european corporate governance institute

# Contracting for Innovation: Vertical Disintegration and Interfirm Collaboration

Law Working Paper N° 118/2009
December 2008

Ronald J. Gilson
Stanford Law School, Columbia Law School and ECGI
Charles F. Sabel
Columbia University - Columbia Law School
Robert E. Scott
Columbia University - Columbia Law School

© Ronald J. Gilson, Charles F. Sabel and Robert E. Scott 2008. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
https://doi.org/10.67929/ecgi-law-118-2009

european corporate governance institute

ECGI Working Paper Series in Law

# Contracting for Innovation: Vertical Disintegration and Interfirm

Working Paper N°.118/2008

December 2008

Ronald J. Gilson
Charles F. Sabel
Robert E. Scott

Forthcoming 109 Colum. L. Rev. (April 2009)

We are especially grateful to Victor Goldberg, who made very substantial contributions to this article. George Geis, Michael Klausner, Richard Langlois, Curtis Milhaupt, Richard Nelson, Daniel Raff, Alan Schwartz, Yane Svetiev and George Triantis provided helpful comments on an earlier draft. We also appreciate the support of the Kauffman Foundation.

© Ronald J. Gilson, Charles F. Sabel and Robert E. Scott 2008. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

## Abstract

Rapidly innovating industries are just not behaving the way theory expected. Conventional industrial organization theory predicts that when parties in the 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, 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 product. In doing this, the parties are developing 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 investment in transaction specific investment in knowledge about their collaborators' capacities, deter opportunism under circumstances when explicit contracting, renegotiation and the anticipation of future dealings cannot.

---

Keywords: vertical integration, contracting, switching costs, innovation

JEL Classifications: D23, G34, L14, L23

Ronald J. Gilson
Stanford Law School
559 Nathan Abbott Way
Stanford, CA 94305-8610
United States
phone: 650-723-0614, fax: 650-725-0253
e-mail: rgilson@leland.stanford.edu

Charles F. Sabel
Columbia University - Columbia Law School
435 West 116th Street
New York, NY 10027
United States
e-mail: csabel@law.columbia.edu

Robert E. Scott
Columbia University - Columbia Law School
435 West 116th Street
New York, NY 10027
United States
phone: 212-854-0072
e-mail: rscott@law.columbia.edu

# CONTRACTING FOR INNOVATION: VERTICAL DISINTEGRATION AND INTERFIRM COLLABORATION

*Ronald J. Gilson,* Charles F. Sabel** & Robert E. Scott****

Forthcoming 109 Colum. L. Rev. (April 2009)

## 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 risk of opportunism will drive them away from contracts 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 relationships; that is, where the future states of those relationships cannot be predicted probabilistically.[2] 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

*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. George Geis, Michael Klausner, Richard Langlois, Curtis Milhaupt, Richard Nelson, Daniel Raff, Alan Schwartz, Yane Svetiev and George Triantis provided helpful comments on an earlier draft. We also appreciate the support of the Kauffman Foundation.

[1] Much of this transaction-cost literature is an extension of the work of Oliver Williamson. See e.g., Oliver E. Williamson, The Economic Institutions of Capitalism (1985); Oliver E. Williamson, Markets and Hierarchies: Analysis and Antitrust Implications (1975); Oliver E. Williamson, Transaction-Cost Economics: the Governance of Contractual Relations, 22 J. L. & Econ. 233 (1979

[2] The reference here is to uncertainty in Knightian terms. See e.g., Frank H. Knight, Risk, Uncertainty and Profit (1921). In Knight's usage there is risk when alternative future states of the world occur with quantifiable probability: the future can be expressed as a probability distribution. Since we can through insurance and other means mitigate or even eliminate the effects of unfavorable states, a risky world is one in which we can with near certainty live in the conditions we choose. The Knightian distinction between risk and uncertainty is a useful way to illustrate the way accelerating technology and global competition have created unique circumstances that resist probabilistic classification. Id. at 197-232.

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 recognize that they cannot themselves maintain cutting-edge technology in every field required for the success of their product. Accordingly, companies are increasingly electing to acquire by contract components that in the past they would have made themselves. Thus, instead of vertical integration, we observe vertical disintegration in a significant number of industries. In the process of vertically disintegrating, firms are developing forms of contracting beyond the reach of contract theory models.

To explain the incursions of contract on the domain of the vertically integrated firm, some theorists emphasize "modularity," The 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. Despite its apparent benefits, however, 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.

---

[3] Francine Lafontaine & Margaret Slade, Vertical Integration and Firm Boundaries: The Evidence, 45 J. Econ. Lit. 629, 649 (2008) (hold-up problem "clearly pose[s] problems for long-term contracting, and those problems are exacerbated in volatile environments") See TAN __ infra.

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

[5] Henry Chesbrough, Towards a Dynamics of Modularity: A Cyclical Model of Technical Advance, *in* Andrea Prencipe, Andrew Davies, and Michael Hobday (eds.), The Business of Systems Integration 181 (2004).

In view of these developments, other theorists suggest that this collaborative process that replaces vertical integration is governed by relational contracting rather than modularity.[6] Here the focus is on the dominance of non-contractual social and network bonds and informal cooperation as the mechanisms that support collaboration by constraining opportunism. But neither of these polar alternatives fits the 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, together with implicit (i.e, non legally enforceable) obligations that establish formal governance structures regulating the exchange of highly revealing information, but do not necessarily impose legally enforceable obligations actually to buy and sell products. This braiding creates an interactive process that constrains opportunism as the parties' investments in detailed knowledge of their respective 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 seek to 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 that we observe in many industries is mediated neither by fully specified explicit contracts 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 changes in firm boundaries are mediated by a new form of contracting---what we call "contracting for innovation."[8] Contracting for

[6] Naomi R. Lamoreaux et al., Beyond Markets and Hierarchies: Toward a new Synthesis of American Business History, 108 Am. Hist. Rev. 404 (2003).

[7] See discussion TAN supra.

[8] Lafontaine et. al., supra note __ provide a useful survey of the theory and empirical studies of the vertical integration decision and the consequent boundary of the firm. 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

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.

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, but it is of a character somewhat different than in the conventional vertical integration literature, and 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 that illustrate 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 the need to acknowledge that there is no theory of *the* firm. Following Bengt Holmström and John

---

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 empirical studies. In most cases, this implies that we equate contracts with arms length transactions and contrast firms' decision to rely on such transactions versus vertical integration.")(citations omitted).

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 in 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 U.S. and other advanced economies has changed in two fundamental ways. First, many kinds of transactions that used to take place within firms—the manufacture of key parts or components, assembly of final products, and research leading to new ones—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 proportion 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

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

[10] The fact of vertical disintegration is not controversial. Thus, while Langlois and Lamoreaux et. al. disagree about both the causes and the response, each treats the phenomenon as fact. Richard N. Langlois, Chandler in Historical Perspective: Markets, Transaction Costs, and Other Organizational Form in History, 5 Enterprise & Society 355 (2004). ("In 1977, ... the large, vertically integrated 'Chandlerian' corporation had dominated the organizational landscape for nearly a century. ... A quarter of a century later, however, the Chandlerian firm no longer dominates the landscape."); Naomi R. Lamoreaux et. al., Against Whig History, 5 Enterprise & Society 376, 377 (2004) ("By the end of the twentieth century, it had become clear that ... the acme of capitalist development – the large, vertically integrated, horizontally diversified, managerially directed corporation – was clearly in retreat.").

organized exchanges of information designed to identify and utilize possibilities for innovation. We label this novel collaborative form: "contracting for innovation."[11]

This collaborative innovation is not just a shift from hierarchy (organization of transactions within firms) to contract. Rather, we will argue that 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 away 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 20th century, the dominant firms in industries such as steel, automobiles, electric machinery and food processing, both in the U.S. and worldwide, used the technologies of the second industrial revolution[12] to achieve dramatic economies of scale through the mass production of standard goods with single-purpose or dedicated machinery. 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 disruption to the production process. Because achieving economies of scale entailed large specific 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 D. Chandler, the pre-eminent historian of the mass-production firm, the

[11] The discussion that follows draws on Charles R. Sabel & Jonathan Zeitlin, Neither Modularity nor Relational Contracting: Inter-Firm Collaboration in the New Economy, 5 Enterprise & Society 388 (2004).

[12] Michael J. Piore & Charles F. Sabel, The Second Industrial Divide: Possibilities for Prosperity (1984).

"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 value of two independently controlled investments is 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 inter-dependent 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 dis-integration of vertical combinations in sectors where they once seemed irrevocably established, or the exploration of collaborative alternatives to full integration in domains unencumbered by their legacy. In certain economically significant sectors, fear of hold-ups, at least in their traditional form, no longer compels firms to vertically integrate.

---

[13] Alfred D. Chandler, Jr., The Visible Hand: The Managerial Revolution in American Business (1977).
[14] See note – supra.
[15] Williamson, Markets and Hierarchies, supra note 1; Williamson, The Economic Institutions of Capitalism, supra note 1.
[16] See Part II infra.

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

---

[17] See e.g., Benjamin Klein, Vertical Integration as Organizational Ownership: The Fisher Body-General Motors Relationship Revisited in The Nature of the Firm: Origins, Evolution and Development 213 (O. Williamson & S. Winter eds., 1991); Benjamin Klein, Fisher-General Motors and the Nature of the Firm, 43 J. L. & Econ. 105 (2000); Ronald Coase, The Conduct of Economics: The Example of Fisher Body and General Motors, 15 J. Econ. & Mgmt. Strategy 255 (2006).

[18] Dave Nelson et al., The Purchasing Machine: How the Top Ten Companies Use the Best Practices to Manage Their Supply Chains (2001); Joel M. Podolny & Karen L. Page, Network Forms of Organization, 24 Ann. Rev. Socio. 57, 68 (1998).

[19] Gary Gereffi, et al., The Governance of Global Value Chains, 12 Rev. Intl. Pol. Econ. 78 (2005); Timothy J. Sturgeon, Modular Production Networks: A New American Model of Industrial Organization, 11 Indus. & Corp. Change 451 (2002).

[20] Richard S. Rosenbloom & William J. Spencer, eds., Engines of Innovation: U.S. Industrial Research at the End of an Era (1996).

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

At the intersection of both these developments---dis-integration 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 subsystems—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 planes.[22] The performance of each of these subsystems depends conversely 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 areas necessary for the various components. Collaboration with groups of key technology suppliers, involving continuing mutual adjustment and exchange of quintessentially proprietary knowledge becomes the norm.[23] Thus, the

[21] Bruce Kogot et al., Interfirm Cooperation and Startup Innovation in the Biotechnology Industry, 15 Strat. Mgmt. J. 387 (1994); Walter W. Powell et al., Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology, 41 Admin. Sci. Q. (1996).
[22] Annabelle Gawer & Michael A. Cusumano, Platform Leadership: How Intel, Microsoft and Cisco Drive Industry Leadership (2002).
[23] 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. 1

recent emergence of a literature on platform industries and the forms of predation that tempt platform owners attests to the diffusion of enduring (though sometimes fraught) cooperation across firm boundaries just as the literature on modules and global value chain attests to the decline of vertical integration.[24]

The replacement across a wide range of industries of vertically integrated firms by inter-firm collaboration thus 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 has 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. 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 holdups and so induces vertical integration. As Adam Smith, thinking of the economies of scale attained in the pin factories of his day put it, "The division of labor is limited by the extent of the market."[25]

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. For present purposes, one set of developments is

---

(2003); David S. Evans, The Antitrust Economics of Multi- Sided Platform Markets, 20 Yale J. Reg. 325 (2003).

[24] David S. Evans & Richard Schmalensee, Some Economic Aspects of Antitrust Analysis in Dynamically Competitive Industries, 2 Innovation Policy and the Economy 1 (J. L. Jaffe & S. Stern eds., 2001); Jean- Charles Rochet & Jean Tirole, Two- Sided Markets: An Overview, (IDEI Working Paper, 2004) available at http://www.frbatlanta.org/filelegacydocs/ep_rochetover.pdf. David S. Evans et al., A Survey of the Economic Role of Software Platforms in Computer-Based Industries, (CESIFO Working Paper No. 1314, October 2004), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=618982 (on file with the *Columbia Law Review*).

[25] Adam Smith, The Wealth of Nations (1776).

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 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 then replayed.

This increasing unpredictability is manifest as the pervasive fear of what Clayton Christensen calls "disruptive" technologies.[26] 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 how to improve on what they already know and 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.[27] 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

[26] Clayton M. Christensen, The Innovator's Dilemma (1997).
[27] Examples of this phenomenon include electric-arc or mini-mill steel producers, hydraulically activated earth-moving equipment, or, in the realm of general production technologies—Japanese or lean production methods. Id. at ---. Christensen argues—unchallenged, so far as we know—that all established technologies are in principle disruptable in this way.

capabilities that underlie success in the industry as currently understood blind the firm to threats from outside of the dominant conception.[28]

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 of these responses—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 which is our focus in this Article.

First consider clusters or industrial districts. Clusters are geographically compact agglomerations of small and medium sized firms in an industry characterized by volatile or rapidly shifting demand, all of which specialize in a particular phase of production or a production process. Finished goods are produced by groups of firms collaborating in rapidly shifting constellations.[29] 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 18th century onwards.[30] Variants are common in more recent industrializers ranging from Japan to Taiwan to Brazil to Kenya, and in the development of Silicon Valley.[31] Since the turbulence in the markets for mass produced goods in the mid-1980s

[28] 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. See Christensen, supra note – at –.
[29] Up to some limit, the more firms in a cluster, the easier it is for each firm to find the partners it needs, 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 Paul Krugman, Increasing Returns and Economic Geography, 99 J. Pol. Econ. 483 (1991); Edgar M. Hoover and Raymond Vernon Anatomy of a Metropolis: The Changing Distribution of People and Jobs within the New York Metropolitan Region (1959).
[30] See Charles F. Sabel & Jonathan Zeitlin, World of Possibilities: Flexibility and Mass Production In Western Industrialization (1997).
[31] On Japan, see David Friedman, The Misunderstood Miracle: Industrial Development And Political Change In Japan (1988); on Silicon Valley, see AnnaLee Saxenian, Regional Advantage: Culture and Competition in Silicon Valley and Route 128 (1994), and Ronald J. Gilson, The

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.

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 U.S. after World War II to develop weapons systems at the then frontier of technological capability.[32] Since no single firm could produce, say, both the inertial guidance system and the rocket motor needed for a missile, coordination 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 elaborated in response to current conditions are common today at the (vastly extended) technical frontier.[33]

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

The preceding examples 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 the clusters is extremely fluid; the creation of new firms, the re-contracting among existing ones, and the circulation of

---

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 (2006); on Brazil and Latin America, see Elisa Giuliani et al., Upgrading in Global Value Chains: Lessons from Latin American Clusters, 33 World Develop. 549 (2005); and on Italy, see Roberta Rabellotti, Anna Carabelli & Giovanna Hirsch, Italian Industrial Districts on the Move: Where are they Going? (forthcoming European and Planning Studies, 2008).

[32] 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.

[33] See, for example, the relation between Intel, the microprocessor manufacturer, and the suppliers of its chip-making equipment. Best, The Geography of Systems Integration, in Prencipe, Davies and Hobday, The Business of Systems Integration at 209-210.

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. 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.[34] Its central claim is that the availability of the new tools of design and production allow 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 inter-firm coordination that products can be decomposed into distinct modules, all of which can be produced in virtual isolation from the others.[35] 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' 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. Langlois focuses, however, only on the opportunity for market substitutes to vertical integration; he does not address the form those arrangements actually take.[36]

---

[34] Richard N. Langlois, supra note __.
[35] Id. at 374.
[36] Langlois, Chandler in a Larger Frame, supra note __. While Langlois acknowledges that much 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. My claim is only that they are 'market' arrangements in the broad sense ...." Richard Langlois, Rejoinder, 5 Enterprise & Soc. 404, __

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 the industrial districts or clusters.[37] They see the new economy as a shift away from coordination by managerial hierarchies in vertically integrated firms towards coordination through long-term relationships, based on "informal restraints on self-interested behaviour," among networks of formally separate firms.[38] Where Langlois emphasizes the resurgence of the market over the visible hand of management, Lamoreaux, Raff and Temin emphasize the resurgence of social bonds as the underpinnings of economic relations. In the pre-Chandlerian era, they write, "business people in industrial communities interacted socially as well as economically, and the resulting multidimensional relationships facilitated cooperation for purposes besides production."[39] Relational coordination is possible again today, they argue, because cost reductions in transportation and communications allow multi-dimensional relationships to develop at a distance, which give 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."[40] Put another way, the chief advantage of formal disintegration of the firm is to create the possibility for sustained informal cooperation between independent producers, a different vector from the highly formalized exchanges stressed by Langlois.

We share Lamoreaux, Raff, and Temin's recognition that uncertainty has driven the disintegration process, and agree that relationships between firms substitute for

---

(2004). As he puts it in Langlois, *The Vanishing Hand*, supra note___, at 376, "As 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."

[37] Lamoreaux et al., Beyond Markets and Hierarchies, supra note __.

[38] Id. at —.

[39] Id. at 34.

[40] Id. at 11.

vertical integration. However, the critical point from our perspective is to understand these relationships: How is collaborative innovation organized?[41]

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. 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 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, 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.[42] Thus, modularization is not a stable strategy because it leaves open too many possibilities for competitive improvements that would cumulatively undermine the initial interfaces.[43]

---

[41] Lamoreaux, Raff and Timins 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 __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 under theorized because current contract theory provides no guide as to how that renegotiation would take place. See TAN __ infra

[42] Clayton M. Christensen, Matt Verlinden, and George Westerman, 11 Indus. & Corp. Change 955 (2002) (2002) (In the hard disk industry, only in the low-performance end of the market could the performance of finished products be predicted from the performance of their component modules)...

[43] 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

In contrast to modular interfaces specified by the standard setter, an interesting set of firms engage 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 under-specified.

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.[44]

Taken together, these iterative, cooperative techniques play an important part in shaping the links that connect firms in the vertically disintegrated economy. As each

---

architecture. This is what Chesbrough calls a "modularity trap." Henry Chesbrough, Towards a Dynamics of Modularity: A Cyclical Model of Technical Advance, in Andrea Prencipe, Andrew Davies, and Michael Hobday (eds.), The Business of Systems Integration 181 (2004). In the firm, "the focus on developing products to compete within the standard eventually erodes the amount of system-level knowledge." Id. at __.

[44] For a fuller discussion, on which this presentation draws, see Susan Helper, John Paul MacDuffie, and Charles F. Sabel, Pragmatic Collaborations: Advancing Knowledge While Controlling Opportunism, 9 Indus. & Corp. Change 443 (2000), and Charles F. Sabel, A Real-time Revolution in Routines, supra note --.

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.

Moreover, these methods also address the governance problems arising from the mutual vulnerability inherent in such open-opened collaboration among different entities. The exchanges of information required for benchmarking, simultaneous engineering, and error detection and correction increases 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, the parties learn their 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 section, 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 one another'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 their special expertise. This on-going 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.

These 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; a model that has as an end point (relatively) complete contingent contracting that specifies the relevant interfaces. On the other hand, the new supply patterns are not regulated simply by non-contractual continuing relations and tit-for-tat enforcement –as suggested by the relational models of Lamoreoux, Raff and Timins.[45] 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 on-going 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.[46] The contracts can be arrayed along a continuum, ranging from contractual relationships that impose no formally enforceable obligations on the parties but that contemplate on-going 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

[45] While we are confident that they would include contracting for innovation within their general framework, as we set out it note __ supra, the hard task is understanding the structures that we actually observe.

[46] 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.

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 spill-over effects of continuous innovation create "coordination cascades"---an innovation by one party that 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. But as theory and conventional legal practice have lagged behind the conditions in the marketplace, transactional lawyers in a number of industries apparently have begun responding to their clients' need to structure new relationships in light of the constraints imposed by uncertainty and have created the novel patterns of contracting whose characteristics we now address.

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. [47] Innovation in one structure must mesh with innovation in the other in order for either to be successful. The wing must not only be compatible with the fuselage; the two must fit. Innovation is thus a collaborative and iterative process rather than a discrete product

[47] 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 (2006).

supplied by a party upstream in the supply chain according to specifications set by a downstream customer. [48]

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, which reflect distinct patterns of collaborative production and supply. The Deere-Stanadyne contract addresses collaboration but without any product/sale obligation;[49] 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;[50] and the Warner-Lambert-Ligand contract covers the collaborative search for a product and the non-collaborative commercialization of it.[51]

As the foregoing suggests, the transactions governed by these contracts share a number of characteristics. First, the primary output is an innovative "product,"[52] 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

---

[48] Id.

[49] 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 established formal programs covering Deere's evaluation of a supplier's characteristics and performance. See TAN supra.

[50] 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 TAN supra.

[51] The Warner-Lambert-Ligand contract concerns a collaborative effort to discover and commercialize pharmaceutical products over a specified term and specifies the options at the end of collaboration if the parties' efforts result in a marketable product. See TAN supra.

[52] We use the term "product" in this section to describe a range of innovative outputs: it may result in a single product, but it also can be a stream of innovations (see e.g., the Apple-SCI, GM and Deere-Stanadyne contracts discussed infra) or intellectual property that results in a single patent or patentable things (see, e.g., the Warner-Lambert-Ligand contract discussed infra).

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 each 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*.[53]

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, specific investments,[54] 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

[53] This discussion draws on Robert E. Scott & George G. Triantis, Incomplete Contracts and the Theory of Contract Design, 56 Case W. Res. L. Rev. 187 (2005), and Robert E. Scott & George G. Triantis, Anticipating Litigation in Contract Design, 115 Yale L. J. 814 (2006).
[54] See TAN __ supra.

so changed that the result is a net loss for the parties jointly.[55] 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 happy 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.[56]

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 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

[55] 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 promise.

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

incentive to misrepresent actuality). It is costly to specify what should happen in different future states, and it is costly to prove what actually did happen. 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 promise 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.[57] Under those conditions, both parties would prefer to design their contract ex ante so as to avoid inefficient production ex post.

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.[58]

---

[57] Charles J. Goetz & Robert E. Scott, Principles of Relational Contracts, 67 Va. L. Rev. 1089 (1981).

[58] For discussion, see Oliver Hart & John Moore, Property Rights and the Nature of the Firm, 98 J. Pol. Econ. 98 (1990); Benjamin Klein, Vertical Integration as Organizational Ownership: The Fisher Body-General Motors Relationship Revisited, J. L. Econ & Org. 199 (1988).

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 results from a promisor with the discretion to adjust performance as conditions change always choosing the best alternative for him, 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.[59] 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

---

[59] Alan Schwartz & Robert E. Scott, *Contract Theory and the Limits of Contract Law*, 113 Yale L.J. 541, 601-05 (2003).

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 its 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 or "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 direct instead 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 tradeoff between the benefits of ex ante precision (with resulting ex 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. 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 and 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. The contract design problem is acute, however, because the collaborative process generates continuous uncertainty. As a result, 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).[60] 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---to perform, to alter the terms of performance, or to abandon the transaction, or whether to 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

[60] See e.g., Eli Lilly & Co. v. Emisphere Technologies, Inc, 408 F. Supp. 2d 668 (2006) ( 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); In the matter of the Arbitration between Pharmacia & Upjohn Co. v. Elan Pharmaceuticals, 781 N. Y.S. 2d 95 (2004) (contractual relationship breaks down when Pfizer acquired Pharmacia and allegedly used information transferred to the collaboration to support a "separate research program."); Static Control Components v. Mitsubishi Kagaku Imaging Corp., 2007 WL 586710 (M.D. N.C.) (plaintiff claims defendant breached confidentiality provision by selling co-developed products to third parties, and favored other competitors over plaintiff).

investment. Contract and property-rights theorists have proposed solutions to the hold-up problem that rely on regulating renegotiation so as 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 that party the rights to control key decisions or property rights in assets specific to the exchange.[61] 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 describes a new pattern of collaborative innovation in the supply chain, one characterized by multiple information flows, iterative design and 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.[62]

---

[61] Sanford J. Grossman & Oliver Hart, The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, 94 J. Pol. Econ. 691 (1986) (property ownership rights used to determine the parties' bargaining position in ex post renegotiation); Oliver Hart & John Moore, Property "Rights and the Nature of the Firm, 98 J. Pol. Econ. 1119 (1990) (same).

[62] The renegotiation models that theorists have developed have the following key assumptions: parties' efforts (i.e., investment decisions) are non-contractible 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

Rather they are contracting over the creation of something whose features—and the complementarities between those features and their own (changing) interests—emerges 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.[63]

In response to these limitations, Baker, Gibbons, and 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.[64] 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 sometime ago, "the goal is to shift the discretion to the party whose misuse of it can be most easily constrained,"[65] rather than to specify the appropriate adjustment. For example, in venture capital contracting, the decision whether to continue a project is

---

bargaining power in the ex post renegotiation to the investing party. See sources cited in note 57 infra

[63] 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 investments 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. Alan Schwartz & Robert E.. Scott, *Precontractual Liability and Preliminary Agreements*, 120 Harv. L. Rev. 661, 685-91 (2007).

[64] George Baker, Robert Gibbons & Kevin J. Murphy, Contracting for Control (mimeo 2006), available at http://www.law.columbia.edu/null/Contracts+Conf+-+April+7-8,+2006+-+Paper+-+Gibbons?exclusive=filemgr.download&file_id=961193&showthumb=0 (on file with the *Columbia Law Review*).

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

shifted to the venture capitalist through staged financing because its decision, unlike that of the entrepreneur, is policed by an effective reputation market.[66]

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 feasible technique for assuring efficient adaptation.[67] 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 offers valuable tools for explaining the contractual patterns that we observe in a small number of co-design contracts. Before doing so, 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. 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

[66] Id. at 1086.

[67] Baker et. al. solve part of this problem by assuming that there is no transaction-specific investment, 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 et. al. actually have in mind. Our analysis here generalizes the Baker et. al. approach to encompass the type of transaction specific investment inherent in collaborative innovation settings.

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 buy a manufacturing plant from the seller of a product and to continue to manufacture and supply that product for a specified term – 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 pharmaceutical products, which would then be guided through the process of regulatory approval and commercialized by only one of the parties.[68]

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 that we have 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 the real contracts: We have no reason to believe that individual lawyers and clients negotiating

[68] 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.missour.edu. These organizations, in turn, obtain most of their contracts from SEC filings on the 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 data base pursuant to a company's request for confidentiality.

and writing individual contracts will craft the efficient structure every time – even in competitive markets efficiency constraints are simply not that binding.[69] 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 Contract.*

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. What is in many ways the commercially most important part of the arrangement, however, concerns as yet unidentified future products and is established by indirection rather than by explicit provisions.

Deere manufactures and sells machinery and equipment used in agriculture, construction, and commercial-residential lawn-garden care. Stanadyne owns and operates design and manufacturing facilities for precision engine components including injection equipment.[70] 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.

[69] The contract terms meet only the weakest test of "birth," not the stronger test of "survival." See Victor P. Goldberg & John R. Erickson, Quality and Price Adjustment in Long-Term Contracts: A Case Study of Petroleum Coke, 30 J, L. & Econ. 369, 371 (1987).
[70] Article I

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 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.[71]

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.[72] 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.[73]

---

[71] Article VIII (B) provides that " If STANADYNE CORPORATION is in compliance with Section V [the Achieving Excellence program described below], all SD Program reductions realized will be shared equally between DEERE and STANADYNE CORPORATION. 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."

[72] The agreement was follow-on to a previous five-year contract that was extended for an additional five years.

[73] Article VIII (B) provides that " If STANADYNE CORPORATION is in compliance with Section V [the Achieving Excellence program described below], all SD Program reductions realized will be shared equally between DEERE and STANADYNE CORPORATION. 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."

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 depends 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 feasible, 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.

From this perspective, 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.[74] 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 Deere Achieving Excellence 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."[75] 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.[76] 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.[77] Under the program, the suppliers' performance is evaluated on a

[74] 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." Article V.
[75] Byron Black, Transforming the Global Value Chain – (200_ ).
[76] Id.
[77] Id. at —.

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 non-enforceability 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 fact finding. In contrast, Deere's Achieving Excellence program is a governance mechanism 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.

Assessment of a supplier's capabilities and character is based on the actual experience of collaboration. It therefore takes time: As a supplier moves from conditional supplier to Partner, both Deere and Stanadyne learn more each other's capabilities and character. Thus, replacing a particular supplier means that Deere would have to incur the costs of learning both the capabilities and character of a new one; making an existing relationship work is often preferable to incurring the information costs associated with a replacement.

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. Moreover, the loss increases 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. This problem has a

symmetric impact on Deere. 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 manage effectively a prior supply relationship raises questions that a new supplier would have to invest in answering.[78]

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 potential customers, a breakdown in the relationship imposes costs beyond the immediate loss of Deere's business or the supplier's product.

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 established a two-step dispute resolution process. First, executives from each firm higher up in the management than the disputing managers would meet in good faith in an attempt to negotiate a resolution.[79] If that failed, the dispute would 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 relied upon an external, pre-existing,

[78] 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 (F. Stephen ed. 1984).

[79] Article XXII (G) (1) 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 DISPUTE, and who are at a higher level of management than the persons who have direct responsibility for the administration of this Agreement.

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. *Apple-SCI Manufacturing Agreement*

The Apple-SCI agreement represents a movement in the direction of legal enforceability compared to the Deere-Stanadyne agreement. Because the Apple-SCI agreement 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 quite 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 $4 billion, twenty plants in eight countries, over 15,000 employees, and over fifty customers, including Hewlett Packard and IBM.[80]

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 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.

[80] Timothy J. Sturgeon, Turnkey Production: The Organizational Delinking of Production from Innovation, in New Product Development and Production Networks 67, 73-74 (U. Jurges ed. 2000).

The sale was said to have been triggered by Apple's inability to meet a surge in demand from its own plants.[81] The terms of the contract seem to have given Apple significant flexibility.[82] The formulation of its purchase obligation in relative terms shifted the risk of demand fluctuation to SCI.[83] 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 holdup 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.[84]

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

---

[81] Timothy J. Sturgeon, Turnkey Production Network: A New American Model of Industrial Organization?, Berkeley Roundtable on the International Economy Working Paper BRIEWP92A (1997), available at https://repositories.elib.org/brie/ BRIEWP92A.

[82] Within the constraints of the percentage volume commitment, Apple had considerable flexibility to respond to market conditions. Apple provides on a monthly basis non-binding forecasts covering the following six months and issues 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 30 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.

[83] 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.

[84] As discussed TAN __ infra, Apple has multiple sources for the products it is purchasing from SCI; that is, Apple will be getting 60 percent of its computers from suppliers other than SCI. 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.

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 acquisition 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; Apple commits 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.

From this perspective, the centerpiece of the contract was the parties' agreement to establish "product plans."[85] 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 pre-production 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 pre-production 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 pre-production review, Apple would give SCI the

[85] Article 4.

go-ahead to begin production.[86] SCI would then produce on a "turnkey" basis in accordance with Apple's specifications and quality requirements.[87]

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. Its request must include cost, scheduling, or 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 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. Nonetheless, the basic outlines are clear. For each new product SCI would propose a price quote; the formula would take into account a number of cost factors, but the basic agreement does not say how these would be weighted. On procured material, SCI would pass through any cost reductions including rebates and discounts. Failure to do so would constitute a material breach and would be grounds for immediate termination, except if the failures were de minimis or accidental and were promptly remedied, with interest. While such a response would appear 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.[88] Importantly, the pricing provisions can function as no more than a

[86] Article 5.
[87] Article 6.
[88] During the three year term the contract also contains three significant explicit remedies covering situations that presented a serious risk to Apple, and which 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." While the consequences of a failure to achieve certification or the loss of the approved supplier label are not spelled out, they

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 would iteratively determine the feasibility of innovations suggested by both parties. For this purpose, the three-year term during which Apple committed to purchase from SCI in order to support the sale of the manufacturing facility serves to facilitate the development of valuable information: Joint development of cumulative changes in product plans over the three year period would generate ever increasing knowledge about each party's capacity for collaborative innovation and good faith dispute resolution, as the products actually being produced reflected less and less of the legacy of Apple's pre-contract products and more and more the product of the collaborative effort. At the end of the three year commitment, the contract no longer bound either party; it left the parties free to go forward, unconstrained by the detailed terms of the contract.

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.

From this perspective, 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

---

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."

its collaborative manufacturers---multiple suppliers pose a risk that Apple will switch from one non-collaborative 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. *Warner-Lambert - Ligand Agreement*

The development of new drugs based on biotechnology often entails contracting across organizational boundaries.[89] 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 both to take the drugs through the arduous process of obtaining FDA approval and then to commercially market the drug.[90] We observe a broad range of collaborative arrangements including joint ventures, licensing agreements, and co-development deals.[91] We focus here on a particular contract: a research, development, and license agreement between Warner-Lambert, a large pharmaceutical company, and Ligand Pharmaceutical, a much smaller biotech company, to discover and/or design small-molecule compounds which act through the estrogen receptors, to develop pharmaceutical products from such

[89] For example, Robinson and Stuart report that some 25 percent of the $25 billion of industrial financed (as opposed to university or other non-profit research) 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).

[90] Id. at 364.

[91] See Leslie Gladstone Restaino, BioPharma Collaborative Agreements: Choosing the right Deal Structure, Metropolitan Corporate Counsel 47 (Nov. 2007), available at http://www.metrocorpcounsel.com/pdf/2007/November/47.pdf (on file with the *Columbia Law* Review).

compounds and to take such products through the FDA approval process and commercialization (the "Warner-Lambert-Ligand agreement").[92]

A brief description of the drug development process provides a context for the Warner-Lambert-Ligand agreement.[93] The initial screening of compounds and preclinical work take 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 products can move to clinical testing of the drug on humans. Clinical testing takes another six to seven years. That period is broken down into three phases: Phase one tests include less 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 could 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,[94] and only about 20 percent of drugs that begin phase one testing are ultimately approved by the FDA.[95] Only upon approval does the pharmaceutical company discover whether the drug will be successful commercially.

In the research stage of the project, Ligand engages 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 succeeds, only a small fraction of the costs would

---

[92] 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 mineralocortoid receptor.

[93] Because of the regulatory structure, the components of the process are standardized. The description of the process of developing a new drug is based on "Drug Discovery and Development" by innovation.org, http://www.innovation.org/drug_discovery/objects/pdf/RD_Brochure.pdf.

[94] DiMasi et. al., Cost of Innovation in the Pharmaceutical Industry, 10 J. Health Econ. 107, __ (1990).

[95] DiMasi et. al, The Price of Innovation: New Evidence of Drug Development Costs, 22 J. Health Econ. 1541, __ (2003).

be associated with the research phase. The major costs of bringing a drug to market are in the later stages in which the manufacturer must prove efficacy and safety through clinical studies in the FDA approval process.[96] Once a successful compound has been identified, Ligand's role is largely over and, as a result, its role in decision-making largely disappears. Most decisions in this phase are at the "sole discretion" of Warner-Lambert.[97]

The research stage is divided into three periods with Warner-Lambert having an option to abandon in the first two. The Exploratory phase is fifteen months; the project would terminate 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 terminates after three years. Funding levels for both periods are specified in dollars per FTE.[98]

During the research stage, the parties' collaboration is centered on a Joint Research Committee (JRC), an elaborate governance structure responsible for reviewing

[96] 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, supra note __; David T. Robinson & Toby E. Stuart, Network Effects in the Governance of Strategic Alliances, 23 J. L. Econ. & Org. 242 (2006); David T. Robinson, Strategic Alliances and the Boundaries of the Firm, 21 Rev. Fin. Stud. 649 (2008). 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.

[97] In addition, Warner-Lambert contemporaneously purchased approximately seven percent of Ligand's equity. Thus, the contract contained two linked but distinct agreements, one covering research and development of small molecule compounds which act through the estrogen receptors, and another covering the financing of the sale of equity in one research partner to other. Such equity investments are not uncommon in research and development joint ventures. Robinson & Stuart, supra note __.

[98] Article 2.8. FTE (Full-time equivalent) is defined in Article 1 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." 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.

and directing all scientific activities in the pre-clinical period. In effect, the JRC determines the research path. The committee would consist of three members from each firm; all decisions must be unanimous.[99] In the event of a disagreement, the dispute would go to Ligand's CEO and the president of Warner-Lambert's Pharmaceutical Research Division for "good faith resolution" within a specified period. If they failed to resolve the dispute, the parties would be 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 carried more weight than the threat of legal action.

In contrast to the JRC collaborative governance structures, Warner-Lambert's options to abandon the project are unilateral. If the science proved 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 would 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 IND within a defined period after termination, it would have to pay royalties.[100]

As the project moves from the research to the development stage, regulatory and market experience becomes more important. The cost of the project, all of which will be borne by Warner-Lambert, also increases exponentially. As a result, both responsibility

[99] Article 3.1.4

[100] 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." Article 12.2 If Warner-Lambert concludes that Ligand's work in the research phase is not satisfactory, it can terminate the agreement at its sole discretion. It 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. If Warner-Lambert is successful, Ligand would be entitled to its milestone payments and royalties. Article 12.8.

and decision-making shift to it: 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 with 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 regulatory approval for, and (c) when and where and how and on what terms and conditions, to market such Products in the Territory."[101]

The gap between contract formation and the appearance of a marketable drug is, we saw, likely to be more than a decade.[102] So the nature of that drug and its potential value (clinically and financially) is unknown at the time of contracting. The uncertainty is reflected in the manner in which Ligand is compensated. First, as noted, it was to be paid for some fraction (perhaps all) of the FTE assigned to the task. Second, it would be paid a fixed fee upon the initial screening of the Warner-Lambert compound library; if Warner-Lambert chose to go ahead with the Extension term, Ligand would receive an additional fixed fee. These could be labeled milestone payments, although the contract does not do so. The agreement did establish a number of specific milestones, and, upon reaching each milestone, Ligand would receive an additional payment. Finally, if the research produced marketable products, Ligand would receive royalty payments on sales.

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 would fail to yield a commercially viable outcome. Even if Warner-Lambert believed 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 motivated Warner-Lambert's decision to abandon might well be difficult for Ligand even to

[101] Article 4.2.
[102] See TAN __ supra.

observe, let alone to verify. This difficulty is addressed by giving Ligand a matching option. If Warner-Lambert decided not to proceed with the development of a particular "collaboration lead compound," then Ligand would have the right to develop and commercialize it.[103] 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.

## 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 affecting their contractual structure.[104] In particular, the design and specification of product characteristics cannot be contracted for ex ante; rather, these will result from repeat collaboration by employees of both firms. The contracting problem is to craft a structure that a) induces efficient, transaction-specific investment by both parties, b) establishes a framework for iterative collaboration and adjustment of the parties' obligations under conditions of continuing uncertainty—responding, that is, to coordination cascades, and c) limits the risk of opportunism that could undermine the incentive to make relation-specific investments in the first place.

The governance structures in the three contracts we examine incorporate a mix of formal contract and informal 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

[103] Article 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).
[104] See TAN supra.

by the expectation of future dealings.[105] 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.[106] 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.[107]

Such mechanisms would be unnecessary if we assumed that these new collaborative arrangements operate in a contractual state of nature, in which good-faith cooperation can be expected from an arm's length partner, without 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 non-cooperatively 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 being a stark current example.[108] Thus, cooperation does not eliminate tension or

[105] See, e.g., Avinash K. Dixit, Lawlessness and Economics: Alternative Modes of Governance 16 (2004).
[106] This is the "folk theorem" of non-cooperative game theory. See, e.g., Id. at 61; Drew Fudenberg & Eric Maskin, The Folk Theorem in Repeated Games with Discounting or with Incomplete Information, 54 Econometrica 533 (1986).
[107] Laura Poppo & Todd Zenger, Do Formal Contracts and Relational Governance Function as Substitutes or Complements, 23 Strat. Mgmt. J. 707 (2002), 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 assessing the actual contractual techniques that create and sustain the relationship. Thus, our focus here on actual contracts and contracting techniques extends their insightful intuition.
[108] See, e.g., Boeing Delays 787 Debut due to Supply Chain and Assembly Problems, Los Angeles Times, April 10, 2008, available at https://articles. latimes.com/april/10/2008/business/ fi-boeing10.

conflict;[109] 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.

A. *Transactional Variations in Collaborative Contract Design.*

In this Part, we argue that 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 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 rights to decision-making is feasible.

Contracting for long-term collaborative innovation is particularly well illustrated by the Deere-Stanadyne agreement,[110] 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

[109] O'Sullivan, supra note __ at ___, details the wide range of conflict that is possible in even a successful collaborative contract for innovation.
[110] See TAN __ supra.

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.[111] Recall that in this agreement the three-year, explicit supply arrangement supported Apple's sale to SCI of Apple's manufacturing plant.[112] 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 60 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/assembly 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

[111] See TAN __ supra.
[112] See TAN __ supra.

relationship by threatening to withhold further cooperation. While each party makes a front-end specific investment---SCI in buying the plant and Apple in 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.

Contracting for a discrete project, which contemplates collaboration during the development stage, but allows for opportunism when the development stage is completed, is well illustrated by the Warner-Lambert - Ligand agreement.[113] Here the contract contemplates 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 commercialization of the compounds – 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's 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 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

[113] See TAN __ supra.

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 –the assignment of decision rights through the use of options -- is both necessary and feasible.

B. *Coordination Cascades and the New Governance Mechanisms*

In the following discussion, we show that the collaborative process is structured to produce information about the parties' capabilities and their capacity for cooperation. The innovative governance terms of these co-design contracts (detailed specification of the collaboration and dispute resolution mechanisms rather than specification of long-term price and quality terms) are driven by the parties' investment in the information that results from collaboration.

1. *The Function of The Collaborative/Learning Phase.*

*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 a new supplier's perspective, the contract is an invitation to enter the Achieving Excellence program, though 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, which provide 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 fact legally unenforceable (a fact that previously had been missed in the large literature that concerns the contract), which the lawyers who wrote the contract would have known.[114] Thus, this more famous contract operated like the Deere-Stanadyne contract: as a way of organizing the parties' learning and continued collaboration, and 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,[115] but Apple monitors, requiring progress reports, pre-production review and progress monitoring to achieve pre-production deliverables.[116] The formal contract thus specifies an iterative process, in which the parties cooperate in designing the product and the manufacturing process. Thereafter, 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 pre-production deliverables

[114] Victor P. Goldberg, Lawyers Asleep at the Wheel? The GM-Fisher Body Contract, Indus. & Corp. Change, (forthcoming 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 which Fisher Body had demonstrated in its relations with demanding customers, such as Chrysler. See Helper, MacDuffie and Sabel, infra note____.
[115] Article 5.2
[116] Article 5.3

does real production begin. The parties agree to co-engineer product cost reductions, set goals and meet quarterly while production is underway. At each step in the pre-production process, the parties learn about each other's technical capacity to accomplish the necessary design tasks, abilities to cooperate in a productive fashion, to share information readily, and to effectively manage the employees involved despite the non-hierarchical characteristics of the relationship. This process takes on greater and greater importance with time. As the product evolves in response to technology and market 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 designates 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 skills, do they have the skills necessary to manage cooperation and, finally, what is their dispute resolution style---do they cooperate or do they 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 about 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 compounds identified.[117]

As suggested above, at this stage 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.

---

[117] 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 TAN __ supra. In two recent articles, Robinson & Stuart explore the role of networks in supporting reputational sanctions in biotech strategic alliances. Robinson & Stuart, Financial Contracting, supra note __; Robinson & Stuart, Network Effects, supra note --. The articles conclude that the centrality of the biotech company in the industry network of alliances is the key to explaining both the use of non-verifiable 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, working paper 2007, available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=944784.

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 others character, it cannot guarantee a cooperative outcome. Common experience teaches that transactions that rely on informal enforcement can break down; relational enforcement requires that each party be able to observe and properly characterize the other's behavior. This 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. Then 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.[118] Lacking clarity, either party may mischaracterize the other's actions. Under these circumstances, without the necessary linkage between action and response, reciprocity will be a less effective mechanism for enforcement.

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,[119] b) the assignment of decision rights to

[118] 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, see Jonathan Bender, Roderick M. Kramer & Suzanne Stout, When in Doubt: Cooperation in a Noisy Prisoner's Dilemma, 35 J. Conflict. Res. 691 (1991).

[119] See e.g., the Warner-Lambert/Ligand contract, where the parties agree to disclose to each other all background technology relevant to the field and helpful to perform the work set out in the Research Plan. Both commit to provide resources necessary to carry out the Plan (Art. 2.1). Each agrees to allocate XX FTEs to the Plan (Art. 3.3). All collaboration discoveries and technology are to be promptly disclosed to either party and monthly reports are required on all biological structures and compounds. 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.(Art. 10.6). SCI agrees to share all production cost information and Apple agrees to share all marketing forecasts.(Art. 10.9). The parties co-design lead time reductions.(Art. 11.7)

a joint project management team subject to a unanimity rule,[120] and c) the appointment of "referees"-- representatives from each firm charged with resolving disputes.[121] This 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.[122] The referee also serves as an informal disciplining mechanism. Superiors are unlikely to look with favor on subordinates who send problems up the line for resolution. The subordinates' job was 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.[123] Finally,

---

[120] 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) consisting of 3 representatives from each firm, one of the Ligand representatives serves as the Chair. The JRC has control and decision rights over the Research Plan and is responsible for co-design and implementation (Art. 3.1.1 -2). The JRC holds quarterly meetings and all decisions are by unanimous vote (Art. 3.1.4).

[121] See, e.g., the Warner-Lambert/Ligand contract which provides that all decisions are by unanimous vote (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).

[122] 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.)

[123] Stewart Macaulay noted this forty years ago:

> Internal 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.

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.

In what follows, we focus on the potential for opportunism resulting from both the collaborative and pie-splitting phases of the transactions, and on the characteristic responses. We take up the collaboration and its consequences in deterring both ongoing and endgame opportunism in turn.

2. *Mechanisms for Deterring On-Going Opportunism*

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.[124] The term is connected most directly in the literature with the inquiry whether rigorous initial competition adequately substitutes for the absence of competition after the first purchase;[125] but it also appears in accounts of search costs,[126] path dependency,[127] first mover advantage,[128] or

---

Stewart Macaulay, Non-Contractual Relations in Business: A Preliminary Study, 28 Am. Soc. Rev. 1 (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.

[124] See, e.g., Joseph Farrell & Paul Klemperer, Coordination and Lock-In; Competition with Switching Costs and Network Effects, in 3 Handbook of Industrial Organization 1970 (M. Armstrong & R. Porter, eds. (2007)(defining switching costs).

[125] See Paul Klemperer, Switching Costs, *New Palgrave Dictionary of* Economic 125 (2nd ed, S. Durlauf and L. Blume, eds. 2008)(central question in literature is whether ex ante competition substitutes for standard period by period competition).

[126] Search costs refer to the costs associated with simply finding out what other parties offer competitive goods and at what price. See, e.g., Alan Schwartz & Louis Wilde, *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, 127 U. Pa. L. Rev. 630 (1979).

[127] 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 non-economic, 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, Positive Feedbacks in the Economy, SCI. AM., Feb. 1990, at 92-99 provides an accessible survey of the concept.. For discussions in an institutional context, see Mark Roe, *Chaos and Evolution in Law*

reputation.[129] For our purposes, the common core is the idea that information costs—for example the costs of learning whether a supplier is good at learning to collaborate, and can be counted on to act in good faith in hard times--can create barriers to replacing one party to a contract with another. In a deep spot market such substitution 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 to exiting a relationship. The switching costs of interest to us here, however, have two additional 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. Switching costs in our sense 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 that would be lost if the relationship terminated, and which would have to be duplicated with any new supplier. Second, these investments are made gradually, rather than 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

---

*and Economics*, 109 Harv. L. Rev. 641, 647-53 (1996); S. J. Leibowitz & Stephen Margolis, *Path Dependence, Lock-In, and History*, 11 J.L. Econ. & Org. 205, 206-08 (1995); and Michael Klausner, *Corporations, Corporate Law, and Networks of Contracts*, 81 Va. L. Rev. 757 (1995).

[128] See, e.g., CARL SHAPIRO & HAL VARIAN, INFORMATION RULES: A STRATEGIC GUIDE TO THE NETWORK ECONOMY (1999).

[129] 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.

capabilities and 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 makes 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, cooperation, and dispute resolution style.[130] 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' investment 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 claims 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

[130] See TAN __ supra.

replace SCI would be costly for Apple for three reasons: a) The delay in securing another supplier given that the specifications and production technology had been co-engineered with SCI, and necessarily reflect the idiosyncrasies of SCI's capabilities; b) Apple's lack of knowledge about the capabilities of potential new suppliers that it had concerning SCI; and c) 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 chicken game---even though both 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 play a cooperative or chicken game. To the extent that styles are not greatly malleable[131]---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 about the output and, simultaneously, about each other's capacity to cooperate, both in solving problems and in resolving the inevitable disputes. 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

[131] The behavioral literature suggests that these preferences are quite stable. Robert E. Scott, A Theory of Self-Enforcing Indefinite Contracts, 103 Colum. L. Rev. 1641(2003) (discussing comfort agreements and other informal agreements to agree that provide key information about the parties' preferences to cooperate).

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.[132]

But the contracts in our group do not square easily with this common understanding of the domain of self-enforcing agreements. The Deere-Stanadyne and the 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 supplier's conduct to potential suppliers which rest, in turn, on conditions that are difficult to sustain.[133] These conditions may not be met or only partially met in the settings of these contracting parties. Indeed, reputational sanctions are best seen as a special, and especially context sensitive, case of switching costs.

The transactions represented by the group of co-design contracts are, for the most part, interactions in heterogeneous markets where reputational constraints are thought to be quite weak.[134] In such an environment, reputation alone is an inadequate means of

---

[132] Rene Sacasas & Don Wiesner, Comfort Letters: The Legal and Business Implications, 104 Banking L. J. 313, 328 (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....Id.

[133] See Gilson, supra note __, at __.

[134] Reputations are most effective in small homogeneous communities, where non-cooperative behavior quickly becomes common knowledge and sanctions against such behavior can be effectively imposed. See e.g., Avner Grief, *Informal Contract Enforcement: Lessons from Medieval Trade* in 2 The New Palgrave Dictionary of Economics and the Law 287 (1998); Janet Landa, *A Theory of the Ethnically Homogeneous Middleman Group: An Institutional Alternative to Contract Law*, 10 J. Legal Stud. 349 (1981). 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, *Merchant Law in a Merchant Court: Rethinking the*

credibly enforcing promises. Even if others can observe the contractual failure, it would be difficult to learn the true reasons why the particular transaction broke down. 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.[135] More precisely, they do not seem to be a tit-for-tat solution to a multi-round prisoners' dilemma in the form of a supply contract. 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.

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.

---

*Code's Search for Immanent Business Norms*, 144 U. Pa. L. Rev. 1765, 1771-77 (1996); Lisa Bernstein, *Private Commercial Law in the Cotton Industry: Creating Cooperation Through Rules, Norms, and Institutions,* 99 Mich. L. Rev. 1724, 1745-54 (2001).

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

The experimental behavioral literature shows that contracting parties are heterogeneous regarding their preferences for cooperation---a significant percentage reward cooperation and punish defection, but an equally large percentage will act in pure self-interest and behave opportunistically.[136] The problem then is to identify who are the cooperators and who are likely to behave opportunistically.[137]

At first cut, the willingness to write a contract that conditions on non-verifiable 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, and therefore renders observable the parties' character with respect to dispute resolution, at a time when specific investments are modest and switching costs are low.[138] Thus, early reciprocation, which serves to

[136] There are two key findings of a substantial body of experimental evidence on people's propensity to reciprocate. First, many people respond cooperatively to generous acts, and, conversely, punish non-cooperative behavior. Second, , the observed preference for reciprocity is heterogeneous. Some people exhibit reciprocal behavior and others are selfish. Taking all the experiments together – gathered from diverse countries and cultures – the fraction of reciprocally fair subjects ranges from 40 to 60% as does the fraction of subjects who are selfish. For discussion, see Ernst Fehr et al, Reciprocity as a Contract Enforcement Device: Experimental Evidence, 65 Econometrica 833 (1997); Ernst Fehr & Klaus M. Schmidt, A Theory of Fairness, Competition and Cooperation, 114 Q.J. Econ. 817 (1999); Mathew Rabin, Incorporating Fairness into Game Theory and Economics, 83 Am. Econ. Rev. 1281 (1993); Ernst Fehr & Simon Gatcher, Fairness and Retaliation: The Economics of Reciprocity, 14 J. Econ. Perspectives 159 (2000). For applications of this experimental evidence to contract and international law, see Scott, Self-Enforcing Agreements, supra note ---; Robert E. Scott & Paul B. Stephan, Self-Enforcing International Agreements and the Limits of Coercion, 2004 Wisc. L. Rev. 551; Robert E. Scott & Paul B. Stephan, The Limits of Leviathan: Contract Theory and the Enforcement of International Law (Cambridge 2006).

[137] 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.

[138] 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 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." John McMillan & Christopher Woodruff, Private Order Under Dysfunctional Public Order, 98 Mich. L. Rev. 2421, 2432 (2000).

gradually lock the parties into a relationship, signals that the parties intend to cooperate, and provide information about their type.

There are two ways that the formal governance structure established in these contracts can function to identify a contracting 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.[139] 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, 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 is thus an example of the expenditure of time for the purposes of communication, thereby increasing the cost of switching.[140] In this case, the transactors are not only subject to observation, but the parties must spend considerable time and effort in executing portions of an agreement that is effectively non-verifiable and thus can only be self-enforcing. Since parties with preferences to cooperate are able to capture the returns to general information about their type through an enhanced reputation for cooperation, they are more willing to spend resources to provide this information.[141] The expenditure of time performing a non-verifiable promise is itself a signal: because these costs make it

---

[138] See George Akerloff, The Market for "Lemons:" Quality Uncertainty and the Market Mechanism, 84 Qu. J. Econ. 355, 366 (1970

[139] See George Akerloff, The Market for "Lemons:" Quality Uncertainty and the Market Mechanism, 84 Qu. J. Econ. 355, 366 (1970).

[140] A. Michael Spence, Time and Communication in Economic and Social Interaction, 87 Qu. J. Econ. 651 (1973

[141] Another way of expressing the point in the text 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. See Joseph E. Stiglitz, The Theory of Screening, Education and the Distribution of Income, 65 Am. Econ. Rev. 283, 287 (1975). 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. See Ronald J. Gilson, Controlling Family Shareholders in Developing Countries: Anchoring Relational Exchange, 60 Stan. L. Rev. 633 (2007) (discussing the requirements of a viable reputation market).

difficult for a competitor who lacked the characteristic nonetheless to imitate the behavior, it may signal a preference for reciprocity.

Our examples of co-design contracts suggest that the widespread use of elaborate, non-verifiable and therefore non-enforceable governance obligations may be in part a function of their properties as screens for the parties' willingness to engage in cooperative behavior. The knowledge of the other parties' capacities, and therefore the cost of acquiring equivalent information about a replacement, increases in parallel with iterative performance. This period of build up of 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. This 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 bi-lateral monopoly and run the risk of 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, or 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 uncertainly---one can not 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 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.[142] 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. *Deterring 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

[142] McMillan & Woodruff make a similar argument in discussing relational contracting in underdeveloped markets with poor legal systems:

> [b]ilateral 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.

McMullan & Woodruff, supra note – at 2432.

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 when decision rights are contractible.

The Warner-Lambert-Ligand contract provides an illustration of the problem and a response. As we have seen,[143] 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. 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 similar fashion. But once the 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 non-cooperative phase.

The Warner-Lambert-Ligand contract illustrates the utility of nesting explicit options---the allocation of sequential ex post decision rights among the parties to deal with opportunism in the post-cooperative 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 can not

[143] See TAN __ supra.

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.[144] Warner-Lambert's reputation with other potential partners might be impaired, as we argued above, a form 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.[145] Given the asymmetry in outside options, the parties have differing threat points, 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[146] and movie development[147]---which also involve innovation without the ability ex ante to condition shifting payouts on outcomes---the problem is addressed

[144] 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 is 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 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.

[145] See TAN __ supra.

[146] See Gilson, supra note __.

[147] Victor P. Goldberg, The Net Profits Puzzle, 97 Colum. L. Rev. 524 (1997).

through nested options. 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 rates. 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 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. 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. [148] The results show that repeated interactions cause a significant increase in the effort levels of the subjects.[149]

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.[150] 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.[151]

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.[152] But once the entire relationship, including its implicit aspects, is

---

[148] See Martin Brown, Armin Falk & Ernst Fehr, Incomplete Contracts and the Nature of Market Interactions, Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 38 (2002).(comparing the effort levels of agents in a one-shot gift exchange game with the effort levels in a similar game in which repeated interaction was an additional material outcome)
[149] Id. See also Ernst Fehr & Klaus M. Schmidt, Theories of Fairness and Reciprocity, note -- supra, at 6-7; Ernst Fehr, Simon Gächter & Georg Kirchsteiger, Gift Exchange and Reciprocity in Competitive Experimental Markets, 42 Eur. Econ. Rev. 1 (1998); Gary Charness, Responsibility and Effort in an Experimental Labor Market, 42 J. Econ. Behav. & Org. 375 (2000); Ernst Fehr & Armin Falk, Wage Rigidity in a Competitive Incomplete Contract Market, 107 J. Pol. Econ. 106 (1999); Simon Gächter & Armin Falk, Reputation or Reciprocity, Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 19 (1999).
[150] Scott & Stephan, The Limits of Leviathan, supra note ---.at 102.
[151].Id.
[152] Ernst Fehr & Simon Gachter, Fairness and Retaliation: The Economics of Reciprocity, J. Econ. Persp., Summer 2000, at 159.

subject to formal enforcement, voluntary reciprocity declines along with the overall level of cooperation.[153] These experimental results suggest that formal legal sanctions and informal sanctions based on reciprocity may well conflict with each another. In other words, formal contracting may "crowd out" behavior based on relational contracting.[154]

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.[155] 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 of the agreement actually enhances cooperation in those dimensions of the agreement that are non-verifiable.[156]

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 on-going relationship, the evidence suggests that explicit mechanisms designed to deter opportunism at the end stage of collaborative contracts may reinforce

[153] Ernst Fehr & Simon Gachter, Do Incentive Contracts Undermine Voluntary Cooperation? (Inst. For Empirical Research in Econ., Univ. of Zurich, Working Paper No. 34, 2002).
[154] For discussion, see Scott, A Theory of Self-Enforcing Agreements, supra note – at 1688-92
[155] See, e.g., Bruno Frey & Reto Jegan, Motivation Crowding Theory, 15 J. Econ. Surveys 589 (2001); Bruno Frey & Matthias, Motivation Transfer Effect, University of Zurich, Institute for Empirical Research in Economics, (mimeo 2001); Iris Bohnet, Bruno Frey, & Steffen Huck; Ernst Fehr & Simon Gachter, supra note ---.
[156] Sergio Lazzarini, Gary J. Miller & Todd R. Zenger, Order with Some Law: Complementarity versus Substitution of Formal and Informal Arrangements, 20 J. L. Econ. & Org. 261 (2004).

the patterns of trust and reciprocity, thereby enabling the parties better to enforce themselves the non-verifiable portions of the relationship.

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

Contracts for innovation create inter-firm 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 of in the 1930s.[157] This brings us back to the second current 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 that the diffusion of new forms of cooperation in recent decades have 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 can 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 need to assure both ex ante and ex post efficiency, and so on. Alternatively, agents can produce the good or service internally by establishing a corresponding hierarchy within the firm. Then the

[157] Ronald H. Coase, The Nature of the Firm, 4 Economica 386 (1937).

costs of transacting are the costs of establishing and operating a bureaucracy.[158] 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.[159] He argued that the asset specificity of complementary investments created the risk of costly holdups, 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.[160]

But, it turns out, Coase in the 1930s 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.[161] 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

[158] In fact the costs of hierarchy have remained under-illuminated in this strand of the literature, but the thrust of the argument is in the direction suggested here. See Paul Milgrom & John Roberts, The Economics of Modern Manufacturing: Technology, Strategy, and Organization, 80 Am. Econ. Rev. 511 (1990).
[159] See note – supra.
[160] See TAN __ supra.
[161] Ronald H. Coase, The Acquisition of Fisher Body by General Motors, 43 J. L. & Econ. 15 (2000).

continuing to operate as an independent company.[162] Most recently, Coase has gone on to argue that the canonical story connecting hold-ups and vertical integration—the final take over 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 suppliers—an early variant of the relations described here—than to protect itself against the (arguably non-existent) threat of a hold-up.[163] 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.[164]

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?[165] 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 firm, Holmström and Roberts broached this more thorough going critique a decade ago.[166] 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 the light of our findings here.

---

[162] Id. at ---.
[163] Ronald H. Coase, The Conduct of Economics: The Example of Fisher Body and General Motors, 15 J. Econ. & Mgmt Strategy, 255 (2006).
[164] See TAN __ supra. Note that this makes the GM-Fisher Body contract analogous to the Deere-Stanadyne contract.
[165] Lafontaine, et. al, supra note __, surveys the empirical transaction cost literature.
[166] See note ---supra.

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.[167] Indeed, much of that literature argues that the chief function of the firm is not to solve hold-up problems but instead to facilitate the flow of the un-codifiable, living knowledge that animates innovation.[168]

Holmström and Roberts' 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; the firm's boundary therefore is dictated by the breadth of hold-up problems.[169] The core of Coase's criticism of transaction cost economics is that the link between the character of transactions costs and

[167] Ikujiro Nonaka & David Teece eds., Managing Industrial Knowledge: Creation, Transfer and Utilization (2001); Margherita Balconi, Andrea Pozzali and Riccardo Viale, The "Codification Debate" Revisited: A Conceptual Framework to Analyze the Role of Tacit Knowledge in Economics, 16 Indus. & Corp. Change 823 (2007).

[168] Bruce Kogut, Joint Ventures: Theoretical and Empirical Perspectives, 9 Strat. Mgmt. J. 319 (1988); Bruce Kogut and Udo Zander, Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology, 3 Org. Science 383 (1992).

[169] 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 was not costly, then a state contingent contract was the appropriate transaction form. If, instead, bounded rationality was still present, information was costly, but opportunism was 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 __.

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,[170] 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 holdup problem, neither does it reflect contract, at least as conventionally understood. The contracts between 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.[171] Rather, the governing mechanism is the "long-term, repeated nature of the interaction,"[172] 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

[170] An Interview with Oliver Williamson, 3 J. Institutional Econ. 373 (2007) (acknowledging that Japanese supplier relations are inconsistent with the transactions costs framework.)
[171] See Holmström & Roberts, supra note – at 8. 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." Oliver Williamson, Transaction Cost Economics: The Governance of Contractual Relations, 22 J.L. & Econ. 233, 241 (1979).
[172] Holstrom & Roberts, supra note _, at 8.

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." 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.

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, bi-lateral 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" 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. It seems quite possible, therefore, 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.[173] 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. 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.[174]

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 in a specialized exchange economy."[175] 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 world of a half century later. To illustrate the limits

---

[173] 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. 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 and from each other. See Charles Sabel, Ungoverned Production: An American View of the Novel Universalism of Japanese Production Methods, Columbia Law School L. & Econ. Working Paper, Feb. 1996); Ronald J. Gilson, Globalizing Corporate Governance: Convergence of Form or Function, 49 Am. J. Comp. L. 329, 332 (2001).

[174] 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." 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. See Coase, The Example of Fisher Body and General Motors, supra note ---.

[175] Coase, The Nature of the Firm, supra note – at --.

of current explanatory schemes, Holmström and Roberts introduce a form of governance—Japanese subcontracting—which 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 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 counter-tendencies even amidst the general, current tendency to vertical dis-integration; 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.[176] Even as U.S. steel firms were divesting their holdings of iron ore, Mittal (now Arcelor/Mittal), a multi-national offshoot of an Indian steel maker, was buying ore reserves to protect itself against the possibility of world-wide limits to supply.[177] It would be a surprise, given the transformations of recent decades, if this shifting of firm boundaries did not produce

[176] Yves Doz and Mikko Kosonen, The Dynamics of Strategic Agility: Nokia's Rollercoaster Experience, 50 Cal. Mgmt Rev. 95 (2008).
[177] Mittal Steel in 2005: Changing the Global Steel Game, Harv. Bus. School Case.

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 solves—and the ones it doesn't.

## 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. This will be our next project.

european corporate governance institute

## about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.org

european corporate governance institute

ECGI Working Paper Series in Law

Editorial Board

| | |
|---|---|
| Editor | Guido Ferrarini, Professor of Law, University of Genova & ECGI |
| Consulting Editors | Theodor Baums, Director of the Institute for Banking Law, Johann Wolfgang Goethe University, Frankfurt & ECGI<br>Paul Davies, Cassel Professor of Commercial Law, London School of Economics and Political Science & ECGI<br>Henry B Hansmann, Augustus E. Lines Professor of Law, Yale Law School & ECGI<br>Klaus J. Hopt, Director, Max Planck Institute for Foreign Private and Private International Law & ECGI<br>Roberta Romano, Allen Duffy/Class of 1960 Professor of Law, Yale Law School & ECGI<br>Eddy Wymeersch, Professor of Commercial Law, University of Ghent & ECGI |
| Editorial Assistant : | Paolo Casini, "G.d'Annunzio" University, Chieti & ECARES,<br>Lidia Tsyganok, ECARES, Université Libre De Bruxelles |

www.ecgi.org\wp