However, the agreement explicitly disavows any obligation on the part of Deere to develop parts with or purchase parts from Stanadyne. Deere’s only obligation is to negotiate in good faith, and that obligation is limited by a broad “meeting competition” clause that gives Deere virtually complete discretion.⁷⁸

The written agreement, therefore, does not commit either party with respect to the development of new products within the existing five-year term, or thereafter. Yet the success of the supply relationship and of both parties’ businesses depend on continued innovations in Deere’s products and therefore in the parts produced by Stanadyne. And precisely because the parties could not specify what innovations would be necessary or feasi-

75. *Id.* art. I(B).

76. Article VIII(D) provides, “If STANADYNE CORPORATION is in compliance with Section V [N.B. Section V obligates Stanadyne to participate in the Achieving Excellence program described *infra*], all SD Program reductions realized will be shared equally between DEERE and STANADYNE CORPORATION.” *Id.* art. VIII(D). “If DEERE determines that STANADYNE CORPORATION is not globally competitive and therefore not in compliance with Section V, however, STANADYNE CORPORATION agrees to pass 100% of cost reductions realized to DEERE until such time as they are in compliance.” *Id.*

77. The agreement was a follow-on that extended a previous five-year contract for an additional five years.

78. See *supra* note 76 (setting out text of relevant contract provision).

ble, or could be produced at a cost-effective price, something other than a state-contingent contract had to govern the parties' ongoing response to uncertainty. Understanding that, it becomes clear that the more significant element of the arrangement is a Deere program for supplier relationships—"Achieving Excellence"—that is identified but not explained in the contract.⁷⁹

The Achieving Excellence program, established in 1991, has three interrelated components: (a) measuring and monitoring performance; (b) providing and transmitting information about the character of the parties; and (c) combining (a) and (b) as the parties learn about both product development and each other over time. The program sets up a hierarchy of suppliers, ranging from "Conditional" at the bottom to "Partner" at the top. A Partner is defined as a "supplier who exceeds our performance standards, has reached world-class levels, and has a high impact on the satisfaction level of our customers."⁸⁰ If a supplier maintains Partner status five years in a row, it goes into the Deere Hall of Fame. In 2006, for example, twenty-one suppliers achieved Hall of Fame status.⁸¹ Suppliers are judged in five categories: quality, delivery, technical support, cost management, and wavelength. The first four are self-explanatory. The last, wavelength, purports to capture the supplier's ability to manage the human underpinnings of collaboration with Deere. It is a composite of initiative, attitude, responsiveness, attention to detail, communication, and safety performance.⁸² Under the program, the suppliers' performance is evaluated on a semi-annual basis by Deere evaluation teams whose composition—including representatives from all plants supplied and various corporate functions—reflects and balances Deere's different interests in its supply base, and thus helps protect both supplier and customer against partial judgments.

This arrangement suggests why the legal unenforceability of the agreement is of little importance to the parties. External verification of the facts surrounding a supplier's collaboration with Deere in developing new products through the legal process would be extremely difficult. Whether, for example, the supplier acted in good faith in refusing or failing to develop and produce a specific product is difficult for a court to determine, especially because, for reasons discussed above, both parties are prone to opportunistic behavior during judicial factfinding. In contrast, Deere's Achieving Excellence program is a governance mechanism

79. The supplier agrees only to participate (but given the absence of any consequences, the commitment is not enforceable): "STANADYNE CORPORATION will strive to meet or exceed all DEERE 'Achieving Excellence' (or 'AE') requirements to reach and maintain 'Partner' status." Deere-Stanadyne Agreement, *supra* note 74, art. V.

80. Deere & Co., Achieving Excellence Indirect Materials & Services 8, at https://jdsupply.deere.com/business_processes/ae_process/ae_indirect_presentation.pdf (last visited Jan. 19, 2009) (on file with the *Columbia Law Review*).

81. E.g., Centro, Inc., Awards, at <http://www.centroinc.com/awards.htm> (last visited Feb. 10, 2009) (on file with the *Columbia Law Review*).

82. Deere & Co., *supra* note 80, at 20.

that allows Deere to act based only on what its own evaluation teams find to be observable, rather than also having to verify those findings to a court. Moreover, the wavelength category puts suppliers on notice that Deere will take into account a supplier's character, in particular a supplier's attitudes toward cooperation.

Furthermore, because assessment of a supplier's capabilities and character is based on the actual experience of collaboration, it takes time, investing both parties in the relationship. As a supplier moves from Conditional supplier to Partner, both Deere and Stanadyne learn more about each other's capabilities and character. Replacing a particular supplier would require that Deere expend time and effort to learn both the capabilities and character of a new one. Thus, making an existing relationship work is often preferable to incurring the information costs associated with a replacement. And the same kind of considerations will constrain the supplier. Losing Deere as a customer eliminates the value of Stanadyne's investment in teaching Deere about its capabilities and character.

Finally, either party's termination of the relationship potentially results in a depreciation of the other party's reputation. Precisely because a party's actual performance is neither observable nor verifiable by other potential partners, a breakdown in the relationship imposes costs beyond the immediate loss of Deere's business or the supplier's product. For Stanadyne, termination by Deere would increase the investment a new customer would have to make in a relationship with Stanadyne, as the potential customer would want to understand why Deere had terminated the supplier. And this problem has a symmetric impact on Deere. Potential suppliers have to learn that Deere has the capabilities and character necessary to warrant the supplier's investment (i.e., that the Achieving Excellence program remains credible). Deere's failure to effectively manage a prior supply relationship raises questions that a new supplier would have to invest in answering.⁸³

Consistent with the parties' needs to demonstrate their trustworthiness and to allow each to learn how the other responds to disagreements (a characteristic critical to ongoing collaborative innovation), the agreement establishes a two-step dispute resolution process. First, executives from each firm higher up in the management than the disputing managers will meet in good faith in an attempt to negotiate a resolution.⁸⁴ If

83. See Victor P. Goldberg, A Relational Exchange Perspective on the Employment Relationship, *in* *Firms, Organization and Labour: Approaches to the Economics of Work Organization* 127, 129 (Frank H. Stephen ed., 1984) (discussing impact ex ante of imperfect information regarding one's contractual counterparty, specifically impact of associated information costs on one's willingness to enter into a bargain).

84. The contract provides, in pertinent part:

The parties shall attempt in good faith to resolve any controversy, claim or dispute of whatever nature arising out of or relating to this Agreement or the breach, termination, enforceability or validity thereof (a "DISPUTE") promptly by negotiation between executives or managers who have authority to settle the

that fails, the dispute will go to arbitrators using Illinois law and the Rules of Conciliation and Arbitration of the International Chamber of Commerce.

Thus, in the contracts between Deere and its suppliers, the nature of the parts to be produced and the technology for producing them changes over the course of the agreement in response to changes in the marketplace. Deere's products and the parts provided by suppliers co-evolve as the interplay between Deere and its suppliers produces new information. Deere and its suppliers rely upon an external, preexisting, informal governance mechanism over the course of their relationship to govern the innovative portion of the supply relationship, and an explicit, legally enforceable contract to govern the actual provision of specific parts.

B. *The Apple–SCI Agreement*

Compared to the Deere–Stanadyne agreement, the Apple–SCI agreement represents a movement toward legal enforceability. Because the Apple–SCI arrangement appears to cover two conceptually separate but related transactions—a one-time sale of a manufacturing facility and an ongoing commitment to collaborative innovation in connection with the manufacture and assembly of Apple computers—we see two different approaches to dealing with the challenge of dividing gains. The agreement has explicit, enforceable contract terms to support the one-time sale, but the commitment to collaborative innovation is protected by implicit, unenforceable terms.

In 1996, Apple sold its Fountain, Colorado manufacturing plant to SCI and simultaneously entered into a three-year contract to purchase a substantial share of its logic boards and personal computers from that plant. The transactions were one element in Apple's strategic decision to rely more on outsourcing. SCI was, at the time, the largest contract manufacturer, with sales of around four billion dollars, twenty plants in eight countries, over fifteen thousand employees, and over fifty customers, including Hewlett Packard and IBM.⁸⁵

The contract itself is fairly straightforward. For an initial three-year term, Apple promised to purchase at least a specified percentage of its logic boards and computers from SCI. Rather than committing to purchase a specific number of units, Apple's commitment was a function of its total purchases; that is, it agreed to buy a specified fraction of its main logic boards and computer systems in each of the three years. For the former, the commitment for the three years was 60/50/40; for the

DISPUTE, and who are at a higher level of management than the persons who have direct responsibility for the administration of this Agreement.

Deere–Stanadyne Agreement, *supra* note 74, art. XXII(G)(1).

85. Timothy J. Sturgeon, Turnkey Production: The Organizational Delinking of Production from Innovation, *in* *New Product Development and Production Networks* 67, 73–74 (Ulrich Jürges ed., 2000).

latter, it was 40/40/30. While the contract could be extended on a year-by-year basis, Apple made no purchase commitment after year three.

The sale was said to have been triggered by Apple's inability to meet a surge in demand from its own plants.⁸⁶ The terms of the contract seem to have given Apple significant flexibility.⁸⁷ The formulation of its purchase obligation in relative terms shifted the risk of demand fluctuation to SCI.⁸⁸ To some extent, this flexibility stands the traditional logic of industrial organization on its head, since we would normally expect to see a greater likelihood of vertical integration by ownership when the possibility of shortages would make the firm subject to the threat of hold-up. In the case of the contract manufacturing sector, however, the ability to respond quickly to demand fluctuations—to bear the risk of either over or under capacity—is central to the package that is offered to customers. Thus, to the extent that explicit legal remedies are insufficient to constrain an SCI hold-up in the first three years, and with respect to any supply beginning in year four, the risk of hold-up must be constrained by something other than explicit contract.⁸⁹

For at least the three-year initial term, the explicit portion of the Apple-SCI contract looks like a straightforward supply agreement, albeit with a formulation that shifts much of the risk of demand fluctuation to the contract manufacturer. This element of the agreement seems to protect SCI's purchase of the manufacturing plant from Apple: For the first three years, Apple will provide sufficient volume to make the plant acqui-

86. Timothy J. Sturgeon, *Turnkey Production Networks: A New American Model of Industrial Organization?* 9–10 & n.4 (Berkeley Roundtable on the Int'l Econ., Working Paper No. 92A, 1997), available at <http://brie.berkeley.edu/publications/wp%2092A.pdf> (on file with the *Columbia Law Review*).

87. Within the constraints of the percentage volume commitment, Apple had considerable flexibility to respond to market conditions. Apple would provide, on a monthly basis, nonbinding forecasts covering the following six months and issue monthly purchase orders on a rolling four-month basis. Apple could increase or decrease the quantity without penalty if it gave satisfactory advance notice. If Apple required greater flexibility, then it would be responsible for any overtime charges and vendor premiums. Apple could cancel any purchase order with thirty days notice provided that it reimburses SCI for costs reasonably incurred. Additionally, the percentage volume commitment itself provided flexibility to Apple to respond to market movements through a make-up clause. If Apple fell below its commitment in the first two years, it could either add the shortfall to its commitment for the third year or pay SCI the profit it would have earned on the shortfall.

88. Because Apple committed only to a specified share of its purchases, SCI bore the risk that reduced demand (for Apple computers or industry-wide) would result in an absolute decline in Apple's purchases from SCI, while increased demand would commit SCI to provide additional product, even though in both cases SCI's share of Apple's purchases would not change.

89. As discussed *supra*, Apple has multiple sources for the products it is purchasing from SCI; that is, Apple will be getting sixty percent of its computers from suppliers other than SCI. See *supra* text accompanying notes 85–86. Multi-sourcing provides some protection against SCI-specific opportunism, but when hold-up is made possible by industry-wide short supply there is no reason not to expect other suppliers to behave just as would SCI.

sition viable. For our purposes, however, the more interesting part of the contract is a second set of obligations concerning co-design of the actual products to be supplied.

Technological change in the computer industry is rapid, and, as noted, the contract is for products not yet known. In the agreement, Apple committed to purchase circuit boards and personal computers, but the specifications for those products are not set out in the contract. Thus, Apple and SCI had to collaborate in defining the performance characteristics of the products to be manufactured by SCI, and in setting the price to be paid.

Understanding that, the centerpiece of the contract was the parties' agreement to establish "product plans."⁹⁰ SCI promised to produce, and Apple promised to purchase, *something* over the three-year term, but the details would be determined collaboratively over the performance period. Apple would provide the first product plan, less than a month into the agreement. Since the first plan likely would largely reflect Apple's existing specifications, little collaboration was necessary. In contrast, subsequent plans, which would have to both anticipate and respond to technology changes, would be prepared collaboratively. In addition to providing a pricing formula for the products, the plans would specify preproduction services including development of assembly and test processes; development of test programs and/or fixtures; and production of prototype and/or validation units. The plan would also include a preproduction delivery and payment schedule. SCI would appoint a test engineer to work with Apple's test engineers and, if necessary, co-locate that engineer at Apple's facilities. SCI would regularly report to Apple and make its facilities available to Apple for inspection on reasonable notice. After a successful preproduction review, Apple would give SCI the go-ahead to begin production.⁹¹ SCI would then produce on a "turnkey" basis in accordance with Apple's specifications and quality requirements.⁹²

The process is not necessarily completed when Apple approves the specifications. Once production has begun, one or the other party might find a possible improvement. If SCI wants to change a component, material, or process it must obtain Apple's written consent. A consent request must include information on any cost, scheduling, and other impacts of the change, and Apple could require sample units. Adoption of the change would be solely at Apple's discretion. If Apple desired a modification in the design, it would have to submit an "engineering change order." Within a week, SCI would have to advise Apple on the cost or other

90. Apple Computer, Inc. & SCI Systems, Inc., Foundation Manufacturing Agreement art. 4 (May 31, 1996) (on file with the *Columbia Law Review*) [hereinafter Apple-SCI Agreement].

91. *Id.* art. 5.6.

92. *Id.* art. 6.

impacts of the change. Again, adoption of the change would be solely at Apple's discretion.

The details on pricing are in an exhibit not included with the contract, but the basic outlines are clear. For each new product, SCI proposes a price quote; the formula takes into account a number of cost factors, but the basic agreement does not say how these are to be weighted. On procured material, SCI passes through any cost reductions including rebates and discounts. Failure to do so constitutes a material breach and is grounds for immediate termination, except if the failures are *de minimis* or accidental and are promptly remedied, with interest. While such a response appears at first to be overkill, SCI's behavior on this verifiable fact could be a plausible proxy for SCI's commitment not to behave opportunistically.⁹³ Importantly, the pricing provisions can function as no more than a focal point for bargaining, since they are of significance only after an agreement is reached over the product plan. Thus, the collaboration that leads up to determining the innovations that will be incorporated in the project allows hard bargaining over pricing despite the contractual pricing formula.

Thus, at the core of the Apple–SCI agreement is a process of collaborative co-design, in which the parties iteratively determine the feasibility of innovations suggested by both parties. The three-year term, during which Apple commits to purchase from SCI in order to support the sale of the manufacturing plant, facilitates the development of valuable information. As the products actually being produced over the three-year period reflect less and less Apple's pre-contract products and more and more the product of the collaborative effort, joint development of cumulative changes in product plans generates ever-increasing knowledge about each party's capacity for collaborative innovation and for good faith dispute resolution. Then, at the end of the three-year commitment,

93. During the three-year term, the contract also contains two significant explicit remedies covering situations that presented a serious risk to Apple and that could arise before switching costs had risen sufficiently to discourage opportunistic behavior. First, the pricing is subject to a significant condition regarding quality:

This Agreement and the Pricing Schedules are based on the assumption that SCI can produce the Products at quality levels suitable for shipment directly to Apple's distribution system. SCI's inability to achieve certification status as defined in Exhibit E, will create a significant increase in costs to Apple. SCI will develop a plan to meet such requirements and understands that failure to achieve certification status within a reasonable time frame may result in disqualification as an approved Apple supplier.

Id. art. 9.3. While the consequences of a failure to achieve certification or the loss of the approved supplier label are not spelled out, they appear to be serious, presumably relieving Apple of future purchase obligations. Second, the collaborative process will give SCI information about Apple's future plans, information that might be valuable to Apple's competitors. The contractual response to SCI misbehavior concerning Apple's competitively sensitive data is draconian: "Apple may terminate this Agreement effective immediately upon written notice to SCI if SCI materially breaches its obligation of confidentiality." *Id.* art. 18.1(b).

the contract no longer binds either party; it leaves the parties free to go forward, unconstrained by the detailed terms of the contract.

Importantly, the Apple–SCI contract differs from the Deere–Stanadyne contract in that both Apple and SCI understand that Apple will have other suppliers providing precisely the products that SCI is providing. Thus, building an informal enforcement mechanism is more difficult: The costs of finding and learning about replacement suppliers are lower because Apple already has this information about its parallel suppliers. In order to understand how the Apple–SCI arrangement operates, therefore, we have to distinguish between manufacturing on the one hand, and collaborative innovation on the other. To the extent that a manufacturer is only a specifications taker—that is, it manufactures to the specifications Apple develops with its collaborative manufacturers—multiple suppliers pose a risk that Apple will switch from one noncollaborative supplier to others. To the collaborative manufacturers, in contrast, the fact of multiple suppliers may not create the same risk. Apple needs to learn about a new supplier’s capacity to make a contribution. To be sure, they can switch production to one of their other suppliers, but there are two frictions. First, some of the suppliers may be just manufacturers and not good at iterative collaboration. Switching more production to them without finding another supplier who can contribute to innovation leaves Apple worse off. Alternatively, suppose all suppliers help innovate but come up with different insights, the best of which Apple builds into the specifications for all manufacturers. Then Apple is choosing how many suppliers to have; a reduction in the optimum number of collaborative innovators is a cost to Apple of cheating.

C. *The Warner-Lambert–Ligand Agreement*

The development of new drugs based on biotechnology often entails contracting across organizational boundaries.⁹⁴ Large pharmaceutical companies frequently lack the depth of scientific knowledge and experience that provide the foundation for biotech research. Smaller biotech firms typically lack the experience and capital to both take the drugs through the arduous process of obtaining FDA approval and commercially market the drug.⁹⁵ Firms engage in a broad range of collaborative arrangements including joint ventures, licensing agreements, and co-de-

94. For example, Robinson and Stuart report that some twenty-five percent of the twenty-six billion dollars of industry-financed (as opposed to university-financed or nonprofit-financed) pharmaceutical research and development was done in collaborative agreements between separate entities. David T. Robinson & Toby E. Stuart, *Financial Contracting in Biotech Strategic Alliances*, 50 J.L. & Econ. 559, 559 (2007) [hereinafter Robinson & Stuart, *Financial Contracting*].

95. Cf. *id.* at 563–64 (detailing high costs associated with pharmaceutical regulatory approval and marketing).

velopment deals.⁹⁶ We focus here on a particular contract (the “Warner-Lambert–Ligand agreement”):⁹⁷ a research, development, and license agreement between Warner-Lambert, a large pharmaceutical company, and Ligand Pharmaceuticals, a much smaller biotech company, to discover and/or design small-molecule compounds that act through the estrogen receptors, to develop those compounds into pharmaceutical products, and to take those products through the FDA approval process and through commercialization.

A brief description of the drug development process provides a context for the Warner-Lambert–Ligand agreement.⁹⁸ The initial screening of compounds and preclinical work takes, on average, three to six years. During that period, the number of compounds under consideration is winnowed from 5,000–10,000 down to a quite small number through scientific and animal testing. At that point, an application for an Investigational New Drug is filed with the FDA. If the FDA approves, the drug can move to clinical testing on humans. Clinical testing takes another six to seven years. That period is broken down into three phases: Phase one tests include fewer than 100 persons, phase two between 100 and 500, and phase three between 1,000 and 5,000. If the drug surmounts these hurdles, the sponsoring company submits a New Drug Application (NDA) with supporting documentation. FDA review of the NDA can take another six months to two years. If the FDA approves, the drug can be brought to market. Estimates are that out of 5,000 to 10,000 compounds, only 250 enter preclinical testing,⁹⁹ and only about twenty percent of drugs that begin phase one testing are ultimately approved by the FDA.¹⁰⁰ Only upon approval does the pharmaceutical company discover whether the drug will be successful commercially.

In the research stage of the project, the Warner-Lambert–Ligand agreement stipulates that Ligand will engage in directed research, with Warner-Lambert providing the bulk of the funding. In this phase, Warner-Lambert monitors the work and has options to abandon in the event that the research proves unpromising. If the project ultimately suc-

96. See, e.g., Leslie Gladstone Restaino, *BioPharma Collaborative Agreements: Choosing the Right Deal Structure*, Metropolitan Corp. Couns., Nov. 2007, at 47, available at <http://www.metrocorpconsult.com/pdf/2007/November/47.pdf> (on file with the *Columbia Law Review*) (discussing history and types of collaborative arrangements).

97. We will also draw upon Ligand’s contract of May 19, 2000 with Bristol-Myers Squibb to discover small-molecule compounds which act as modulators of the mineralocorticoid receptor.

98. Because of the regulatory structure, the components of the process are standardized. This description of the drug-development process is based on Innovation.org, *Drug Discovery and Development* (2007), at http://www.innovation.org/drug_discovery/objects/pdf/RD_Brochure.pdf (on file with the *Columbia Law Review*).

99. Nat’l Inst. of Health, Dep’t of Health and Human Servs., *A Plan to Ensure Taxpayers’ Interests Are Protected* (2001), available at <http://www.nih.gov/news/070101wyden.htm> (on file with the *Columbia Law Review*).

100. Joseph A. DiMasi et al., *The Price of Innovation: New Evidence of Drug Development Costs*, 22 J. Health Econ. 151, 165 (2003).

ceeds, only a small fraction of the costs would be associated with the research phase. The major costs of bringing a drug to market are incurred in the later stages, in which the manufacturer must prove efficacy and safety through clinical studies in the FDA approval process.¹⁰¹ Once a successful compound has been identified, Ligand's role is largely over, and, as a result, its role in decisionmaking largely disappears. Most decisions in this phase are at the "sole discretion" of Warner-Lambert.¹⁰²

The research stage is divided into three periods, with Warner-Lambert having an option to abandon in the first two. The Exploratory phase lasts fifteen months, after which the project terminates unless Warner-Lambert gives at least one month's written notice of its intention to enter into the second phase, the Extension term. The Extension term is three years. Funding levels for both periods are specified in dollars per Full-Time Equivalent (FTE).¹⁰³

During the research stage, the parties' collaboration is centered on a Joint Research Committee (JRC), an elaborate governance structure responsible for reviewing and directing all scientific activities in the preclinical period. In effect, the JRC determines the research path. The com-

101. A growing literature concerning strategic alliances emphasizes three different elements: (1) the role of strategic alliances as an alternative financing vehicle to venture capital; (2) the role of networks in developing reputations that help support strategic alliances, and (3) the choice of a strategic alliance through which to carry on an activity as opposed to undertaking the activity within the existing entity. See, e.g., Robinson & Stuart, *Financial Contracting*, supra note 94, at 561–62 (comparing strategic alliances with venture capital arrangements); David T. Robinson & Toby E. Stuart, *Network Effects in the Governance of Strategic Alliances*, 23 *J.L. Econ. & Org.* 242, 243 (2007) [hereinafter Robinson & Stuart, *Network Effects*] (arguing that "the stock of past alliances gives rise to a communication network that affects the allocation of control in strategic alliance agreements by allowing long-term reputational concerns to affect the terms of specific contracts"); David T. Robinson, *Strategic Alliances and the Boundaries of the Firm*, 21 *Rev. Fin. Stud.* 649, 651–52 (2008) (finding that firms prefer strategic alliances to internal expansion when new activity pursued is risky and/or unrelated to firm's primary line of business). The three categories, however, share one common characteristic: In their analysis of the terms of strategic alliances, the need to support the collaborative innovation that is at the heart of the substantive transaction is typically ignored.

102. Warner-Lambert Co. & Ligand Pharmaceuticals Inc., *Research, Development, and License Agreement* art. 3.1.1 (Sept. 1, 1999) (on file with the *Columbia Law Review*) [hereinafter *Warner-Lambert–Ligand Agreement*]. In addition, Warner-Lambert contemporaneously purchased approximately seven percent of Ligand's equity. Thus, the contract contained two linked but distinct agreements: one covering the research and development of small-molecule compounds that act through the estrogen receptors and another covering the financing of the sale of equity in one research partner to another. Such equity investments are not uncommon in research-and-development joint ventures. See Robinson & Stuart, *Network Effects*, supra note 101, at 242–43.

103. Warner-Lambert–Ligand Agreement, supra note 102, art. 2.8. "Full-Time Equivalent" (FTE) is defined as "one or more researchers with appropriate qualifications employed by Ligand or Warner-Lambert and assigned to work on the Collaboration with such time and effort to constitute one such researcher working on the Collaboration on a full-time basis for no less than *** hours per year." *Id.* art. 1. Warner-Lambert then has the right to add further extension terms. To trigger those extensions, Warner-Lambert has to provide written notice and a financial commitment to support a certain level of activity.

mittee consists of three members from each firm, and all decisions must be unanimous.¹⁰⁴ In the event of a disagreement, the dispute goes to Ligand's CEO and the president of Warner-Lambert's Pharmaceutical Research Division for "good faith resolution" within a specified period.¹⁰⁵ If they fail to resolve the dispute, the parties are free to pursue legal remedies. As with the other agreements we discuss, the threat to line managers of having to explain to senior executives of both companies the failure to effectively cooperate likely carries more weight than the threat of legal action.

Nevertheless, in contrast to the JRC collaborative governance structures, Warner-Lambert's options to abandon the project are unilateral. If the science proves unpromising, it can terminate with little or no direct cost. There is an indirect cost, however, in that it agrees not to pursue research in the field for a specified period. If Warner-Lambert terminates after the Extension period, the likelihood that something valuable has been produced in the interim will have increased, and the termination might be opportunistic. The contract reflects that concern. If Warner-Lambert were to conduct independent research on a collaborative compound in the field and file an Investigational New Drug Application within a defined period after termination, it would have to pay royalties.¹⁰⁶

As the project moves from the research to the development stage, regulatory and market experience become more important. The cost of the project, all of which will be borne by Warner-Lambert, also increases exponentially. As a result, both responsibility and decisionmaking shift to Warner-Lambert: The JRC's role ends with the completion of the research phase. Under the agreement, the JRC recommends which compounds to pursue, but the decision to go forward rests in Warner-Lambert's sole discretion. Warner-Lambert promises to "use diligent efforts to pursue the Clinical Development and commercialization of each Collaboration Lead Compound at its own expense"; however, it "shall have the sole discretion to determine (a) which Products to develop or market or to continue to develop or market, (b) which Products to seek

104. *Id.* art. 3.1.4.

105. *Id.* art. 3.2.

106. *Id.* art. 12.3. The termination provisions in the contract are quite complicated. Warner-Lambert, as noted in the text, has the option to abandon twice during the research term, at the end of the Exploratory and Extension terms. In both instances, the background technology would be returned. If Warner-Lambert terminates after the Exploratory term, it would grant Ligand an exclusive royalty-free license to use a number of compounds of Ligand's choice that have exhibited "field activity." *Id.* art. 12.2. If Warner-Lambert concludes that Ligand's work in the research phase is unsatisfactory, it can unilaterally terminate the agreement. In such a case, Warner-Lambert would have exclusive rights to develop a certain number of background technology compounds in the field, and it could choose which; the rights to the others would revert to Ligand. Moreover, if Warner-Lambert were successful, Ligand would be entitled to milestone payments and royalties. *Id.* art. 12.8.

regulatory approval for, and (c) when and where and how and on what terms and conditions, to market such Products in the Territory.”¹⁰⁷

The significant *ex ante* uncertainty associated with the collaboration is reflected in the manner in which Ligand is compensated. As we saw, the gap between contract formation and the appearance of a marketable drug is likely to be more than a decade.¹⁰⁸ So the nature of that drug and its potential value (clinically and financially) is unknown at the time of contracting. Because of that uncertainty, Ligand’s compensation is carefully structured. First, as noted, it is to be paid for some fraction (perhaps all) of the FTE assigned to the task. Second, it is paid a fixed fee upon the initial screening of the Warner-Lambert compound library; if Warner-Lambert chooses to go ahead with the Extension term, Ligand will receive an additional fixed fee. These could be labeled milestone payments, although the contract does not do so. The agreement does establish a number of specific milestones, and, upon reaching each milestone, Ligand will receive an additional payment. Finally, if the research produces marketable products, Ligand will receive royalty payments on sales.

The risk of opportunistic behavior by Warner-Lambert is mitigated by matching options. Warner-Lambert might decide not to proceed to the development stage for a variety of reasons. The most significant is a genuine belief that the project will fail to yield a commercially viable outcome. Even if Warner-Lambert believes the project viable, it might want to defer development in favor of a more promising alternative in another field. Or it might act opportunistically, feigning disappointment with the intent of renegotiating the financial terms. Which of these alternatives in fact motivates Warner-Lambert’s decision to abandon might well be difficult for Ligand even to observe, let alone to verify. This difficulty is addressed by giving Ligand a matching option. If Warner-Lambert decides not to proceed with the development of a particular “collaboration lead compound,” then Ligand has the right to develop and commercialize it.¹⁰⁹ To decide whether to exercise its option, Ligand would need only to be able to assess the commercial viability of the product; realistically, it would have to convince a replacement partner of the product’s viability. Of course, the later in the process Warner-Lambert exercises its option to abandon, the more observable should be the product’s viability, and the more effective Ligand’s matching option in deterring opportunism.

107. *Id.* art. 4.2.

108. See *supra* text accompanying notes 98–100.

109. Warner-Lambert–Ligand Agreement, *supra* note 102, art. 5.3.1. Ligand’s right is qualified; it cannot go forward if Warner-Lambert is commercializing the compound for another use or has a competing product (either existing or in the pipeline).

IV. A THEORY OF CONTRACTUAL COLLABORATION AND CO-DESIGN

As we discussed in Part II, transactions involving collaborative innovation across organizational boundaries have distinctive features that affect their contractual structure.¹¹⁰ In particular, the design and specification of product characteristics cannot be contracted for *ex ante*; rather, they will result from repeated collaboration by employees of both firms. The contracting problem is to craft a structure that (1) induces efficient, transaction-specific investment by both parties; (2) establishes a framework for iterative collaboration and adjustment of the parties' obligations under conditions of continuing uncertainty—responding, that is, to coordination cascades; and (3) limits the risk of opportunism that could undermine the incentive to make relation-specific investments in the first place. Each of our contract exemplars addresses those three aims.

The three contracts do so by incorporating a mix of formal and informal contract mechanisms. Successful collaborative innovation requires ample knowledge of the collaborating parties' capabilities and substantial confidence in the parties' future cooperative behavior. Neither the knowledge nor the confidence can be acquired or assured by formal contract alone. At the same time, the commercial context makes it infeasible to build up both over time through repeated exchanges policed by the expectation of future dealings.¹¹¹ In the arrangements of interest here, the innovative product (or products) must be created early in the relationship; there is thus no assurance of a number of future rounds sufficiently large that the expectation of a long-term relationship and the discipline of repeated dealings will protect against opportunistic behavior.¹¹² Instead, in these contracts, formal contracting operates importantly to facilitate the development of informal contracting structures that police the parties' expectations of capability, cooperation, and trust.¹¹³

110. See *supra* Part II.A.

111. See generally Avinash K. Dixit, *Lawlessness and Economics: Alternative Modes of Governance* 14–21 (2004) (outlining theory of repeated games in which “the prospect that future benefits depend on current performance may suffice to induce” contractual parties to behave cooperatively).

112. The theorem of repeated dealings underlying this conclusion is the “folk theorem” of noncooperative game theory. See generally *id.* at 61 (noting that certain models of repeated games with imperfect monitoring are limited in practical applicability because of their “folk theorem” focus on the “limiting case where . . . players are extremely patient”); Drew Fudenberg & Eric Maskin, *The Folk Theorem in Repeated Games with Discounting or with Incomplete Information*, 54 *Econometrica* 533, 536–39 (1986) (outlining “the classical folk theorem”).

113. Laura Poppo and Todd Zenger raise the potential for this interactive relation between formal and informal contracting but base their analysis on interviews with executives involved in the contracting process rather than on assessments of the actual contractual techniques that create and sustain the relationship. See Laura Poppo & Todd Zenger, *Do Formal Contracts and Relational Governance Function as Substitutes or Complements?*, 23 *Strategic Mgmt. J.* 707, 712–15, 721 (2002). Thus, our focus here on actual contracts and contracting techniques extends their insightful intuition.

Such mechanisms would be unnecessary if we assumed that these new collaborative arrangements operated in a contractual state of nature, in which good faith cooperation could be expected from an arm's length partner despite the absence of effective enforcement and despite the parties' inability to specify the substance of their joint efforts. We make no such assumption. Rather, as we explain more fully in Part IV.C, process-oriented, formal contracting supports the rapid development of informal contracting techniques that address, in turn, the substantive elements of the parties' performance. In the end, the same parties that must cooperate to create the innovation also must bargain noncooperatively over how the gains created by that innovation are shared, and, in some—perhaps many—cases, the collaborative effort will be less successful than anticipated. The delay of delivery of the Boeing 787 because of slow performance by some suppliers is a stark, current example.¹¹⁴ Thus, cooperation does not eliminate tension or conflict;¹¹⁵ the package of interrelated formal and informal contractual techniques described below allows the cooperation necessary for iterative collaborative innovation while also allowing the parties to accommodate their conflicting interests in division of joint gains.

In the discussion that follows, we set out a theory that seeks to explain how the exemplar contracts we study are structured so as to support collaborative contracting. We begin in Part IV.A by describing the transactional variations in our exemplar contracts, focusing primarily on whether the intended collaboration is long-term or of a relatively discrete duration. Despite these variations, the analysis in Part IV.B identifies a key similarity: Each contract creates a governance mechanism that is structured to produce information about the parties' capabilities to innovate and their capacities to cooperate. In Part IV.C, we describe the key enforcement mechanisms that result from this governance structure: a rich braiding of formal and informal terms that deters opportunism during the collaborative/learning phase of the contract and a series of nested options that applies once the parties reach an end stage in their collaboration.

A. Transactional Variations in Collaborative Contract Design

The characteristic differences in the new contractual patterns we observe are driven by the nature of the barriers to ex post opportunism, which in turn are dictated by the substance of the transaction. The structure of contracts for collaborative innovation differ most importantly depending on whether the contemplated collaboration is long-term—involving an ongoing stream of interactive innovations—or whether it

114. See, e.g., Boeing Announces Another 787 Delay, L.A. Times, Apr. 10, 2008, at C6.

115. Alan O'Sullivan details the wide range of conflict that is possible even in a successful collaborative contract for innovation. See O'Sullivan, *supra* note 51, at 236–44.

involves a discrete project aimed at producing a single innovation such as a patentable product or process. As we will see, in the former case the barrier to ex post opportunism arises from the collaboration process itself: The continued presence of uncertainty makes impossible the ex ante allocation of ex post decision power through assigning to one party options to take action like termination. In the discrete project setting, informal mechanisms operating during the collaborative period discourage opportunism—in particular, the appropriation of jointly produced information for private purposes—but the parties have to fear opportunistic renegotiation once the cooperative stage of the project is completed. The only issue then remaining is division of the gains from prior cooperation. As a result, an explicit constraint on opportunism must be employed, but, at this stage, the uncertainty having been resolved, the contract theory solution of allocating decisionmaking rights is feasible.

At one end of the spectrum is contracting for long-term collaborative innovation, particularly well illustrated by the Deere–Stanadyne agreement,¹¹⁶ which is almost entirely focused on building knowledge of Stanadyne’s capabilities as a collaborator in future innovation. To be sure, the agreement provides the terms of actual purchases, should Deere in fact make them, but the agreement itself does not commit Deere to purchase anything at all. Instead, the bulk of the work is done by non-contractual programs that create settings for Stanadyne and other suppliers to demonstrate their capacity for collaboration over time by ascending the ranks of favored suppliers. The agreement plainly contemplates a long-term relationship between Deere and Stanadyne, despite the impossibility of specifying what products will turn out to be needed in the face of the uncertain future of Deere’s market. As each party learns about the other’s capabilities and character, the costs of extracting private benefits at the risk of undermining collective gains continue to rise.

The Apple–SCI agreement presents a mixed case between an explicit, three-year supply contract and an implicit, long-term contract that contemplates collaborative innovation.¹¹⁷ Recall that in this agreement the three-year, explicit supply arrangement supported Apple’s sale to SCI of Apple’s manufacturing plant.¹¹⁸ In the short run, the product that SCI would manufacture for Apple was fully specified: The agreement contemplated a turnkey sale to SCI with SCI immediately charged with producing sixty percent of Apple’s requirements for precisely the same computers Apple produced at the same plant until the sale. However, that product would evolve over time in step with changes in markets and technology, with the differences from the original specifications accumulating. The agreement provided an elaborate planning process to ensure that the manufacturer and assembler (SCI) and the seller of the product (Apple) would jointly address the changes in the manufacturing/assem-

116. See *supra* Part III.A (detailing provisions of Deere–Stanadyne agreement).

117. See *supra* Part III.B (detailing provisions of Apple–SCI agreement).

118. See *supra* notes 85–89 and accompanying text.

bly process and in the product. The process calls for continued improvement in performance and cost through collaboration and jointly determined benchmarking.

As Apple's three-year purchase commitment winds down, the explicit terms of the contract no longer protect SCI's investment in the plant from the threat of termination by Apple. At the same time, Apple's own exposure grows as the manufacturing process reflects increasingly more input by SCI through the co-design process, and therefore a more and more complex bundle of explicit and tacit knowledge. As a result, at one time or another, either party may be in a position to extract a larger part of the gain from the relationship by threatening to withhold further cooperation. While each party makes a front-end specific investment—SCI in buying the plant and Apple in agreeing to purchase specific amounts of its supplies from SCI and thus relinquishing control over a large part of its computer production—these threats may be credible because the specific investments are not necessarily symmetric; each party is vulnerable to hold-up at different times during the collaborative period.

Thus, a co-design relationship in the prototypical design and adaptation arrangement has to address the traditional hold-up problem associated with specific investment, but must do so in an environment where ex post decision rights are not contractible *ex ante*—the iterative and collaborative process makes it impossible to specify what decisions will need to be made. For the project to work, the parties must rely on an informal arrangement to constrain opportunistic renegotiation of the division of the value created by success during the three-year period of the contract, and in setting the terms of the parties' relationship thereafter.

At the opposite end of the spectrum from long-term collaboration is contracting for a discrete project. Discrete-project contracting contemplates collaboration during the development stage but allows for opportunism when the development stage is completed. The Warner-Lambert–Ligand agreement provides a good illustration.¹¹⁹ In that agreement, the parties contemplate a joint effort to develop pharmaceutical products having specified capabilities, with Ligand playing the leading role in the research-and-development stage and Warner-Lambert playing the primary role in the compounds' commercialization: clinical trials, FDA approval, and the like. The problem is that once the compound is developed—once the collaborative effort has been completed—Warner-Lambert seems to be in a position to take advantage of the sequential structure of the arrangement. After Ligand has performed its portion of the effort to identify particular compounds, Warner-Lambert must initiate the regulatory process and commercialization, which enormously increases its financial commitment. At that point, the potential for opportunistic renegotiation resurfaces. Warner-Lambert might seek to

119. See *supra* Part III.C (detailing provisions of Warner-Lambert–Ligand agreement).

renegotiate downward the previously contracted royalty terms through which Ligand will share in the value created by the collaboration. At this stage of the relationship, uncertainty has been resolved and the respective decision rights of the parties are therefore contractible—that is, there is a discrete decision to go forward with commercializing the product rather than a series of iterations as in the collaborative stage. Hence, an explicit contractual device is both necessary and feasible.

B. *Coordination Cascades and the New Governance Mechanisms*

While the contractual mechanisms employed will vary depending on the needs of the parties—especially, the length of the collaboration—the mechanisms themselves fall into several functional categories. Since the collaboration is designed to produce information about the parties' capabilities and their capacities for cooperation, the collaborative/learning phase initially requires that both parties invest in producing information. In turn, mutual investment in information deters opportunism during the collaboration by locking in the participants and by screening out naturally opportunistic counterparties. Finally, allocating decision rights at the end stage deters opportunism for those collaborations that are not open-ended. In the following discussion, we describe these functional categories and the contractual mechanisms that fall within them.

1. *The Function of the Collaborative/Learning Phase.* — A key set of contractual mechanisms regulates the collaborative/learning process. The set comprises two types: those mechanisms designed to facilitate iterative investments in information and those designed to improve the quality of the information produced. We discuss each in turn.

a. *Iterative Investments in Information.* — Our three illustrative contracts—Deere–Stanadyne, Apple–SCI, and Warner–Lambert–Ligand—share a common feature. Each establishes formal governance structures: processes of interaction and dispute resolution. One might imagine many reasons for writing down adaptation protocols: The process builds consensus, enhances learning, minimizes misunderstanding, and the like. But none of those reasons explains why the elaborate governance structures in these contracts are made part of a formal contract. The first step, therefore, is to find a theoretical framework that addresses that question.

The role of the contract as a nexus for the parties to invest in learning about each other's capabilities is seen most starkly in the Deere–Stanadyne agreement, where it is the only binding element of the contract. From Stanadyne's (or any new supplier's) perspective, the contract is an invitation to enter the Achieving Excellence program, through which the parties will learn about each other, and Deere's experience will move the supplier up the supplier status ladder. The parties' investments in the relationship grow as the process continues, providing the supplier the assurance that Deere will place orders with it. Recall that neither the contract itself, nor the Achieving Excellence program, obligates Deere to purchase any product at all. In short, the contract operates entirely as a

governance structure that facilitates learning about the parties' capabilities.

From this perspective, the Deere–Stanadyne contract resembles a more famous contract that, over the years, has been the focus of a great deal of academic attention: the General Motors–Fisher Body supply contract for the supply of auto bodies to GM in the 1920s. As Victor Goldberg has recently shown, the General Motors–Fisher Body supply contract was, in truth, legally unenforceable, a fact that previously had been missed in the large literature concerning the contract but that the lawyers who wrote the contract would have known.¹²⁰ Thus, this more famous contract operated like the Deere–Stanadyne contract: as a way of organizing the parties' learning and continued collaboration, their expectations with the investment protected by ever increasing mutual dependency.

A governance arrangement is also at the center of the Apple–SCI agreement. The parties commit to prepare a product plan, which will contain the specifications, quality requirements, price schedule, and other terms. The contract gives the parties joint responsibility for all manufacturing design and equipment technology, testing, tooling, and the like. SCI controls the test engineering process,¹²¹ but Apple monitors, requiring progress reports, preproduction review, and progress monitoring to achieve preproduction deliverables.¹²² The formal contract thus specifies an iterative process in which the parties cooperate in designing the product and the manufacturing process. The contract allocates initial responsibility to SCI for test engineering while Apple monitors their performance and receives progress reports leading to an agreement on product characteristics and the manufacturing process, the substance of which is not specified in the contract. Only after Apple signs off on the preproduction deliverables does real production begin. The parties agree to co-engineer product cost reductions, to set goals, and to meet quarterly while production is underway. At each step in the preproduction process, the parties learn about each other's technical capacity to accomplish the necessary design tasks and ability to cooperate in a productive fashion, to share information readily, and to effectively manage the employees involved despite the nonhierarchical characteristics of the relationship. This process takes on greater and greater importance with time. As the product evolves in response to technology and market

120. See Victor P. Goldberg, *Lawyers Asleep at the Wheel? The GM-Fisher Body Contract*, 17 *Indus. & Corp. Change* 1071, 1076–80 (2008). In particular, it seems that GM wanted to learn from the Fisher brothers, both of whom were given seats on GM's board, how better to organize flexible, co-design relations with its other suppliers—a capacity that Fisher Body had demonstrated in its relations with demanding customers like Chrysler. See Helper et al., *supra* note 47, at 452–53.

121. Apple–SCI Agreement, *supra* note 90, art. 5.2.

122. *Id.* art. 5.3.

changes, legacy specifications and processes become less important; the collaborative process represents the future.

Paralleling the iterative co-design process, the Apple–SCI contract also specifies a dispute resolution process, which appears to develop information about each party’s character. The contract provides that each party designate one person to be the single operations manager for the day-to-day administration of the agreement. When the two individuals cannot resolve an issue, the issue moves up (and then back down) the hierarchies in a fashion that must be observable to each company’s designee, if the process is to work. And of course, another critical piece of information—whether it works—also will be observable.

The Apple–SCI agreement thus has parallel formal structures that function to provide critical information about the characteristics central to the success of a project requiring collaborative innovation: Do the parties have the technical capabilities? Do the parties have the skills necessary to manage cooperation? And what are the parties’ dispute-resolution styles—does each cooperate or fight? The continued interaction over the initial three-year period, especially as the results of the collaboration represent a higher percentage of the product being produced, leads to increasing information on each of these dimensions.

The Warner–Lambert–Ligand agreement has similar explicit terms that dictate a structure for the continued interaction of the parties, which in turn results in a learning process similar to that in the Deere–Stanadyne and Apple–SCI agreements. During the first stage of the contract, Ligand uses its proprietary technology to identify compounds that hold promise for estrogen therapy. Warner–Lambert then has the option, based on Ligand’s performance, of extending the arrangement into a developmental period, when a “Joint Research Committee” composed of three representatives from each firm manages further research. Decisions must be unanimous; disagreements within the committee are resolved in conference between one senior manager from each company. Both parties have a large incentive to collaborate in the pre-exercise period, and to learn about the other party’s capabilities and characteristics. Ligand’s incentive is to provide enough information to cause Warner–Lambert to exercise its option to continue the project. Warner–Lambert’s incentive is that the value of its option goes up as it gains more information and is better able to predict the probability distribution of the payoffs to further investment in the compounds identified.¹²³

123. The option to extend also operates as a valuable right to abandon the project, either because the project is not working or because Warner–Lambert is not satisfied with Ligand’s performance. See generally *supra* note 103 (discussing Warner–Lambert extension option). In two recent articles, Robinson and Stuart explore the role of networks in supporting reputational sanctions in biotech strategic alliances. Robinson & Stuart, *Financial Contracting*, *supra* note 94; Robinson & Stuart, *Network Effects*, *supra* note 101. The articles conclude that the centrality of the biotech company in the industry

In sum, the three contracts share a common structure: explicit mechanisms with respect to collaboration and dispute resolution. Thus, during the collaboration period, the production of information about the contracting parties is supported by the formally specified collaboration and dispute resolution processes.

b. *The Role of the Contract Referee Mechanism.* — Despite the central role of the collaborative/learning phase in providing the parties valuable information about each other's character, it cannot guarantee a cooperative outcome. Common experience teaches that transactions relying on informal enforcement can break down because relational enforcement requires that each party be able to observe and properly characterize the other's behavior. The required transparency dissipates when adaptations are complex and the sequence of performances are interrelated, precisely the circumstance in contracts for collaborative innovation, where debate over the right strategy is to be expected. In such circumstances, the parties' signals are noisy: They do not perfectly demonstrate whether the disagreement is in good faith, part of the cooperative process, or, instead, an indication of opportunism. In complex interactions, disputes may or may not reflect a failure to cooperate; what one party intends as a cooperative response—"good idea but what about . . ."—may be mistakenly interpreted as a defection.¹²⁴ Lacking clarity, either party may mischaracterize the other's actions. In these circumstances, without the necessary linkage between action and response, reciprocity will be a less effective mechanism for enforcement.

network of alliances is the key to explaining both the use of nonverifiable contract terms like best efforts and the commitment to treat the alliance as seriously as the biotech does other alliances in the governing contracts. Centrality is treated as a proxy for reputation; the conclusion is that reputation operates as an implicit enforcement mechanism. While we recognize the role of reputation as one element of switching costs, we remain skeptical about the extent to which reputation can carry the weight Robinson and Stuart assign to it. Most important, it is extremely difficult for third parties, however well-connected, to observe the conduct of the parties. Suppose a venture fails. Given the very low likelihood of finding a successful drug, the most reasonable inference is that the outcome is the result of bad luck, not poor skills or bad faith. From this perspective, reputation is hard to gain, but it is also hard to lose. Both require repetitive results to separate the signal from the noise. From a contracting for innovation perspective, a better explanation for the use of terms that are difficult to observe or verify is to help set expectations for the nature of the parties' ongoing collaboration. The point is not to impose a standard that will trigger sanctions but to identify a goal that will help organize the collaborative effort. See Oliver Hart & John Moore, *Contracts as Reference Points* 1, 4–13 (Nov. 2006) (unpublished manuscript, on file with the *Columbia Law Review*), available at <http://ssrn.com/abstract=944784> (developing contractual model based on view that "a contract provides a reference point for the parties' trading relationship").

124. If one party mistakenly observes a defection when the other party intended to cooperate, a tit-for-tat strategy will collapse into repetitive retaliation. For an accessible account of the problem for cooperation created by "noise" impacting parties' perceptions of their counterparties' actions, see generally Jonathan Bendor et al., *When in Doubt . . . : Cooperation in a Noisy Prisoner's Dilemma*, 35 J. Conflict Res. 691, 712–14 (1991).

The risk of misinterpreting the other party's actions is mitigated in our exemplar contracts by what we call the "contract referee" mechanism. This part of the governance structure typically combines three key elements: (a) the commitment to share and exchange information during the collaboration,¹²⁵ (b) the assignment of decision rights to a joint project management team subject to a unanimity rule,¹²⁶ and (c) the appointment of "referees"—representatives from each firm charged with resolving disputes.¹²⁷

The contract referee mechanism has several effects. First, the referees provide information concerning the nature of a complex interaction that others cannot obtain directly. A referee can clarify misunderstandings early, avoiding false negatives—i.e., the interpretation of the other's behavior as a defection. When she finds that a defection has indeed occurred, a referee can, by "blowing the whistle" while providing for a fast and low-cost resolution to the dispute, forestall disproportionate responses by the aggrieved party. These steps promote the development of a cooperative equilibrium and reinforce the "lock-in" effects that have been experimentally observed.¹²⁸ The referee also serves as an informal

125. See, for example, the Warner-Lambert-Ligand contract, where the parties agree to disclose to each other all background technology relevant to the field and helpful to performing the work set out in the Research Plan. Both commit to provide resources necessary to carry out the Plan. Warner-Lambert-Ligand Agreement, *supra* note 102, art. 2.1. Each agrees to allocate a certain number of FTEs to the Plan. *Id.* art. 2.3. All collaboration discoveries and technology are to be "promptly disclosed" to either party, and monthly reports are required on all biological structures and compounds. *Id.* art. 2.7. The contract calls for maintenance and inspection of records. Similarly, the Apple-SCI contract provides that the parties will co-engineer product cost reductions, setting goals and meeting every three months during the term. Apple-SCI Agreement, *supra* note 90, art. 10.6. SCI agrees to share all production-cost information, and Apple agrees to share all marketing forecasts. *Id.* arts. 10.6(b), 11.1. The parties co-design lead-time reductions. *Id.* art. 11.7.

126. While a unanimity rule is common to many of these contracts, it is not an essential feature. Typically, a unanimity rule exacerbates hold-up problems but here parties deliberately contract into it. In this context, not exercising the option to behave opportunistically by taking advantage of the unanimity rule is a credible signal of cooperation. The Warner-Lambert-Ligand contract is typical in providing for such an elaborate management structure. The Joint Research Committee (JRC) consists of three representatives from each firm, with one of the Ligand representatives serving as the Chair. Warner-Lambert-Ligand Agreement, *supra* note 102, art. 3.1.1. The JRC has control and decision rights over the Research Plan and is responsible for co-design and implementation. *Id.* art. 3.1.2. The JRC holds quarterly meetings, and all decisions are by unanimous vote. *Id.* arts. 3.1.3–1.4.

127. See, for example, the Warner-Lambert-Ligand contract, which provides that all decisions are by unanimous vote, *id.* art. 3.1.4, and disagreements are resolved by the CEO of Ligand and the president of Warner-Lambert's Pharmaceutical division (or their designees). *Id.* art. 3.2.

128. See, e.g., John McMillan & Christopher Woodruff, Private Order Under Dysfunctional Public Order, 98 Mich. L. Rev. 2421, 2422 (2000) (noting how buyer may be "locked in with a particular seller, either because the seller is a monopolist or because the buyer would face high costs of locating an alternative seller[, and thus] the seller can make the contract self-enforcing by cutting off further dealings").

disciplining mechanism. Superiors are unlikely to look with favor on subordinates who send problems up the line for resolution. The subordinates' job is to resolve problems, not escalate them. Second, the use of collaborative teams also disciplines shirking, particularly if there is a unanimity requirement and all will be punished for the sins of a few.¹²⁹ Third, the collaboration process produces symmetrical information about two key variables: the value created by the collaboration and the preferences for reciprocal cooperation. So long as both parties are symmetrically informed and temptations to play chicken are muted, then collaborative interaction—Coasian bargaining—will lead to efficient decisions. Thus, the contract referee mechanism functions both to reduce moral hazard and to improve the quality of the information shared by the parties.

Next, we focus on the potential for opportunism resulting from first the collaborative and then the pie-splitting phases of the transactions, and on the characteristic responses.

2. *Mechanisms for Deterring Ongoing Opportunism.* — The risk of opportunistic behavior during the collaboration increases as the parties make asymmetric relation-specific investments. However, the information produced by the contractual mechanisms discussed above alleviates this problem in two ways. First, it builds switching costs and thus locks the parties into the relationship. And second, it screens out naturally opportunistic counterparties. We address each feature in turn.

a. *Building Switching Costs During the Collaborative/Learning Phase.* — Central to our analysis of these new arrangements is the role of switching costs—costs that a party must incur to change from one counterparty to another.¹³⁰ The term is connected most directly in the literature with inquiry into whether rigorous initial competition adequately compensates for the absence of competition after the first purchase;¹³¹ but it also ap-

129. Stewart Macaulay noted this forty years ago:

[I]nternal sanctions will induce performance. For example, sales personnel must face angry customers when there has been a late or defective performance. The salesmen do not enjoy this and will put pressure on the production personnel responsible for the default. If the production personnel default too often, they will be fired.

Stewart Macaulay, *Non-Contractual Relations in Business: A Preliminary Study*, 28 *Am. Soc. Rev.* 55, 63 (1963). Macaulay also described feedback mechanisms where buyers issue "report cards" to management of sellers, report cards that can then be used to discipline personnel. See *id.*

130. See generally Joseph Farrell & Paul Klemperer, *Coordination and Lock-In: Competition with Switching Costs and Network Effects*, in 3 *Handbook of Industrial Organization* 1967, 1972–74 (Mark Armstrong & Ronald H. Porter eds., 2007) (defining switching costs).

131. Paul Klemperer, *Switching Costs*, in 8 *The New Palgrave Dictionary of Economics* 125, 125 (Steven N. Durlauf & Lawrence E. Blume eds., 2d ed. 2008).

pears in accounts of search costs,¹³² path dependency,¹³³ first-mover advantage,¹³⁴ and reputation.¹³⁵ For our purposes, the common core idea is that information costs—for example, the cost of learning whether a supplier is generally collaborative, and can therefore be counted on to act in good faith in hard times—can create barriers to replacing one counterparty with another. In a deep spot market, such replacement is costless. But in markets where learning about the quality of potential substitute suppliers and their products is time consuming and expensive, there can be significant barriers—switching costs—to exiting a relationship.

The switching costs of interest to us have two peculiar characteristics. First, the raising of barriers to exit is not simply a feature of the context or market, as in search costs, nor is it the result of a unilateral effort by one side to make it difficult for others to exit, as with first-mover advantages. These switching costs are produced, rather, by a joint effort of the parties—an instrumental effort that is central to the structure of the relationship. As the Deere–Stanadyne contract discussed above illustrates, the parties make large investments in relation-specific information concerning each other's capabilities, investments that would be lost if the relationship terminated and that would have to be duplicated with any new supplier. Second, these investments are made gradually, rather than

132. Search costs are the costs associated with finding out what other parties offer competitive goods and at what price. E.g., Alan Schwartz & Louis L. Wilde, *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, 127 U. Pa. L. Rev. 630, 635, 641–46 (1979) (defining search costs and describing how switching costs interact with search costs to influence consumer behavior in sequential search model).

133. In a path dependent environment, factors such as increasing returns and network externalities result in an equilibrium that may not be the “most” efficient. Initial conditions, determined by serendipity or factors traditionally viewed as noneconomic, such as culture or politics, can move the system down a particular path. Later on, moving off that path—switching—to a better position may be extremely difficult because of large transition costs. W. Brian Arthur provides an accessible survey of the concept. See W. Brian Arthur, *Positive Feedbacks in the Economy*, *Sci. Am.*, Feb. 1990, at 92, 92–99. For discussions in an institutional context, see Michael Klausner, *Corporations, Corporate Law, and Networks of Contracts*, 81 Va. L. Rev. 757, 763–64, 772–825 (1995) (showing path-dependence effect of network externalities on corporate contract terms); S.J. Leibowitz & Stephen E. Margolis, *Path Dependence, Lock-In, and History*, 11 J.L. Econ. & Org. 205, 206–23 (1995) (defining and distinguishing three forms of path dependence, illustrated by video-recorder market); Mark J. Roe, *Commentary, Chaos and Evolution in Law and Economics*, 109 Harv. L. Rev. 641, 646–53 (1996) (explaining how “weak-form,” “semi-strong form,” and “strong-form” path dependence differ and how they affect institutional form).

134. See, e.g., Carl Shapiro & Hal R. Varian, *Information Rules: A Strategic Guide to the Network Economy* 168–69 (1999) (explaining how first-mover advantage can be exploited to increase customer switching costs, thus “mak[ing] rival entry unattractive”).

135. In circumstances where third parties cannot perfectly observe the conduct of parties to a contract, one party exiting the relationship may raise questions about that party's reliability that require investment by potential contracting parties in order to answer. Thus, that a party risks reputation costs by exiting is simply a particular information cost associated with switching.

all at the outset of the relationship. In effect, the switching costs in the contracts we analyze here increase in step with the learning generated by the project. In this sense they are inherent in the collaborative effort.

Precisely how do switching costs constrain opportunism during the collaborative/learning phase of the contract? Recall that the parties' efforts are not contractible because of the need for continuous mutual adaptation as collaboration resolves uncertainty one step at a time. We have seen that the formal mechanisms of our three contracts induced the parties to develop information about their respective capabilities and to develop the human capital necessary for successful collaboration. But the flip side of this reciprocal learning—and the coordinate development of collaborative skills—is that each party successively raises its switching costs: The longer the interaction, the more each party knows about the other, the higher the switching costs, and the greater the constraint on opportunism. Switching costs may also be increased by the industry context; they rise as the technology race in a given industry intensifies and as the delay associated with reproducing the information necessary to work effectively with a different supplier or customer becomes more costly. Thus, both contract and context make backing out costly for one or both of the parties. In many environments, switching costs or barriers to exit are understood as unfortunate frictions that undermine access to competitive alternatives. In the cases we examine, however, switching costs result from deliberate contractual choices designed to structure the parties' relationship efficiently.

Consider the Deere–Stanadyne agreement. As we have already stressed, two elements of the relationship serve to build switching costs as the parties invest in information about each other's capabilities for innovation and cooperation, and about each other's dispute resolution style.¹³⁶ First, with the passage of time, Deere's products and the parts being supplied by Stanadyne evolve in response to changes in the market, changes in technology, and the parties' joint innovation. The differences between the initial product (and the process by which it was produced) and the current one is thus a measure of the growing size of the switching costs resulting from the parties' investments in learning about each other. Second, the structure of the Achieving Excellence program, which contemplates the supplier's rise in status with the passage of time and the growth of Deere's experience, provides a formal parallel process: rank in the hierarchy of supplier classifications increases as switching costs grow.

The Apple–SCI agreement reflects a similar role for switching costs as a constraint on opportunistic claiming of too much of the gain from collaborative innovation. To the extent that Apple's other contractual suppliers could not immediately expand their production or were less effective at collaborative innovation, adding a new supplier to replace SCI would be costly for Apple for three reasons: (1) the delay in securing

136. See *supra* text accompanying notes 77–83.

another supplier given that the specifications and production technology had been co-engineered with SCI, and necessarily reflect the idiosyncrasies of SCI's capabilities; (2) Apple's lack of knowledge about the capabilities of potential new suppliers that it had concerning SCI; and (3) Apple's need to learn about the way potential replacements handled disputes, where such knowledge can only develop iteratively over time. Even expanding the production assigned to an existing supplier is costly because, by reducing the number of suppliers, it gives the remaining suppliers greater leverage and reduces the number of partners in innovation.

As each party's switching costs rise, the relation between them more closely approximates a bilateral monopoly. At the limit, it resembles a purely competitive, zero-sum chicken game: Even though both parties have high switching costs, there is a temptation to bluff defection so as to secure a larger portion of the *ex post* surplus. This is where the dispute resolution process becomes salient. It provides the opportunity for an iterative learning process concerning each party's dispute resolution style—whether they are cooperative or competitive (as in a chicken game). To the extent that styles are not greatly malleable¹³⁷—that is, people (and institutions) do not easily switch from cooperators to defectors—then the explicit contract provisions covering dispute resolution serve the same learning function along this dimension as the explicit contract provisions governing the staffing, budget, and interaction of the collaborative innovation effort.

In sum, our exemplar contracts systematically generate symmetrical investments in information through governance mechanisms designed to create knowledge both about the product and about each other's capacity to cooperate in problem-solving and in dispute resolution. A plausible hypothesis, therefore, is that these contracts are self-enforcing in the standard fashion: either the parties are relying on reputational sanctions (also a form of switching costs) or on the overhang of future interactions (cheating by a party in one round will be punished by the other in the next) to make their promises credible. Indeed, the analogous practice of firms issuing legally unenforceable "comfort letters" to prospective lenders has been explained as a reputational signal that makes the agreements self-enforcing by establishing the party's obligations even if there is no formal enforcement mechanism.¹³⁸ But the contracts in our group do not square easily with this common understanding of the domain of self-enforcing agreements. For example, the Deere–Stanadyne and the

137. The behavioral literature suggests that these preferences are quite stable. Robert E. Scott, *A Theory of Self-Enforcing Indefinite Agreements*, 103 *Colum. L. Rev.* 1641, 1663–65 (2003) [hereinafter Scott, *Indefinite Agreements*].

138. Rene Sacasas & Don Wiesner, *Comfort Letters: The Legal and Business Implications*, 104 *Banking L.J.* 313, 329 (1987) ("Legally vague promises and inferences from cautious language are not always valueless in business Custom shows that memorializing even a weak legal commitment carries some moral and business weight. The letter can be shown to others, and reputations can be injured by the writer's breach of faith.").

Apple-SCI agreements are plainly intended to be long-term supply arrangements, but the source of the constraints on opportunism as the supply relationship goes forward is not readily apparent.

First, reputational sanctions—that is, that potential replacement suppliers will be more skeptical of doing business with an opportunistic buyer—depend on assumptions concerning the observability of the buyer's conduct to potential suppliers which rest, in turn, on conditions that are difficult to sustain.¹³⁹ These conditions may not always be met or may sometimes be only partially met. Indeed, reputational sanctions are best seen as a special, and especially context sensitive, case of switching costs. In fact, the transactions represented by our group of co-design contracts are, for the most part, transactions in heterogeneous markets where reputational constraints are thought to be quite weak.¹⁴⁰ In such an environment, reputation alone is an inadequate means of credibly enforcing promises. Even if others can observe the contractual failure, it would be difficult to learn the true reasons for the transactional breakdown. Absent much of the information the parties can learn from iterative cooperation, the mere fact of breakdown is not sufficient to impose a reputational cost on either party. If informal (or relational) enforcement is to be a satisfactory explanation for this innovative collaboration, its domain therefore must be extended to encompass switching costs.

Second, the contracts do not themselves try to detail the future relationship between the parties beyond the initial arrangement. Thus, the expectation of a long-term relationship does not appear to rest on the discipline of repeated dealings, the standard foundation for relational contracts that do not depend on reputational sanctions.¹⁴¹ More precisely, the contracts do not seem to be a tit-for-tat solution to a multi-

139. Gilson, *Venture Capital*, supra note 69, at 1086.

140. Reputations are most effective in small homogeneous communities, where noncooperative behavior quickly becomes common knowledge and sanctions against such behavior can be effectively imposed. See, e.g., Avner Greif, *Informal Contract Enforcement: Lessons from Medieval Trade*, in 2 *The New Palgrave Dictionary of Economics and the Law* 287, 288 (Peter Newman ed., 1998); Janet T. Landa, *A Theory of the Ethnically Homogeneous Middleman Group: An Institutional Alternative to Contract Law*, 10 *J. Legal Stud.* 349, 359 (1981) (discussing ethnic community ties). Reputational sanctions can also be effective in industries with established trade associations that can both identify the bad behavior and coordinate sanctions. See Lisa Bernstein, *Private Commercial Law in the Cotton Industry: Creating Cooperation Through Rules, Norms, and Institutions*, 99 *Mich. L. Rev.* 1724, 1745–54, 1762–87 (2001) (discussing existence, importance, and supporting institutional conditions of reputation-based nonlegal sanctions in cotton industry); cf. Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms*, 144 *U. Pa. L. Rev.* 1765, 1771–77, 1787–95, 1815–20 (1996) (concluding that formalist contract enforcement by private system of merchant courts established by National Grain and Feed Association results from recognition of, and to support, role of extralegal, reputational sanctions).

141. Neither contract contains the sort of explicit price and quantity adjustment provisions, common to long-term contracts, which are designed to deal with ex post opportunism in the face of either exogenous events or attempts to take advantage of the other party's high cost of changing to another supplier/customer. In the absence of these

round prisoners' dilemma in the form of a supply contract.¹⁴² Rather, the contract structure is designed to support collaboration and co-design, not cooperation and retaliation. The prisoner's dilemma game would be quite noisy: Cheating would be hard to detect because of the continual uncertainty arising from the parties' ongoing innovation. Thus, neither of the standard self-enforcement mechanisms fits our contract exemplars.

b. *Switching Costs as Screens for Reciprocity*. — A question we have posed earlier recurs: Why do these contracts contain elaborate procedural mechanisms to govern environments where the speed of adaptation renders largely ineffective any recourse to standard legal enforcement? One hypothesis is that these agreements, in addition to being designed for parties to learn about each other's competence and about market conditions, are designed to allow parties to learn about each other's preferences to behave reciprocally and to cooperate in resolving disputes.

The experimental behavioral literature shows that contracting parties are heterogeneous regarding their preferences for cooperation. A significant percentage of transactors generally reward cooperation and punish defection, but an equally large percentage generally act in pure self-interest and behave opportunistically.¹⁴³ Thus, the challenge is to

terms, instances of iterated cooperative adjustment would be unlikely to emerge spontaneously.

142. See generally Robert Axelrod, *The Evolution of Cooperation* 27–54 (1984) (explaining success of “tit-for-tat” strategy—e.g., matching the cooperate/defect decision made by one's opponent in the previous round—in multi-round prisoner dilemma “tournaments”).

143. The substantial body of experimental evidence on people's propensity to reciprocate yields two key findings. First, many people respond cooperatively to generous acts and, conversely, punish noncooperative behavior. Such people exhibit reciprocal behavior that varies from that which a purely self-interested person would exhibit. Second, reciprocal propensity differs in degree from person to person: Some people exhibit reciprocal behavior, and others exhibit self-interested behavior. Taking all the experiments together—gathered from diverse countries and cultures—the fraction of reciprocating subjects ranges from forty to sixty percent, as does the fraction of self-interested subjects. For discussion, see Ernst Fehr & Simon Gächter, *Fairness and Retaliation: The Economics of Reciprocity*, *J. Econ. Persp.*, Summer 2000, at 159, 162 [hereinafter Fehr & Gächter, *Fairness and Retaliation*] (“Many studies have carried out detailed analyses of individual decisions and found that the fraction of subjects exhibiting reciprocal choices is between forty and sixty-six percent.”); Ernst Fehr et al., *Reciprocity as a Contract Enforcement Device: Experimental Evidence*, 65 *Econometrica* 833, 850 (1997) (finding roughly half of subjects punishing shirkers, and roughly half rewarding non-shirkers); Ernst Fehr & Klaus M. Schmidt, *A Theory of Fairness, Competition, and Cooperation*, 114 *Q.J. Econ.* 817, 818 (1999) (suggesting that observed behaviors can be explained by assuming that some actors are self-interested, while some are motivated by fairness); Matthew Rabin, *Incorporating Fairness into Game Theory and Economics*, 83 *Am. Econ. Rev.* 1281, 1283 (1993) (describing experiment showing contribution rate at forty to sixty percent of socially optimal level in one-shot public-goods-decision games). For applications of this experimental evidence to contract and international law, see Robert E. Scott & Paul B. Stephan, *The Limits of Leviathan: Contract Theory and the Enforcement of International Law* 88–94, 122–27 (Cambridge 2006) [hereinafter Scott & Stephan, *Leviathan*]; Robert E. Scott & Paul B. Stephan, *Self-Enforcing International*

identify who are the cooperators and who are likely to behave opportunistically.¹⁴⁴

The willingness to write a contract that conditions on nonverifiable procedural factors is not, by itself, a reliable signal that either party is a cooperator. But the contracts in our group have an additional feature: The procedural mechanism that generates information about the feasibility of the project itself also creates opportunities to cooperate at early stages of the relationship. The mechanism therefore also renders observable the parties' characters with respect to dispute resolution, at a time when specific investments are modest and switching costs are low.¹⁴⁵ Early reciprocation, which serves to gradually lock the parties into a relationship, signals that the parties intend to cooperate and provides information about their types.

The formal governance structure established in these contracts functions in two ways to identify a party's type. First, the obligation to share information in the collaborative/learning phase of the contract provides opportunities to observe the behavior of the other in response to opportunities to reciprocate. This gives each party the opportunity to acquire knowledge of the other's propensities.¹⁴⁶ To be sure, some parties may attempt to "act reciprocally" during the initial collaborative process only to turn to opportunistic behavior when the investments are much greater. But that problem is mitigated by the second way in which the formal governance structure signals a party's type. In addition to observation, the collaborative learning process serves to separate in time the opportunity to reciprocate from the end-stage transaction that is contemplated. It thus forces the expenditure of time for the purposes of communication, thereby increasing the cost of switching.¹⁴⁷ In this case, the transactors are not only subject to observation, but the parties must spend considera-

Agreements and the Limits of Coercion, 2004 Wis. L. Rev. 551, 564–68, 582–96; Scott, Indefinite Agreements, *supra* note 137, at 1661–75.

144. To be sure, the experimental evidence relates to the behavior of individuals and not to the behavior of firms. Nevertheless, the collaborative process that these contracts stimulate is undertaken by dedicated teams of individuals who are personally invested in the success of the co-design initiative. Thus, if the behavioral literature on reciprocity is at all relevant to institutions, it would be in the contracting environments we are describing.

145. There is anecdotal evidence that managers believe that preferences for reciprocity are stable and can be observed even in institutional settings. John McMillan and Christopher Woodruff report for example that one manager believed that "[p]eople show their personality through their actions in difficult times," so after a few transactions this manager selected a few customers to concentrate on." McMillan & Woodruff, *supra* note 128, at 2432.

146. See, e.g., George A. Akerlof, The Market for "Lemons": Quality Uncertainty and the Market Mechanism, 84 Q.J. Econ. 488, 497–99 (1970) (discussing how creditors in India use knowledge of debtors' propensities gained from repeated dealings to reduce credit risk).

147. See generally A. Michael Spence, Time and Communication in Economic and Social Interaction, 87 Q.J. Econ. 651, 651 (1973) ("[T]he fact that [communication] takes time . . . imparts signaling power to it").

ble time and effort in executing portions of an agreement, time and effort that is effectively nonverifiable and thus can only be self-enforcing. Because cooperative parties can internalize the returns to general information about their type through an enhanced reputation for cooperation, they are more willing to spend resources, like time and effort, to provide that information.¹⁴⁸ The expenditure of time performing a nonverifiable promise is itself a signal: Because these costs make it difficult for a competitor who lacks the characteristic nonetheless to imitate the behavior, it may signal a preference for reciprocity.

In sum, our examples of co-design contracts suggest that the widespread use of elaborate, nonverifiable (and therefore unenforceable) governance obligations may be in part a function of their properties as screens for the parties' willingness to engage in cooperative behavior. Knowledge of the other parties' propensities, and therefore the cost of acquiring equivalent information about any replacement, increases in parallel with iterative performance. This period of building up switching costs ranges in our exemplar contracts from fifteen months in the Warner-Lambert-Ligand contract to three years in the Apple-SCI contract and potentially much longer in the Deere-Stanadyne contract. It is the unique and innovative feature of these contracts: They use the process of collaboration to generate new information in two ways—to create the innovative product and to bind each party to the other in a process of symmetrical investments so that neither one has a hold-up advantage over the other at any point in time. To be sure, the parties are then in a bilateral monopoly and run the risk of playing chicken games in the predictable event of disagreement. But this risk is reduced by the fact that the iterated co-design process also contains a mechanism that allows early exit, for example, Deere's right not to purchase any product at all from a particular supplier and Warner-Lambert's right to elect early termination of the project.

To summarize the argument to this point: A contract that requires relation-specific investments *ex ante* but where the resolution of uncertainty will require *ex post* renegotiation will not work unless opportunism is effectively constrained. Assigning *ex post* decision rights *ex ante* will not work in the face of ongoing collaboration and uncertainty—one cannot tell ahead of time to whom decision rights should be given. The mutual raising of switching costs will constrain opportunism even in the face of uncertainty. One function of the elaborate governance and dispute

148. Joseph E. Stiglitz, *The Theory of "Screening," Education, and the Distribution of Income*, 65 *Am. Econ. Rev.* 283, 287 (1975). Another way of expressing this point is that an opportunist can perhaps dupe a single contracting partner and capture a larger surplus, but, once his type is revealed, he is less able to replicate the transaction at low cost. It must be stressed, however, that establishing or eroding a reputation requires that the relevant behavior (or a credible signal of it) be observable not only to one's existing trading partners, but also to potential trading partners, a more complicated issue. Ronald J. Gilson, *Controlling Family Shareholders in Developing Countries: Anchoring Relational Exchange*, 60 *Stan. L. Rev.* 633, 638 (2007).

resolution mechanisms in contracting for innovation is precisely to facilitate this mutual raising of switching costs: The contracts provide a screening and learning process about the parties' propensities to behave opportunistically that would be costly to duplicate with a new partner.¹⁴⁹ Thus, we see a braiding of explicit and implicit contracting that supports a co-design contract: explicit provisions that create knowledge and routines that raise switching costs and a dispute resolution mechanism that builds mutual knowledge of the propensity to reciprocate and deters behavior that could undermine the cooperative equilibrium. In this way, the collaborative mechanism that produces the information necessary to the project's success also provides the constraint on opportunism that allows collaboration—and innovation—to continue.

3. *Detering Opportunism at the End Stage: Dividing the Pie with Options.*— Thus far, we have focused on situations where the relationship potentially is long term, uncertainty continues, and there is no final period in which collaboration has ended. However, in some contracts that contemplate collaborative innovation, the period for collaboration has a predictable end, at which time the potential for opportunism reappears because the parties face a division of the surplus created by the collaboration. Here, the contractual response that supports collaborative innovation by constraining opportunism takes a different form.

If the parties' relationship moves from a collaboration stage to an end stage, the governance arrangements that operated to support collaboration no longer function. Recall that the relational governance structure worked because information about the collaborative process was symmetric and the parties could collectively discipline opportunism. But once the period of collaboration ends, the potential for opportunism reemerges. At this point, however, uncertainty is resolved and decision and control rights are contractible *ex post*—they can be assigned by explicit contract. The exemplar contracts suggest, therefore, that parties may have learned to guard against the risk of opportunism by writing contracts with two sets of terms: an implicit (and flexible) set of terms that support self-enforcement where rights and obligations are not reasonably contractible, and an explicit and precise set of terms for legal enforcement where decision rights are contractible.

149. McMillan and Woodruff make a similar argument in discussing relational contracting in underdeveloped markets with poor legal systems:

Bilateral cooperation may evolve in relationships via a process in which the potential loss from having a trading partner defect is kept small in initial transactions and allowed to increase as the relationship progresses. . . . Early in the relationship, the two sides of the trading relationship test each other. As trading continues, experience with the trading partner provides information. . . . The data show that longer-lasting relationships involve significantly more trust. After two years of dealings, the amount of trade credit offered is on average fourteen percentage points higher than at the start of the relationship.

McMillan & Woodruff, *supra* note 128, at 2432.

The Warner-Lambert–Ligand contract provides an illustration of the problem and a response. As we have seen, Warner-Lambert has an incentive to insist *ex post* on lowering the royalty payments to Ligand once it learns that the drug developed is promising.¹⁵⁰ (In contrast, in the Apple–SCI contract, the pricing terms are subject to review after the production plan is executed, but it is expected that cooperative innovation will continue so that high switching costs continue to constrain opportunism. The Deere–Stanadyne contract operates in a fashion similar to the Apple–SCI contract.) But once the Warner-Lambert–Ligand collaboration process ends and Ligand delivers a compound, Warner-Lambert no longer requires Ligand’s cooperation, so switching costs are no longer relevant; a different mechanism for constraining opportunism is necessary to support the overall collaborative innovation. In this situation, the *ex ante* contract can supplement the high switching costs (including reputation effects) operative during the cooperative phase with complementary explicit provisions that assign designated decision rights to a specific party during the noncooperative phase.

The Warner-Lambert–Ligand contract supplements switching costs by using nested explicit options—the allocation of sequential *ex post* decision rights among the parties to deal with opportunism in the postcooperative phase. The structure of the contract gives Warner-Lambert an initial option, exercisable after the development period ends, to extend the contract for another three years to develop the “lead collaboration compounds” identified by Ligand during the prior fifteen months. If Warner-Lambert exercises its option, it effectively owns the rights to the lead collaboration compounds, subject to a hierarchy of license and royalty payments set out in the original agreement. This is a point at which Warner-Lambert can act opportunistically *ex post*. Since Ligand cannot further develop a lead collaboration compound itself, Warner-Lambert can threaten to exercise the option to extend, but only at a lower royalty rate. To be sure, the *ex ante* royalty arrangement gives Warner-Lambert a substantial economic incentive to exercise its option to develop a promising project. But the typical “big pharma” has a much better capacity to sustain a delay in developing a promising project than does an undercapitalized biotech.¹⁵¹ Warner-Lambert’s reputation with other potential partners might be impaired, as we argued above, a form

150. See *supra* text accompanying notes 106–109.

151. Because Warner-Lambert will receive the great majority of the proceeds from the commercialization of a compound, one can argue that Ligand will not need to fear opportunistic renegotiation of the royalty rate. The problem, however, is that the proceeds from the compound are likely a small part of Warner-Lambert’s revenues, but a large part of Ligand’s revenues, which affects the parties’ bargaining power. Additionally, a reduction in the royalty rate is attractive to Warner-Lambert because it will be pure profit. Finally, the cost of delay may be of less significance to Warner-Lambert. First, so long as the delay does not reduce the length of patent protection, the cost is only a present value issue. Second, so long as Warner-Lambert’s commercialization pipeline does not have an infinite capacity—that is, that other projects compete for space in it—the cost of delay is

of switching cost that may continue into the end stage, but this depends on the motivation for Warner-Lambert's renegotiation being observable to other firms with whom Warner-Lambert might work in the future.¹⁵² Given the asymmetry in outside options, the parties have differing threat points, and the resulting chicken game would appear to favor Warner-Lambert over Ligand. What explicit contract terms guard against this opportunistic threat to renegotiate?

This problem cannot be addressed simply by linking the royalty to the particular results of the co-design collaborative process. The very point of the process is that the parties cannot detail *ex ante* the different states of the world that the collaborative innovation may reveal, so that one cannot write even a rough state-contingent contract linking royalties to outcomes (other than that the rates are, in the end, applied to sales). A different explicit technique is necessary. In somewhat analogous situations like venture capital contracting and movie development—which also involve innovation without the ability *ex ante* to condition shifting payouts on outcomes—the problem is addressed through nested options.¹⁵³ This appears to be the technique used in the exemplar contracts we observe.

The Warner-Lambert–Ligand contract creates an explicit nested options mechanism that prevents opportunistic renegotiation at the end stage. If Warner-Lambert does not extend, then the rights to compounds developed by Ligand employees remain with Ligand. Of course, other large pharmaceutical companies to whom Ligand then might market the compound would be concerned over the signal given by Warner-Lambert's decision not to exercise its option to extend. However, by this point in the project, the commercial promise of the compounds identified by Ligand likely is observable to other pharmaceutical companies. Thus, if the observability assumption is correct, and if Warner-Lambert seeks to renegotiate the license and royalty fees opportunistically, then Ligand can market the compound to other pharmaceutical companies if Warner-Lambert declines to extend.

Of course, the potential for opportunistic renegotiation by Warner-Lambert does not disappear when it exercises its option to extend. It can renegotiate at any point until the compound clears the FDA—Warner-Lambert can threaten not to take the compound through animal trials, clinical trials, etc., unless Ligand agrees to lower the licensing and royalty

only the difference, if any, between the value of the Ligand developed project and the value of the project that replaces the Ligand project in the queue.

152. See *supra* note 135 (discussing reputation cost as switching cost); *supra* note 139 and accompanying text (discussing importance of observability to effectiveness of reputation cost as switching cost).

153. See Gilson, *Venture Capital*, *supra* note 69, at 1073–74, 1078–81 (describing use of nested options in venture capital contracting); Victor P. Goldberg, *The Net Profits Puzzle*, 97 *Colum. L. Rev.* 524, 538–42 (1997) (describing use of nested options in movie development).

rates. Nevertheless, at this stage as well, the explicit contract structure addresses the problem. Warner-Lambert has an option to abandon development of a compound during the extension period, but the rights to that compound then revert to Ligand. Since the commercial promise of the compound will be more transparent with every post-extension stage, Ligand's ability to market the compound to Warner-Lambert's competitors will increase accordingly, serving as a growing check on opportunistic renegotiations.

C. Risks and Rewards of Braiding Explicit and Implicit Contracts

What are the consequences of a contractual innovation where parties write contracts that contain some terms that are self-enforcing and others designed for legal enforcement? The answer to this question depends on the possible effects of the alternative means of enforcement on each other, and a growing experimental literature is suggestive of the character of the interaction. First, the experimental evidence suggests that the various means of self-enforcement—retaliatory threats, reputational sanctions, and reciprocity—complement each other. For example, experiments have compared the effort levels of subjects given a single, anonymous opportunity to respond to a generous offer with the effort levels in a similar game in which repeated interactions created an additional opportunity to retaliate against selfish behavior.¹⁵⁴ The results show that repeated interactions cause a significant increase in the effort levels of the subjects.¹⁵⁵

154. See Martin Brown et al., Contractual Incompleteness and the Nature of Market Interactions 2 (Inst. for Empirical Research in Econ., Univ. of Zurich, Working Paper No. 38, 2002), available at <http://www.iew.unizh.ch/wp/iewwp038.pdf> (on file with the *Columbia Law Review*) (comparing effort levels of agents in one-shot gift exchange game with effort levels in similar game in which repeated interaction was an additional material outcome).

155. *Id.* at 27–31. For other experimental evidence bearing on the character of repeated interactions, see, e.g., Gary Charness, Responsibility and Effort in an Experimental Labor Market, 42 J. Econ. Behav. & Org. 375, 380–81 (2000) (finding third-party, rather than random, determination of wages in employment relationship diminishes effort level of employee by permitting employee to share some amount of responsibility, for final outcome, with human, third party rather than having only himself to bear full responsibility); Ernst Fehr et al., Gift Exchange and Reciprocity in Competitive Experimental Markets, 42 Eur. Econ. Rev. 1, 11–22 (1998) (finding that “opportunities for reciprocity may significantly alter market outcomes”); Ernst Fehr & Armin Falk, Wage Rigidity in a Competitive Incomplete Contract Market, 107 J. Pol. Econ. 106, 117–31 (1999) (finding that, in labor markets characterized by incomplete employment contracts, employers behave reciprocally, when unemployment is high, by refusing to cut wages); Simon Gächter & Armin Falk, Reputation and Reciprocity: Consequences for the Labour Relation, 104 Scandinavian J. Econ. 1, 3, 7–17 (2002) (finding that, while reciprocity significantly affects outcomes of both one-shot and repeated games, effect in repeated games is stronger because relational settings constrain behavior of otherwise selfish actors, causing them to imitate reciprocal behavior); Ernst Fehr & Klaus M. Schmidt, Theories of Fairness and Reciprocity—Evidence and Economic Applications 4–9 (Ctr. Econ. Stud. & Ifo Inst. Econ. Research, Working Paper No. 403, 2000), available at <http://ssrn.com/>

This result makes sense. Informal sanctions are imposed implicitly and *ex post*. Thus, for example, a cooperator can punish a shirker's defection after the fact without risking offense by announcing in advance that there will be a sanction for defection. Reciprocation also may lead to a virtuous cycle, in which engaging in cooperative behavior increases one's preference for more cooperative behavior.¹⁵⁶ Successful cooperation that generates a reputation for trustworthiness or produces returns in ongoing transactions furthers self-interest and this feedback effect may strengthen the willingness to reciprocate.¹⁵⁷

How, then, do explicit contractual obligations interact with a governance structure designed to screen for and motivate reciprocity? Here, the data indicate that absent a legally enforceable obligation, reciprocity—operating alone—generates high levels of cooperative behavior.¹⁵⁸ But the data also indicate that, once the entire relationship, including its implicit aspects, is subject to formal enforcement, voluntary reciprocity declines along with the overall level of cooperation.¹⁵⁹ These experimental results suggest that formal legal sanctions and informal sanctions based on reciprocity may well conflict with each other. In other words, formal contracting may “crowd out” behavior based on relational contracting.¹⁶⁰

A careful examination of the experimental evidence shows, however, that the crowding-out phenomenon is complex. A number of studies have confirmed the crowding-out hypothesis in interactions between individuals, where the parties must choose either informal or formal enforcement: The choice of formal enforcement uniformly suppresses reciprocity.¹⁶¹ But recent experiments show that, where there is some probability that the same buyers and sellers will continue transacting in the next period, formal enforcement that is limited only to the verifiable dimensions

abstract=255223 (on file with the *Columbia Law Review*) (reviewing experimental evidence on fairness and reciprocity).

156. Scott & Stephan, *Leviathan*, supra note 143, at 102.

157. *Id.*

158. Fehr & Gächter, *Fairness and Retaliation*, supra note 143, at 160–70.

159. Ernst Fehr & Simon Gächter, *Do Incentive Contracts Undermine Voluntary Cooperation?* 1 (Inst. for Empirical Research in Econ., Univ. of Zurich, Working Paper No. 34, 2002), available at <http://ssrn.com/abstract=313028> (on file with the *Columbia Law Review*) [hereinafter Fehr & Gächter, *Incentive Contracts*].

160. For discussion, see Scott, *Indefinite Agreements*, supra note 137, at 1688–92.

161. See, e.g., Iris Bohnet et al., *More Order with Less Law: On Contract Enforcement, Trust, and Crowding*, 95 *Am. Pol. Sci. Rev.* 131, 132 (2001) (finding that “intermediate levels of contract enforcement lead to crowding out, but low levels induce crowding in”); Bruno Frey & Reto Jegen, *Motivation Crowding Theory*, 15 *J. Econ. Survs.* 589, 596–606 (2001) (surveying empirical literature in support of motivation crowding theory); Fehr & Gächter, *Incentive Contracts*, supra note 159, at 15–30 (reporting results of study showing that “voluntary cooperation may . . . be undermined by incentive contracts”); Bruno S. Frey & Matthias Benz, *Motivation Transfer Effect* 1, 9–13 (June 12, 2001) (unpublished manuscript, on file with the *Columbia Law Review*), available at <http://www.iew.uzh.ch/home/benz/downloads/MTE.pdf> (finding “experimental evidence . . . [that] incentives can reduce intrinsic motivation even in areas *not* subject to the incentive mechanism”).

of the agreement actually enhances cooperation in those dimensions of the agreement that are nonverifiable.¹⁶²

Assuming, as before, that these experimental results regarding individual behavior also hold for behavior in the small teams at the core of the institutional designs we are examining, the data point consistently in the same direction. Explicit contracting can complement and support relational governance structures when the contracting parties deploy it to supplement the implicit enforcement mechanisms. Thus, where, as in the Warner-Lambert-Ligand contract, the explicit mechanisms are designed to legally enforce only the verifiable terms of a contract, and where the parties believe in the prospect of an ongoing relationship, the evidence suggests that explicit mechanisms designed to deter opportunism at the end stage of collaborative contracts may reinforce the patterns of trust and reciprocity, thereby enabling the parties better to enforce themselves the nonverifiable portions of the relationship.

V. COMING FULL CIRCLE: CONTRACTING FOR INNOVATION AND THE THEORY OF THE FIRM

Contracts for innovation create interfirm governance structures, which in turn can produce changes in the organization of the individual firms. The innovations these contracts produce can substantially redirect the parties' investment strategies, or lead to the creation of new firms. It is thus in the nature of these agreements to blur the distinction between contract and organization, or market and firm, that has been at the core of efforts to construct a theory of the firm since Ronald Coase's pioneering work in the 1930s.¹⁶³ This brings us back to the second gap between theory and current practice that our contracting for innovation analysis identifies. In this Part, we suggest how analysis of contracting for innovation may focus efforts to resolve key controversies regarding the organization of economic activity, controversies that the diffusion of new forms of cooperation in recent decades has made central.

A. *From One-to-One to Many-to-Many: Current Understandings of the Boundaries of the Firm*

Coase's original insight was to apply marginalist thinking to the organization of transactions, and especially to the selection of the instruments governing them. In determining whether and how to undertake the incremental or marginal transaction, Coase reasoned, economic agents face a fundamental choice. They could turn to the market for the required good or service. In that case, the costs of the transaction are the costs of contracting—finding partners, agreeing on prices, addressing the

162. See Sergio G. Lazzarini et al., Order with Some Law: Complementarity Versus Substitution of Formal and Informal Arrangements, 20 J.L. Econ. & Org. 261, 261 (2004).

163. R.H. Coase, The Nature of the Firm, 4 *Economica* 386 (1937) [hereinafter Coase, Nature of the Firm].

need to assure both *ex ante* and *ex post* efficiency, and so on. Alternatively, agents could produce the good or service internally by establishing a corresponding hierarchy within the firm. Then the costs of transacting are the costs of establishing and operating a bureaucracy.¹⁶⁴ The relative transaction costs would vary with the nature of the transaction; all else being equal, agents prefer in each case the lower cost transaction form. The nature of the firm, now understood as the kinds of transactions for which it is the lowest cost provider, would thus be revealed by the agents' calculations of the costs and benefits of substituting a property rights-based hierarchy for contract, or vice versa, in organizing the relevant transactions.

An intuitively appealing way to operationalize this general idea is to distinguish various types of transaction costs and identify patterns in their incidence that explain salient features of industrial organization. This is precisely what Williamson did.¹⁶⁵ He argued that the asset specificity of complementary investments created the risk of costly hold-ups, and that it was infeasible to mitigate that risk by drafting a complete contract covering all the contingencies that might affect the division of returns from the joint project. Hence, the costs of coordinating such investments through the market were prohibitively high and the transactions were instead organized within the firm. The result was the vertical integration widely observed up through the 1980s and that Chandler famously chronicled.¹⁶⁶

But, it turns out, Coase had anticipated and rejected on empirical grounds precisely this analysis and the generalization to which it led. As he later explained, managers were much less concerned by the prospect of hold-ups than he had speculated they might be.¹⁶⁷ If they worried about the problem at all, he reports, they thought it largely susceptible to contractual remedies. He was particularly impressed by the experience of A.O. Smith, a leader in industrial automation, based in Milwaukee, which supplied a large share of General Motors' auto frames for decades while continuing to operate as an independent company.¹⁶⁸ Most recently, Coase has gone on to argue that the canonical story connecting hold-ups and vertical integration—the final takeover of the Fisher Body Company in 1926 by General Motors—misconstrues the motives of the transaction: On the new account, General Motors' aim was to collaborate more closely with the Fisher brothers in managing innovative relations with suppli-

164. In fact, the costs of hierarchy have remained underilluminated in this strand of the literature, but the thrust of the argument is in the direction suggested here. See, e.g., Paul Milgrom & John Roberts, *The Economics of Modern Manufacturing: Technology, Strategy, and Organization*, 80 *Am. Econ. Rev.* 511, 526–27 (1990).

165. See sources cited *supra* note 1.

166. See *supra* notes 12–13 and accompanying text.

167. R.H. Coase, *The Acquisition of Fisher Body by General Motors*, 43 *J.L. & Econ.* 15, 16 (2000) (finding no hold-up in Fisher Body-General Motors relationship, previously thought to be standard example of vertical integration in response to hold-up problem).

168. *Id.* at 18.

ers—an early variant of the relations described here—rather than to protect itself against the (arguably nonexistent) threat of a hold-up.¹⁶⁹ This assessment, in turn, is consistent with Goldberg’s conclusion that the GM–Fisher Body contract, like the Deere–Stanadyne contract examined earlier, was legally unenforceable and that the parties knew it.¹⁷⁰

Although we are sympathetic to Coase’s rejection of the Fisher Body story, and of the generalization it is supposed to illustrate, his criticism is incomplete. Is the problem, as he sometimes suggests, the insufficient attention to empirical detail in transaction cost analysis?¹⁷¹ If so, what account of the pattern of transaction costs, and their connection to governance instruments, fits the facts better than the hold-up story? Or is the problem deeper, in the very effort to establish a parsimonious list of transactions inherently “suited” either to markets or to firms? And if there is no such list, how, if at all, should we think of the “nature” of the firm?

In a provocative review of the transaction cost and property rights literature on the nature of the firm, Bengt Holmström and John Roberts broached this more thoroughgoing critique a decade ago.¹⁷² Their arguments can be seen as a generalization of Coase’s objections (albeit a generalization that raises questions about the original dichotomy of market and firm). Their conclusions invite further development in light of our findings here.

Holmström and Roberts advance two claims. First, they argue that the coordination of investments, as exemplified in the threat of hold-ups, is but one of the types of governance concerns that managers must engage. Another pervasive concern regards agency problems: ensuring that agents use the discretion they are accorded to pursue the goals set by their principals, rather than pursuing their private interests. A vast and venerable literature in economics, organizational sociology, and law attests to the centrality of this concern both within the firm and between the firm and its suppliers. Yet another problem regards diffusion of knowledge relevant to improvement and innovation. Here too, a large literature in both economics and sociology documents management’s ongoing efforts to create stable, rule-based structures to ensure the circulation of knowledge which is itself too fluid and imperfectly understood—too tacit, as it is often said—to be reducible to rules.¹⁷³ Indeed, much of that literature argues that the chief function of the firm is not to solve

169. Coase, *Conduct of Economics*, *supra* note 17, at 274.

170. See *supra* text accompanying note 120. Note that this makes the GM–Fisher Body contract analogous to the Deere–Stanadyne contract.

171. Lafontaine & Slade, *supra* note 3, surveys the empirical transaction cost literature.

172. See Holmström & Roberts, *supra* note 9.

173. See, e.g., Margherita Balconi et al., *The “Codification Debate” Revisited: A Conceptual Framework to Analyze the Role of Tacit Knowledge in Economics*, 16 *Indus. & Corp. Change* 823, 825–45 (2007) (reviewing, clarifying, and reconceptualizing ideas and economic roles of tacit knowledge and codification efforts); *Managing Industrial*

hold-up problems but instead to facilitate the flow of the uncodifiable, living knowledge that animates innovation.¹⁷⁴

Holmström and Roberts's second claim goes to the relation between the kinds of governance problems and the instruments available to address them. In transaction-cost economics, at its most reductive, this relation was one-to-one: Hold-up problems can only be solved by vertical integration in the firm and the firm's boundary therefore is dictated by the breadth of hold-up problems.¹⁷⁵ The core of Coase's criticism of transaction-cost economics is that the link between the character of transactions costs and organizational form is less precise; there is more than one organizational response to particular transactions costs. The relation is, at least, one to many: Hold-up problems can be solved by contract (as in the case of A.O. Smith) as well as by vertical integration. Holmström and Roberts suggest that the relation is many-to-many: There are various governance tasks and various instruments for managing them. Each task can be addressed by more than one instrument, and each instrument can, alone or in combination with others, be used to address more than one task.

The core illustration of this many-to-many relation is Japanese-style subcontracting in the automobile industry in the mid-1990s. Holmström and Roberts note that, contrary to the predictions of transaction-cost economics,¹⁷⁶ the Japanese subcontractors (and their emulators in other countries) invest in the co-design of specialized parts and components, and in relation-specific equipment (such as expensive, metal-forming dies) despite the hold-up and other risks such investments entail. But if the organization of Japanese supply chains does not reflect recourse to vertical integration to solve the hold-up problem, neither does it reflect contract, at least as conventionally understood. The contracts between

Knowledge: Creation, Transfer and Utilization (Ikujiro Nonaka & David Teece eds., 2001) (collecting articles).

174. See, e.g., Bruce Kogut, *Joint Ventures: Theoretical and Empirical Perspectives*, 9 *Strategic Mgmt. J.* 319, 322–24 (1988) (citing transfer of organizational knowledge as motivation for joint ventures); Bruce Kogut & Udo Zander, *Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology*, 3 *Org. Sci.* 383, 383–96 (1992) (“[W]hat firms do better than markets is the sharing and transfer of the knowledge of individuals and groups within an organization.”).

175. As an example, Williamson sets out a stylized mapping of transaction form based on the presence of three forms of transaction costs: bounded rationality, opportunism, and costly information. If bounded rationality and opportunism were present, but information were not costly, then a state-contingent contract was the appropriate transaction form. If, instead, bounded rationality were still present, information were costly, but opportunism were not present, then a relational contract was the form that resulted. Finally, if all three forms of transaction costs were present, the transaction would be integrated within a firm. Williamson, *Economic Institutions*, *supra* note 1, at 18–35.

176. See generally An Interview with Oliver Williamson, 3 *J. Institutional Econ.* 373, 384 (2007) (acknowledging that transaction-cost economics requires some adjustment to explain Japanese supplier relations and declining to “hold that transaction cost economics is the only interpretation”).

the suppliers and their customers, Holmström and Roberts write, “are short and remarkably imprecise, essentially committing the parties only to work together to resolve difficulties as they emerge,” and to renegotiate prices regularly.¹⁷⁷ Rather, the governing mechanism is the “long-term, repeated nature of the interaction,”¹⁷⁸ based on shared understandings and expectations, a variant of Lamoreaux, Raff, and Temin’s relational view of the new economy. This shared experience is reinforced, at least in some cases, by supplier associations capable of facilitating the imposition of reputational sanctions on powerful customers who neglect the obligations of reciprocity. But Holmström and Roberts also observe that within these long-term relations the auto companies “carefully monitor supplier behavior—including cost reduction, quality levels and improvements, general cooperativeness, and so on.”¹⁷⁹ Thus, long-term relations of the Japanese kind are (along with vertical integration and contract) one of several possible responses to hold-up problems, and organizational structure can comprehend all these responses, as well as address other incentive or governance issues such as agency problems and the management of information flows.

Holmström and Roberts do not discuss the general relation of firms to markets in light of the many-to-many mapping they find between contemporary governance tasks and instruments. But it is clear that the original dichotomy of hierarchy and contract cannot survive the proliferation of cases and relations they report. For one thing, there are no governance tasks done only by firms or only by the market. Vertically integrated firms compete with firms that secure inputs or distribution by contracting with independent suppliers or distributors. For another, there are entire classes of governance mechanisms that simply cannot be usefully categorized as either hierarchy or contract. Information flows in both firms and markets are often managed neither by hierarchy nor contract, but rather by federated structures or peer-to-peer networks, such as supplier clubs or benchmarking groups. Seen in this light, the essence or nature of the firm is not to solve this or that governance problem. The firm does not in this sense have an essence or nature: It bundles governance instruments as the calculus of advantage in particular contexts suggests, and retains that form as a result of path dependency even as changed circumstances cause new competitors to adopt different arrangements.

177. See Holmström & Roberts, *supra* note 9, at 81. Note that this formulation of Japanese contracting almost exactly parallels Williamson’s description of contracting in a world where the parties will not behave opportunistically: “A general clause, to which both parties would agree, to the effect that ‘I will behave responsibly rather than seek individual advantage when an occasion to adapt arises,’ would, in the absence of opportunism, suffice.” Williamson, *Transaction-Cost Economics*, *supra* note 1, at 241.

178. Holmström & Roberts, *supra* note 9, at 81.

179. *Id.* at 82.

B. Contracting for Innovation and the Boundaries of the Firm

Our findings corroborate, help complete, and prompt extension of those of Holmström and Roberts. First, contracting for innovation can be thought of as a further development of the many-to-many relation of governance instruments to tasks. Instead of matching several instruments to several problems, and vice versa, contracting for innovation creates a single, novel regime that fuses and transforms elements of contract, bilateral governance, and hierarchical management. It uses this regime to coordinate investment, resolve agency problems, and direct information flows in a context in which the skills necessary for product development cannot be cabined within a single firm. The key innovation is creating a regime in which the regular and reciprocal provision of information about each party's capacity and willingness to cooperate teaches the parties how to collaborate more effectively, binding them more tightly to imprecisely defined common projects through increased switching costs resulting from that process—or alerting them to possible breakdowns before the costs of failure in the relation become ruinous. The “short and remarkably imprecise”¹⁸⁰ contracts of the Japanese style, when braided with a governance process that supports mutual learning, become a regime that generates quite precise expectations and obligations. The tacit knowledge of innovation, often held to require the carefully controlled environment of the firm, is made explicit enough to be reviewed across organizational boundaries (even if it is far from being fully formalized) and thereby opens an entity to cutting-edge technology lodged in other entities. This regime is not a governance panacea or all-purpose tool adapted to all occasions. Rather, as we have argued, it is particularly suited to situations where parties with distinctly different but complementary capacities jointly undertake to explore, and possibly exploit, an uncertain domain, a description that predicts the prevalence of this organizational form in settings where innovation is central to success.

In speaking of contracting for innovation as a further development of many-to-many governance, we are looking backward and forward in debates about governance, and in the evolution of governance instruments themselves. First, looking backward, we mean to raise the possibility that the complex governance arrangements Holmström and Roberts observed a decade ago were a precursor to (or at least inspiration for) contracting for innovation. In particular, Japanese subcontracting and production methods were widespread and prominently discussed in two industries—automobiles and electronics—in the years before contracting for innovation emerged in both.¹⁸¹ It seems quite possible, therefore,

180. *Id.* at 81.

181. See Jeffrey H. Dyer, *Collaborative Advantage: Winning Through Extended Enterprise Supplier Networks* 169–70 (2000) (advocating “extended enterprise” strategy of creating value from teaming up with other companies in automobile industry); Toshihiro Nishiguchi, *Strategic Industrial Sourcing: The Japanese Advantage* 6 (1994) (“One important reason for the competitiveness of Japanese producers is the nature of Japanese

that firms in these industries experimented with ways to achieve “Japanese” collaboration without relying on features of Japanese culture (high trust) or institutional experience (supplier clubs) or Japanese-style corporate governance not easily transferable to their settings.¹⁸² Along the way, they would have stumbled upon the information exchange and governance methods reported here, and, having learned of the advantages of these methods in practice, they would, sooner or later, have asked a lawyer to capture their defining features in contracts with potential partners.¹⁸³ Japanese subcontracting would thus have evolved into contracting for innovation. Of course, this just-so story must be verified, or replaced by a more accurate account, and in either case there remains the problem of identifying the way in which contracting for innovation arose in the pharmaceutical industry where, to our knowledge, the Japanese production model was unlikely to have been immediately influential. But contracting for innovation came from somewhere, and this seems a plausible hypothesis.¹⁸⁴

Looking forward, a second and more fundamental point raised by the emergence of contracting for innovation is that governance instruments have evolved—and almost surely will continue to do so. Coase’s original contribution, though deeply informed by knowledge of the precise practices of his day, was intended as a contribution to a noncontextual economic theory. Its aim was “to discover why a firm emerges at all

subcontracting, which emphasizes synergistic problem solving, rather than antagonistic bargaining, between organizations.”).

182. For example, early in the comparative corporate governance debate, Masahiko Aoki argued that Japanese-style corporate governance was a critical component of Japanese-style manufacturing. See Masahiko Aoki, *Toward an Economic Model of the Japanese Firm*, 28 *J. Econ. Lit.* 1 (1990). In fact, other forms of corporate governance proved far more flexible than Aoki expected, allowing Japanese manufacturing techniques to be widely replicated in countries whose corporate governance regimes were dramatically different from Japan’s and from each other’s. See Ronald J. Gilson, *Globalizing Corporate Governance: Convergence of Form or Function*, 49 *Am. J. Comp. L.* 329, 332 (2001); Charles Sabel, *Un governed Production: An American View of the Novel Universalism of Japanese Production Methods*, in *Convergence and Persistence in Corporate Governance* 310, 311–12 (Jeffrey N. Gordon & Mark J. Roe eds., 2004).

183. For a case study of the repeated emendation of a contract between a software producer and a customer to reflect the lessons of their on-going relation, see Kyle J. Mayer & Nicholas S. Argyres, *Learning to Contract: Evidence from the Personal Computer Industry*, 15 *Org. Sci.* 394 (2004).

184. As we have seen, the contract between Fisher Body and General Motors might be considered a more remote ancestor of contracting for innovation. It provided that Fisher Body would employ “the most modern, efficient and economical methods, machinery and devices consistent with good workmanship.” Coase, *Conduct of Economics*, *supra* note 17, at 266. Without the information exchange mechanism of contracting for innovation, such a term could seem placatory. With this mechanism, the term could provide an important and reliable discipline on a partnership even if, like the Deere–Stanadyne relationship, it was a creature of an unenforceable contract. Cf. *id.* at 265–66 (suggesting that close relationship between Fisher Body and GM made deviation from efficiency requirement highly unlikely).

in a specialized exchange economy.”¹⁸⁵ Transaction-cost economics, in effect, takes as the field of analysis the era of (then) modern manufacturing: For that reason the Fisher Body story of 1926 is presumed relevant to the world of a half century later. To illustrate the limits of current explanatory schemes, Holmström and Roberts introduce a form of governance—Japanese subcontracting—that happens to be new. But the novelty of Japanese governance, and the continuing evolution of instruments revealed by the diffusion of contracting for innovation, is theoretically significant in itself. It strongly suggests that the list of governance instruments is open-ended. Or, looking backward and forward, we can say that firms and governance forms co-evolve, with adjustments to new contexts by the firms leading to innovations in governance—which in turn change the context to which firms adjust. On this view, of course, firms do not have essences. More exactly, the firm in every epoch takes the shape necessary for the most pressing of the prevailing governance problems: risk in the last century, uncertainty at the start of this one.

There follows a final observation and qualification regarding these changes in firm organization and governance instrument. In reporting on the emergence of a novel governance instrument that supports new forms of cooperation, we underscore the continuing importance of (changes in) context not to suggest that the history of industrial organization has or can come to an end. The firm, now as before, has a future, and that future will be different from its past. For example, there are countertendencies even amidst the general, current tendency to vertical disintegration; some are likely to remain marginal exceptions, others, perhaps presaging new developments, are reversions to apparently superseded forms. Thus, as cell phones become on the one hand fashion items and on the other mobile Internet portals, Nokia, a leading producer, is both (re-)integrating into manufacturing—to cut the time to market for its fashion-sensitive products—and opening its research operations to a wide range of collaborators—to scan for the innovations needed to compete in the market for mobile portals.¹⁸⁶ Even as U.S. steel firms were divesting their holdings of iron ore, Mittal (now ArcelorMittal), a multinational offshoot of an Indian steel maker, was buying ore reserves to protect itself against the possibility of worldwide limits to supply.¹⁸⁷ It would be a surprise, given the transformations of recent decades, if this shifting of firm boundaries did not produce further innovations in governance. Contracting for innovation is a new tool. The best way to understand why it was made, and what it is good for, is to look carefully at the problems it does solve—and the ones it does not.

185. Coase, *Nature of the Firm*, *supra* note 163, at 390.

186. See Yves Doz & Mikko Kosonen, *The Dynamics of Strategic Agility: Nokia's Rollercoaster Experience*, *Cal. Mgmt. Rev.*, Spring 2008, at 95, 98–108.

187. Pankaj Ghemawat & Ravi Madhavan, *Mittal Steel in 2005: Changing the Global Steel Game 11–12* (Harvard Bus. Sch. Case, 2005) (on file with the *Columbia Law Review*).

CONCLUSION

In this Article, we have offered a hypothesis concerning how entities contract for innovation across organizational boundaries as an alternative to vertical integration, and also sought to place the contracting for innovation phenomenon in a more textured account of a theory of the firm and of contract theory. The need for the elaboration of theory was demonstrated by the practices that actually have developed in industries driven by innovation, where the exigencies of doing business pushed practice ahead of theory. Our analysis was illustrated and given plausibility by three exemplars of how the practices of participants contracting for innovation illustrate the gaps in both of these areas of theory. The next step is to move from plausibility to proof, which will require both qualitative and quantitative data on parties' efforts to contract for innovation sufficient to give confidence that our account captures current practice, or to instruct us in what we have missed. That will be our next project.