From Incentives to Control to Adaptation:
Exploring Interactions Between Formal and Relational Governance

George Baker, Robert Gibbons, and Kevin J. Murphy

February 18, 2022

*In 1991 we began to model interactions between formal and relational incentive contracts. We were motivated by a case study on compensation, and we saw this work as a contribution to agency theory. By the time the paper was published (*QJE*, 1994), we had begun to view the research agenda more broadly—with connections to organizational culture (Kreps, 1990), the theory of firms' boundaries (Coase, 1937), and more. Eventually, we built from this initial work, analyzing delegation within organizations as necessarily informal (*JLEO*, 1999), and moving beyond* relational agency *(where a principal motivates an agent through various promises, not limited to compensation) to* structuring relationships *(where parties choose their formal governance structure to facilitate their relational contract). The latter led to our relational analysis of when classic buyer-supplier interactions should be governed under integration and when under non-integration (*QJE*, 2002) and to our working paper (last revised in 2011) on how formal contracts between firms might facilitate "relational adaptation" as events unfold. In this essay we sketch theoretical, empirical, and methodological lessons we learned during this twenty-year journey.*

1. Introduction

The editors of this special issue invited contributors to describe (a) the intellectual and personal context that produced a particular paper, (b) the lessons of that paper, and (c) how those lessons have been or could be applied or extended. For "Subjective Performance Measures in Optimal Incentive Contracts," published in the *Quarterly Journal of Economics* in 1994, our story starts in an empty classroom in the basement of Harvard Business School in the fall of 1991, where our shared interest in a case study on incentive compensation led us to cover many large blackboards with algebra.

By 1991, we had absorbed Kerr's (1975) "On the Folly of Rewarding A, While Hoping for B," so we began our 1994 paper with well-known examples in which employees took actions arguably to increase their compensation, even if those actions decreased long-run firm value. We would now add mortgage brokers at Washington Mutual (and other similar lenders) who, in the years leading up to the 2007-09 financial crisis, were rewarded for writing subprime mortgages to virtually anyone who applied.[1] Or, Wells Fargo branch workers and managers whose aggressive cross-selling goals lead to the creation of two million fake customer accounts by late 2016.[2] In short, decades ago and still today, "Business history is littered with firms that got what they paid for" (BGM 1994: 1125).[3]

Our interest in the case study on compensation at Lincoln Electric (Fast and Berg, 1975) arose because Lincoln appeared to do things differently: in addition to a formal piece rate rewarding a worker for output, compensation also included a discretionary bonus based on the worker's dependability, quality, ideas, and cooperation—behaviors harder to measure than the quantity the worker produced. We wondered not only whether parties in an ongoing relationship could potentially "contract" on things they both observe, even if such contracts could not be enforced in court, but also whether the possibility of such informal (*i.e.*, self-enforced) contracts could mitigate the distortions created by formal (*i.e.*, court-enforceable) contracts. More generally, we wondered how formal and informal contracts might interact.

As we describe in Section 2, our 1994 model thus explored the joint use of (i) formal incentives based on objective performance measures with (ii) promised bonuses based on aspects of performance that courts cannot observe. With 30-30 hindsight—*i.e.*, the awareness that comes from 30 years of considering an issue—we could have called this paper "Interactions Between Formal and Relational Incentive Contracts." Instead, the paper refers to "subjective" performance measures, whereas now we would call a bonus "discretionary" if it depends on things that both parties observe (saving "subjective" for things they might interpret differently). In addition, the paper refers to "implicit" contracts, whereas now we would call them "relational" (to emphasize the importance of an ongoing relationship for such contracts, and to avoid the impression that nuanced shared understandings could arise and function effectively if left implicit).

Those discussions in a basement classroom launched twenty years of joint research—a highlight of each of our careers. Some of this research continued to study *relational agency*, by which we mean settings where a principal influences an agent's actions through credible promises

[1] See Peter S. Goodman and Gretchen Morgenson, "By Saying Yes, WaMu Built Empire on Shaky Loans," *New York Times* (December 27, 2008).

[2] See Emily Glazer, "Well Fargo Fined for Sales Scam—Bank to pay $185 million as regulators say its selling tactics included creating fake accounts" *Wall Street Journal* (September 9, 2016).

[3] For brevity, here and below, quotations from our own work appear with "BGM," the year and the page(s).

(modeled as repeated-game equilibria). For example, in Section 2, in addition to describing our 1994 model, we also discuss "Informal Authority in Organizations," published in the *Journal of Law, Economics, and Organization* in 1999. In this model, the principal can create incentives for the agent by promising to implement projects discovered and proposed by the agent, provided that the project's payoff is not too negative for the principal. If the agent's proposals will be accepted frequently, one might call this situation "empowerment," but our point, discussed in Section 2, is that such empowerment, like discretionary bonuses, requires a relational contract.

While our substantive interest in the discretionary bonuses at Lincoln Electric led us to explore other informal ways that principals may create incentives for agents (such as the form of empowerment just described), unexpected theoretical findings in the 1994 paper also sparked a new line of inquiry, which we call *structuring relationships*. In fact, the last section of our 1999 paper briefly illustrated this new inquiry: having analyzed a boss's ability to create incentives for a subordinate by promising to implement proposed projects that are not too harmful for the boss, we then considered spinning off the agent's activities as a new firm that the agent owns, thus transforming the parties' relationship from a boss and subordinate within one organization to something like an alliance between two organizations. Relational contracts may still be an important aspect of value creation between such non-integrated organizations, but now it may be the former subordinate who is tempted to renege (by implementing projects that harm the former boss), rather than the boss (by failing to implement projects that would be exceptionally good for the subordinate).

Having discussed relational agency in Section 2, we consider structuring relationships in Section 3. In the former, the roles of principal and agent are fixed; in the latter, choosing who controls what decisions—roughly speaking, endogenizing who is the principal and who the agent—influences the parties' temptations to renege on a given relational contract. Our discussion in Section 3 reverses the emphasis of much of the literature on the formal design of organizations and contracts: rather than choosing the formal design for its direct effects (perhaps anticipating that relational contracting might further improve the parties' payoffs), in our approach the formal design is chosen to facilitate the parties' relationship.

In Section 3.1 we discuss "Relational Contracts and the Theory of the Firm," published in the *Quarterly Journal of Economics* in 2002, where we emphasized that not only organizations but also "[b]usiness dealings are … riddled with relational contracts" (2002: 39) —such as the informal promises often involved in trade credits, supply assurance, hand-in-glove supply relationships, alliances, and more. In this paper we studied supply relationships between upstream and downstream parties, analyzing whether the parties are better able to induce efficient specific investments when they are integrated or non-integrated, adding to this classic analysis our new appreciation for relational contracts between firms. Among other results, we found that the temptation to renege on a given relational contract depends on the formal governance structure (here, integration or non-integration of the supply chain), dictating which formal governance structure is optimal for that relational contract.

And in Section 3.2 we discuss our unpublished working paper "Relational Adaptation," last revised in the Fall of 2011—twenty years after our collaboration began. In this paper we explored formal contracts that move decision rights across fixed firm boundaries. To assess whether we were modeling the important drivers of such formal governance structures, we conducted a series of detailed interviews with practitioners who negotiate and manage such contracts. Familiar ideas—such as inefficient hold-ups motivated by specific investments and inadequate investments motivated by bargaining over returns—were rarely mentioned by these practitioners. Rather, two ideas emerged as especially important factors determining the form and performance of collaborations between firms, such as alliances: spillovers (or externalities) from the collaboration onto the partners' main businesses; and the need for governance structures to induce efficient responses to changing circumstances, since contracts often cannot. We therefore made these two ideas central to our 2011 model.

There were thus two differences between our 2002 and 2011 papers. First, whereas the 2002 paper asked whether relational contracts are better facilitated by integration or non-integration (but ignored formal contracting beyond this integration decision), the 2011 paper focused on non-integrated parties and asked how formal contracts between the parties (modeled as a formal allocation of decision rights) might facilitate relational contracting. Second, whereas the 2002 studied how the combination of formal governance and relational contracting could improve the efficiency of *ex ante* specific investments, the 2011 paper studied how this combination could improve the efficiency of *ex post* adaptation to changing circumstances. In the end, however, the similarities between these two papers proved more important than the differences: they both emphasize relationships between organizations, and in both papers we established that (as we elaborate below), given a particular relational contract, the optimal formal allocation of control for that contract minimizes the maximum aggregate temptation to renege on that contract.

All four of our papers discussed above share the ideas that (1) ongoing relationships may allow parties to contract on outcomes not verifiable by third parties (*e.g.*, courts), but (2) there are limits on such contracting, and that, because of these limits, (3) there may also be an important role for formal contracting (*e.g.*, formal incentives based on objective performance measures, or formal contracts moving control rights between firms). But the possibility of formal contracting can create costs as well as benefits. In particular, from our 1994 paper onward, we have explored the idea that (2′) relational contracting may be limited by the possibility of formal contracting *after* reneging on the relational contract. Indeed, in both Sections 2 and 3 below we show that this "fallback payoff" (*i.e.*, the payoff from efficient spot governance after reneging) can be imperfect yet sufficiently attractive that *no* relational contracting is feasible. In such settings, the parties might be better off if they could destroy the possibility of formal contracting (such as by making it prohibitively costly to collect particular performance measures), thereby worsening the fallback payoff after reneging and relaxing this particular limit on relational contracting.

To close this Introduction and launch our discussion of two decades of research, we recall our perspective in 1991. At that point, we had each done work on agency problems. Much of our work (and the literature as a whole) was rooted in static models of formal contracts with full commitment by the principal; some papers considered dynamic settings that again had full commitment by the principal within each period but issues such as learning created new dynamics across periods. In contrast, our 1994 paper was motivated by the prospect that discretionary bonuses could mitigate the distortions caused by formal incentive contracts, leading us to analyze commitment problems within a period and their (partial) resolution in repeated games.

Beyond adding a new analytical technique (repeated games) to our modeling toolkit, the 1994 paper shaped our future thinking and research in four additional ways. First, the paper was directly motivated by the world—specifically, by the discretionary bonuses at Lincoln Electric. Second, this paper opened our eyes to the many informal aspects of organizations, well beyond compensation. Third, the paper also led us to appreciate relationships *between* organizations. And fourth, our 1994 paper taught us that the formal and informal aspects of organizations not only co-exist but interact. More specifically: (a) the optimal formal governance structure typically depends on the extent to which relational contracting is possible, and (b) rather than accept exogenous features of the environment as the sole determinants of the parties' temptations to renege on a given relational contract, the parties often can design their formal governance structure to reduce these temptations. We illustrate these points in Sections 2 and 3, and we return to them in the Conclusion in Section 4.

2. Relational Agency

In this essay, we define *relational agency* to mean principal-agent relationships where the principal influences the agent's actions through credible promises (modeled as repeated-game equilibria). As noted above, both discretionary bonuses and some forms of empowerment fit this definition.

2.1 Formal and Relational Incentive Contracts

In 1991, we had absorbed the agency theory of Holmstrom and Milgrom (1991) and Baker (1992) concerning distortions that could be caused by formal incentive contracts. In addition, we were also beginning to appreciate that the substantial existing literature on repeated games—*e.g.*, from Friedman (1971) through Abreu (1988)—might offer tools for exploring the informal aspects of organizations and other relationships. Bull (1987), MacLeod and Malcomson (1989), and Kreps (1990) were early analyses of such informal aspects, but their analyses were not as rooted in agency theory as we intended our model to be. More specifically, the case study of compensation at Lincoln Electric motivated us to study the combined use of formal and relational incentive contracts.[4]

In our 1994 paper, we used Baker's (1992) model to capture distortions that could be caused by formal incentive contracts. This choice was not parochial. Rather, to emphasize the key issue in our analysis—namely, the potential interactions between formal and relational incentive contracting—we simplified the relational contracting by assuming that the agent is risk-neutral, which Baker's model studied but Holmstrom and Milgrom's (1991) did not. But in 1991 we did

[4] For subsequent models analyzing formal and relational incentive contracting, see Section 5 of Malcomson (2013).

not know Feltham and Xie's (1994) model, which provides a simpler account of distortions from formal incentive contracts for risk-neutral agents. If we were writing (or teaching) our 1994 paper today, we would begin from Feltham and Xie's model, so we do that here. We relegate much of the formulation and analysis of the resulting model to Appendix 1; here we loosely sketch an illustrative formulation, moving quickly to the results of the model and where they took us next.

Consider a repeated game between a firm and a worker, where both parties share the discount rate r. The worker takes actions (a_1, a_2) that influence both an objective performance measure (p) and the worker's total contribution to firm value (y). As discussed above, there may be important differences between y and p. For example, suppose that y and p are binary, with realizations of zero or one, and that the worker's actions affect the (conditionally independent) probabilities of $y = 1$ and $p = 1$ as follows:

$$Prob(y = 1) = f_1 a_1 + f_2 a_2 \quad \text{and}$$

$$Prob(p = 1) = h_1 a_1 + h_2 a_2 \, ,$$

where $f_1/f_2 \neq h_1/h_2$. As one illustration, if a_1 represents effort on the quantity of parts produced and a_2 represents effort on their quality then $f_2 > 0$ and $h_2 = 0$ captures a situation like Lincoln Electric: the piece rate rewards only quantity when the firm would like the worker to focus on quality as well. As another illustration, if a_2 represents the gaming seen at Washington Mutual or Wells Fargo then $f_2 = 0$ and $h_2 > 0$ captures a situation where the worker can increase the performance measure without affecting her actual contribution to the value of the firm.

To explore the combined use of formal and discretionary incentives, our 1994 paper analyzed a model in which $w = s + bp + By$, where bp is enforced by a court but By is at the discretion of the firm—that is, By is a relational contract. The timing of events within each period is as follows. First, the firm offers the worker a compensation package (s, b, B). Second, the worker either accepts the compensation package or rejects it, in which case the worker receives payoff $\overline{u}$ and the firm $\overline{\pi}$, Third, if the worker accepts, then the worker chooses actions (a_1, a_2) at cost $c(a_1, a_2)$. Fourth, the firm and the worker observe the realization of the worker's contribution (y) and the firm and the worker (and, if necessary, a court) observe the realization of the objective performance measure (p). Finally, the firm pays the bonus bp dictated by the formal contract and chooses whether to pay the discretionary bonus By indicated by the relational contract.

Suppose the worker's cost function is $c(a_1, a_2) = \frac{1}{2}(a_1^2 + a_2^2)$. First-best actions maximize the expected value of $y(a_1, a_2) - c(a_1, a_2)$, whereas if the worker believes that the firm will pay the discretionary bonus (as well as the salary and bonus dictated by the formal contract) then the worker chooses actions to maximize the expected value of

$$s + bp(a_1, a_2) + By(a_1, a_2) - c(a_1, a_2)\ .$$

The first-best actions satisfy $a_1 = f_1$ and $a_2 = f_2$, whereas the worker's optimal actions in response to the contract satisfy $a_1 = bh_1 + Bf_1$ and $a_2 = bh_2 + Bf_2$.

First-best actions will be achieved if $B = 1$ and $b = 0$, but this may not be feasible: $B = 1$ may create too strong an incentive for the principal to renege on the relational contract. If the first-best is not feasible, the second-best may involve $B < 1$ and $b > 0$, or the parties may not be able to use relational contracts at all, meaning that $B = 0$ and $b = b^{spot}$ (*i.e.*, the formal contract is the efficient formal contract in the one-shot game). Which of these occurs is determined by the firm's "dynamic enforcement constraint" (Levin, 2003: 842), which we analyze assuming trigger strategies: after any reneging, the parties revert to the efficient formal contract in the one-shot game thereafter.

A novel feature of our model (for its time) is that formal contracting plays two roles: first, the weight *b* on the performance measure *p* influences the worker's actions in the current period; second, the weight b^{spot} in the efficient formal contract in the one-shot game would be used after reneging. To see these two roles in the dynamic enforcement constraint, define $V(b, B)$ as the expected gross surplus created by (b, B) and V^{spot} as the expected gross surplus created by the efficient formal contract in the one-shot game. That is, $V(b, B)$ is the expectation of $y(a_1, a_2) - c(a_1, a_2)$ when the worker's actions satisfy $a_1 = bh_1 + Bf_1$ and $a_2 = bh_2 + Bf_2$, and V^{spot} is that expectation when the worker's actions satisfy $a_1 = b^{spot}h_1$ and $a_2 = b^{spot}h_2$. Importantly, because of the parties' combined outside options $\overline{v} \equiv \overline{u} + \overline{\pi}$, efficient formal contracting alone may not suffice for the firm to attract the worker: we may have $V^{spot} < \overline{v}$.

After *p* and *y* have been realized, the firm is obligated by the formal contract to pay *s* and *bp*, so paying the discretionary bonus produces payoff $y - s - bp - By$ this period and some expected payoff π_C in each future period, whereas not paying the discretionary bonus produces $y - s - bp$

now and some expected payoff π_D thereafter, where the subscripts C and D connote cooperation and defection, respectively. If $V^{spot} < \overline{v}$ then the largest that π_C could be is the expected net surplus $V(b,B) - \overline{u}$ (*i.e.*, net of the worker's opportunity cost), and π_D will be $\overline{\pi}$. If giving the firm the entirety of $V(b,B) - \overline{u}$ will not deter the firm from reneging on the discretionary bonus, then nothing will persuade it to pay the bonus, so we have the necessary condition:

$$\{y - (s + bp + By)\} + \frac{1}{r}[V(b,B) - \overline{u}] \geq \{y - (s + bp)\} + \frac{1}{r}\overline{\pi}.$$

This necessary condition must hold for all y, for which it suffices that it to hold at $y = 1$, in which case the condition simplifies to

$$B \leq \frac{1}{r}(V(b,B) - \overline{v}). \qquad (1)$$

This necessary condition is now standard in the literature. Furthermore, as is now widely known (but should be celebrated nonetheless), this necessary condition is also sufficient, providing a full characterization of the relational contracts (here, trigger-strategy equilibria of the repeated game) that are feasible in this model; see MacLeod and Malcomson (1989) for the first version of this powerful result.

Substantively, the novel aspect of our 1994 paper was our inclusion of both formal and relational contracting in incentive schemes. Analytically, this novelty arises through the two roles of formal contracting described above: b on the equilibrium path and b^{spot} off. The former arises in the $V(b,B)$ term in (1) above; to explore the latter, we next suppose that $V^{spot} \geq \bar{v}$. In this case, the largest that π_C could be is again $V(b,B) - \overline{u}$ but now π_D will be $V^{spot} - \overline{u}$, so the necessary condition becomes

$$\{y - (s + bp + By)\} + \frac{1}{r}[V(b,B) - \overline{u}] \geq \{y - (s + bp)\} + \frac{1}{r}[V^{spot} - \overline{u}],$$

which simplifies to

$$B \leq \frac{1}{r}(V(b,B) - V^{spot}). \qquad (2)$$

Combining the necessary conditions (1) and (2) then yields

$$B \leq \frac{1}{r}(V(b,B) - \max\{V^{spot}, \overline{v}\}), \qquad (3)$$

which is the core analytical result of our 1994 paper: as is now familiar, the relationship must create enough net surplus for the promised bonus to be credible; more specific to our setting, with its combination of formal and relational contracting, the net surplus is relative to max $\{V^{spot}, \bar{v}\}$— the larger of the gross surplus from efficient spot contracting and the sum of the parties' outside options.

Figure 1 shows how the optimal incentive strengths and associated gross surpluses vary with the distortion in the objective measure, p. The top panel shows the optimal b and B (both ranging between 0 and 1) for contracts based on only the objective performance measure (b-only), contracts based on only the discretionary measure (B-only) and contracts based on both b and B. The bottom panel shows $V(B)$) and $V(b, B)$—the expected gross surpluses created by the contract based on only B and by the contract based on both b and B, respectively. These results depend on three parameters: the parties' discount rate r (assumed fixed in the Figure), the parties' combined outside options $\bar{v}$, and the distortions created by the objective performance measure. Following Baker (2002), we summarize the distortion via the parameter θ, which takes its minimum value when $\frac{f_1}{f_2} = \frac{h_1}{h_2}$ and its maximum value when the **f** and **h** vectors are orthogonal; see Appendix 1.

Figure 1 illustrates four substantive results more general than the Figure's specific parameters. In 1991, we expected only the first of these results, and the last two led us to the new line of inquiry described in Section 3.

Result 1: For performance measures with intermediate distortion, the formal contract (b) and the relational contract (B) are substitutes. Simply put, the more that appropriate incentives can be created via By, the less the firm needs to use the distorted incentive bp. More specifically, the firm uses the relational contract By as much as it can, limited by the dynamic enforcement constraint; and the firm then addresses the remainder of the incentive problem, of size (1-B), with the formal incentive contract, putting smaller weight on the objective performance measure if it creates more distortions.

Result 2: For high levels of distortion, where $V^{spot} < \bar{v}$, the formal and relational contracts are complements: the strengths of both decline as the performance measure gets worse. In this region, an improved formal contract (a move to the left in the figure) increases the value of the relationship relative to the fallback (which is fixed at $\bar{\pi}$ for the firm, because formal contracting alone cannot attract the worker), thereby allowing the relational incentive strength B to be higher without violating the dynamic enforcement constraint. In short, when $V^{spot} < \overline{v}$, the more that can be done via bp, the more that can be done via By.

Result 3: For low distortion, *no relational contract is possible.* Strikingly, even mediocre formal contracts can prevent the existence of all relational contracts, including relational contracts that would be superior to the formal contract. Result 3 is the opposite of Result 2: now the possibility of formal contracting after reneging on the relational contract *reduces* the value of the relationship, making only weaker relational incentives possible (and, for many parameters, making no relational incentives possible).

Result 4: For intermediate distortion, the firm may be better off if it were not possible to use p at all, instead relying solely on relational contracting (without the possibility of formal contracting after reneging). Defining $V^{rel}(B)$ to be the expected value of $y(a_1, a_2) - c(a_1, a_2)$ when the worker's actions satisfy $a_1 = Bf_1$ and $a_2 = Bf_2$, for intermediate values of distortion we can have $V^{rel}(B) > V(b, B)$: the possibility of formal contracting after reneging does more harm than the possibility of combining formal and relational contracting does good.[5]

We did not expect Results 2-4 when we began this work; after discovering them, we thought each could be developed in interesting directions. In particular, in Result 3 we heard echoes of Coase's (1937) and Williamson's (1975) work on the boundary of the firm. The distortions induced by the formal contract *bp* seemed potentially related to the difficulties in contracting that Coase and Williamson argued might cause parties to prefer integration to the "market failure" of non-integration. And the relational contract *By* seemed potentially related to Williamson's discussion of "atmosphere" within firms (in Chapter 2) and of Simon's (1951) model of the employment relationship (in Chapter 4): in these chapters, Williamson emphasized relational aspects of firms and employment as a reason that parties might prefer integration to the contracting difficulties of non-integration.

[5] For related analyses of ostensibly useful changes—such as the introduction of markets or money into informal trading environments—that can have adverse effects because fallback payoffs increase and relational contracting becomes more difficult, see Kranton (1996), Prendergast and Stole (1999), Di Tella and MacCulloch (2002), and Dhillon and Rigolini (2011).

In these terms, our Result 3—that mediocre formal contracts can prevent the existence of all relational contracts—might be interpreted as suggesting that a mediocre market can prevent the existence of what could have been a more effective firm. Of course, such an interpretation would equate non-integration with formal contracting, allowing relational contracting only within firms. To consider these issues on a solid footing (including allowing relational contracting between non-integrated parties), we built our 2002 model discussed in Section 3. Before turning to that discussion, however, we briefly discuss our 1999 model, in which empowerment, like discretionary bonuses, requires a relational contract.

2.2 Informal Authority

While our 1994 paper was intended as a contribution to agency theory, unexpected theoretical results in that paper pointed us towards new domains, such as the analysis of firms' boundaries launched by Coase and Williamson. But our 1994 paper did more than expand our research focus from agency theory towards other aspects of organizational economics. More importantly, this paper also opened our eyes to how things often get done in organizations: *managers make promises*.

In this sub-section we briefly sketch an example of a managerial process beyond incentive schemes that again depends on relational contracting. As noted in the Introduction, we use our 1999 paper for this purpose, thus illustrating a second model of *relational agency*—now one where the principal can create incentives for the agent by promising to implement projects discovered and proposed by the agent, provided that the project's payoff is not too negative for the principal.[6]

To analyze the incentives created by such promises, we began from an economic environment similar to Aghion and Tirole's (1997) static model. Consider a corporate parent and a product-development lab owned by the parent. If the lab incurs cost $c(a)$ then either the lab develops a new product (with probability a) or not (with probability $1 - a$). If a new product is developed, the parent can then market it, in which case the lab receives x and the parent receives y; otherwise, both parties receive zero. Suppose that the lab always likes its products marketed ($x > 0$) but a new product could either cannibalize or complement the parent's existing products. Specifically,

[6] This exposition of our 1999 paper borrows from Gibbons (2010).

y takes three values, $y_H > 0 > y_M > y_L$, and conditional on a new product being developed, the probabilities of different values of y are independent of the lab's choice of a.

In a one-shot setting, the parent will market a product only if $y > 0$ ($i.e.$, $y = y_H$). Denote the associated expected gross surplus by V_H. In addition, let $\overline{u}$ and $\overline{\pi}$ denote the lab's and the parent's expected payoffs if the lab works on its own projects rather than working for the parent in a given period, where $\overline{u} + \overline{\pi} = \overline{v}$.

Of course, things can be different in an ongoing relationship. If the parent's credible promise changed from "we will market all products with $y \geq y_H$" to "we will market all products with $y \geq y_M$" the lab's effort would increase. Indeed, the lab's effort would increase further if the parent's credible promise changed to "we will market all products with $y \geq y_L$."

Consider a relational contract in which the parent promises to market products with $y \geq y_M$, and let V_{HM} denote the expected gross surplus from this contract. The parent will be tempted *not* to market a product with $y = y_M$, and the size of the parent's reneging temptation will be $-y_M > 0$, analogous to B in Section 2.1. This relational contract is an equilibrium of the repeated game if

$$-y_M \leq \frac{1}{r}(V_{HM} - \max\{V_H, \overline{v}\}), \tag{4}$$

analogous to condition (3) in Section 2.1.

One might describe the credible promise "market $y \geq y_M$" (*i.e.*, implementing much but not all of what the lab discovers) as limited empowerment of the lab and "market $y \geq y_L$" (*i.e.,* implementing all of what the lab discovers) as full empowerment. That is, in our approach, the extent of empowerment is defined by the worst of the products (from the parent's perspective) that the parent promises to market—here, y_M or y_L—which then determines (i) the expected net surplus of the relationship, (ii) the parent's maximum reneging temptation, and, hence, (iii) the credibility of the promise. As is now widely appreciated, it is the interplay between (i) and (ii), subject to (iii), that causes relational contracting in repeated games to produce different results those of formal contracting in static models; we return to this below.

In our 1999 paper, we postulated that, *within* organizations, "subordinates' decision rights [are] loaned, not owned" (1999: 56), in the sense that the superior can revoke the subordinate's control over future decisions (and, for some kinds of decisions, the superior can overrule the subordinate's current decision). We therefore interpret empowerment within organizations to be a credible promise by the superior, not a formal contract enforceable by the court. As a result, we interpret allocating formal authority to the lab to be spinning off the lab's activities into a new firm that the lab owns—transforming the parties' relationship from a boss and subordinate within one organization to something like an alliance between two organizations.

Suppose such a spin-off has occurred, so the lab has formal authority over marketing. In a one-shot setting, the lab will market a product for any y because $x > 0$. Denote the associated expected gross surplus by V_{HML}, which could be greater or less than the expected gross surplus in a one-shot setting where the parent has formal authority, V_H. As in Section 2.1 and throughout this essay, we assume not only trigger strategies in our repeated-game analyses but also that the parties revert to efficient spot governance after reneging—here, with expected gross surplus $V^{spot} = max\{V_H, V_{HML}\}$.

After a spin-off has occurred, suppose the parties attempt to implement the same relational contract analyzed above: market products with $y \geq y_M$. The lab will be tempted to market a product with $y = y_L$, and the size of the lab's reneging temptation will be $x > 0$. This same relational contract is thus an equilibrium of the repeated game under this new allocation of formal authority if

$$x \leq \frac{1}{r}(V_{HM} - \max\{V^{spot}, \overline{v}\}). \tag{5}$$

And given that the parties will revert to efficient spot governance after reneging, (4) becomes

$$-y_M \leq \frac{1}{r}(V_{HM} - \max\{V^{spot}, \overline{v}\}). \tag{6}$$

At sufficiently low discount rates, the relational contract "market products with $y \geq y_M$" is feasible under either allocation of formal authority. However, when $x < -y_M$, there is a range of discount rates where (5) holds but (6) does not, so the relational contract is feasible only when the lab has formal authority; analogously, when $x > -y_M$, there is a range of discount rates where the relational contract is feasible only when the parent has formal authority. This model thus provides a rationale for choosing the formal allocation of control to facilitate a relational contract—illustrating the idea of *structuring relationships* that we discuss in Section 3.

To conclude this sub-section, we compare our relational model of empowerment to Aghion and Tirole's (1997) static model and draw more general lessons about differences between relational and static models. In our model, formal delegation to the lab is a spin-off, creating a new organization, whereas Aghion and Tirole allow formal delegation within an organization: the boss conveys all control over project selection to the subordinate, but the boss remains the subordinate's boss. In both models, the lab's effort increases if formal authority moves from the parent to the lab. Thus, the *consequences* of moving formal authority are similar, but the *interpretations* of how formal authority can be moved are very different: a spin-off in our case, and a formal contract within an organization in theirs. Accordingly, the two depictions of empowerment *within* organizations are also very different. In our model, the level of empowerment is defined by the worst of the products that the parent will market (y_M or y_L), and the best feasible level of empowerment is determined by (i)-(iii) above: the interplay between the expected net surplus of the relationship and the parent's maximum reneging temptation, subject to the credibility of the promise. In contrast, in their model, even if enriched to allow different levels of empowerment, temptations play no role, leading the two models to make different predictions about when empowerment will occur, how it will be managed, and what outcomes it will produce.

The contrast between relational and static models is more general than this comparison of our relational model of empowerment to Aghion and Tirole's static one. As a striking example, Levin (2003: 842) shows that in a hidden-information setting often analyzed using static models of mechanism design, if the principal cannot commit to the mechanism but instead makes a promise of the kind studied in the repeated games discussed in this essay, two of the most familiar results from static mechanism design disappear (no distortion at the top, and separation of all types). In short, static models are not a reliable reduced-form for relational analyses.

3. Structuring Relationships

While the primary focus of our 1994 and 1999 papers was relational contracting inside an organization, between a boss and a subordinate, we came to appreciate that relational contracts also play important roles between organizations. Such conscious coordination between organizations contradicts Robertson's (1923: 85) early image of firms as "islands of conscious power … like lumps of butter coagulating in a pail of buttermilk." And Coase's (1937) sharp distinction between firms and markets perpetuated thinking like Robertson's for many decades thereafter.[7]

Eventually, reality began to intrude. For example, Richardson (1972: 895) argued that "Firms are not islands but are linked together in patterns of co-operation and affiliation. Planned co-ordination does not stop at the frontier of the individual firm but can be effected through co-operation between firms." Over the next several decades, a few scholars made important progress on this subject.[8] Nonetheless, most doctoral textbooks in microeconomics continued to see collusion as the leading example of relationships between firms.

In this section we discuss our 2002 paper (in Section 3.1) and our 2011 paper (in Section 3.2). While these papers had different motivations and analyzed different models, we now see their similarities as their most important lesson. For example, in both papers, we find that, given a particular relational contract, the optimal formal governance structure for that contract minimizes the maximum aggregate temptation to renege on that contract. In addition, both papers surface versions of the two roles for formal governance introduced in Section 2: in this section, the current formal governance structure can influence the parties' reneging temptations this period, and the prospect of a formal governance structure (perhaps a different one) after reneging can influence the net expected total payoff from the relationship.

3.1. Relational Contracts and Firm Boundaries

Having come to appreciate the productive uses of relationships between organizations, two questions spurred us to build an initial model. First, we wondered if the possibility of relational contracting between organizations should affect their boundaries. That is, two firms engaged in relational contracting could have merged, and two divisions could have dis-integrated. When should two parties structure their relational contract under integration and when under non-integration? And second, recalling Result 3 from our 1994 paper—that mediocre formal contracts can prevent the existence of relational contracts—we wondered if the possibility of mediocre outsourcing could play a similar role: prevent the existence of what would have been a more effective firm (had relational contracting been feasible within the organization).

---

[7] Interestingly, Coase asserted that "it is not possible to draw a hard and fast line which determines whether there is a firm or not" (p. 392), but this aspect of his seminal article is rarely noted.

[8] For example, see Klein and Murphy (1988), Klein (1996) and Klein (2000).

We analyzed these two questions in our 2002 paper. Specifically, we considered an economic environment similar to Grossman and Hart (1986), consisting of an upstream party, a downstream party, and an asset used by the upstream party to produce an intermediate good that could be used by the downstream party. We imagined that the downstream party already owned an unmodeled second asset, so if the downstream party also owned the modeled asset then the downstream party had integrated control over both assets, whereas if the upstream party owned the modeled asset then control was non-integrated.

We defined ownership of the modeled asset as conveying ownership of the intermediate good. Thus, if the downstream party owns the modeled asset then he can simply take the intermediate good, refusing to pay the upstream party anything for it. Alternatively, if the upstream party owns the asset, she can put the intermediate good to an alternative use (*e.g.*, use it herself, or sell it to a third party), with no repercussions from the downstream party.

We assumed that, after the intermediate good is produced but before it is used, both the upstream and the downstream parties (but no courts) observe the values of the intermediate good both to the downstream party and in its alternative use. In current parlance, the values of the intermediate good are observable but not contractible. What *is* contractible, however, is ownership of the modeled asset. That is, in our 2002 paper, ownership of the modeled asset defines the candidate formal governance structures for the upstream and downstream parties' interaction: integration (*i.e.*, the downstream party owns the modeled asset) or non-integration (*i.e.*, the upstream party owns the asset).

We analyzed whether, in a repeated game, it is optimal for the asset to be owned by the downstream or upstream party. For ease of exposition, here we borrow notation from the model in Section 2.1 (as opposed to the notation in our 2002 paper). We solved for the optimal relational contract, $B_I(y)$, under integration and the optimal relational contract, $B_{NI}(y)$, under non-integration. Under integration, the downstream party already owns the intermediate good, so the bonus $B_I(y)$ is an incentive contract, rewarding the upstream party for delivering an intermediate good that the downstream party values, as in our 1994 model. In contrast, under non-integration, the upstream party initially owns the intermediate good, so the bonus $B_{NI}(y)$ is a sales contract, rewarding the upstream party for creating the good and transferring it to the downstream party.

The main analytical result from the model (obvious in retrospect) was that a given relational contract—say, some $B^*(y)$—typically has different reneging temptations under integration than under non-integration. Under integration it is of course the downstream party who is tempted to renege because he can simply take the intermediate good; in contrast, under non-integration it may be the upstream party who is tempted to renege if the value of the intermediate good's alternative use is unusually high.

The result that a given relational contract has different reneging temptations under different formal governance structures echoes our finding in Section 2.2 (discussing our 1999 paper) concerning formal authority. This result from our 2002 paper taught us to turn the make-or-buy decision on its head: instead of asking whether integration or non-integration is preferred, and then fleshing out the details of the preferred formal governance structure, the model suggests that we ask what relational contract would be most effective, and then choose the governance structure (integration or non-integration) to minimize the reneging temptation. In short, in this model *the integration decision is made in the service of the relationship*.[9]

Our 2002 paper also delivered several additional results, beyond this main one. We conclude this sub-section by summarizing some of them informally.

First, there are applications of the main result. For example, a relational contract with high-powered incentives has a larger reneging temptation under integration than under non-integration—because under integration the downstream party can renege on all promised incentives yet still own the intermediate good, whereas under non-integration the upstream party can insist that the downstream pay at least the intermediate good's value in its alternative use. Likewise, greater variation in the extreme values of the intermediate good's alternative use create larger reneging temptations under non-integration than under integration—because under non-integration the upstream party can threaten to put the intermediate good to its alternative use, whereas under integration the downstream party can ignore the value of the alternative use. These findings relate to observations that (a) incentives are typically lower-powered in firms than in markets and (b) supply assurance can motivate vertical integration, respectively.

[9] For related work on structuring relationships, see Halonen (2002) and subsequent work described in Section 10 of Malcomson (2013).

Second, there are results that formalize and enrich intuitions and examples from our earlier work. For example, in the spirit of Result 3 from our 1994 paper, but now allowing relational contracts both within and between firms, the 2002 paper shows that the possibility of mediocre non-integration (which one might call a mediocre spot market) could indeed prevent the existence of what would have been a more effective firm, had relational contracting been feasible within the firm. In addition, as can be illustrated in our discussion of the 1999 paper above—specifically, inequalities (5) or (6)—the 2002 paper shows that the optimal choice of formal governance structure can depend on the extent to which the parties can use relational contracting (modelled as the discount rate, r).

Finally, the 2002 paper produced its own example that we formalized and enriched in other work. In the 2002 paper, we showed that it is impossible for a firm to mimic the spot-market outcome. More precisely, even if some relational contracting is possible under integration, one relational contract that is never feasible under integration is the spot outcome under non-integration. We elaborated on this idea in our 2001 paper, "Bringing the Market Inside the Firm?", which questioned the presumption that the transfer-pricing problem could be solved by charging market prices for internal transactions, or that the capital-allocation problem could be solved by charging the external cost of capital for internal investments. We generalized this idea in our 2011 paper, to which we turn next.

3.2. Relational Adaptation

Our 2002 paper was intended as a contribution to the theory of the firm—construed narrowly as the study of firms' boundaries. One of our motivations for writing that paper was the realization that there are important relational contracts both within organizations and between organizations. Having written the 2002 paper, however, we realized that it did not address another key aspect of the governance of relationships between organizations: formal contracts.

To understand the drivers of such formal contracts, we conducted detailed discussions with practitioners. For example, Judy Lewent (then CFO and EVP of Merck) summarized her firm's work with many other firms in the biotech-pharmaceuticals industry by saying "Our best deals are those where we put the contract in a drawer and build a relationship on top."[10] That is, unlike our

[10] Personal communication.

1994 paper, where payment of the term *bp* surfaces the formal contract each period, important aspects of these formal contracts set long-run governance structures, such as which party controls the marketing of new products produced by the partnership. What then matters for value creation is how the parties exercise these control rights, not merely who has them.

In addition, while our 2002 paper focused on *ex ante* incentive problems caused by specific investments, practitioners rarely mentioned these and other problems familiar from the academic literature. Instead, they focused on problems of efficient *ex post* adaptation to changing circumstances, beyond what could be foreseen in formal contracts. Of course, this adaptation problem has a long history in economics, such as Hayek's (1945: 521) emphasis on "rapid adaptation to changes in the particular circumstances of time and place" or Simon's (1951) model of an employment relationship where the boss makes decisions after uncertainty is resolved. These two ideas—"contracting for control" in the service of a relational contract, and efficient adaptation—motivated our 2011 paper.

More specifically, in our 2011 paper we developed and analyzed a model of *relational adaptation*, where parties can negotiate formal governance structures that allocate particular decision rights to particular parties but must rely on relational contracts to influence how those who have control decide to exercise it. In our 1994, 1999, and 2002 papers described in Sections 2.1, 2.2, and 3.1, our interest was in (a) how relational contracts might mitigate various problems in formal contracting and (b) how formal governance structures might then be chosen not for their direct effects (as analyzed in static models) but instead to facilitate relational contracting. In those papers we therefore were happy to borrow (or modify) rather than invent static models to launch our relational analyses—Baker (1992) in 1994, Aghion and Tirole (1997) in 1999, and Grossman and Hart (1986) in 2002. We might have done the same in our 2011 paper, but there was no static model to be borrowed, so we had to invent one.

Our 2011 paper began with an elemental model of adaptation where two parties determine which of them should control a single decision right. In this sub-section, we sketch five aspects of the 2011 paper (which is attached as Appendix 2): (1) the elemental model of adaptation under contracting for control; (2) relational contracting in a simplified example of that elemental model; (3) costly formal contracting in that example; (4) Williamson's (1991) assertion that courts exercise "forbearance" in hearing disputes within firms; and (5) the general adaptation model in our 2011 paper and its main result.

3.2.1 Elemental Model of Adaptation

Consider a single decision right that can be assigned to either of two parties, A or B. The parties are risk-neutral and have private benefits, π_A and π_B, that depend on a state of nature, s, and also on the decision, d. The first-best decision in state s, denoted $d^{FB}(s)$, maximizes $\pi_A(s,d) + \pi_B(s,d)$, producing a total payoff of $V^{FB}(s) = \pi_A(s, d^{FB}(s)) + \pi_B(s, d^{FB}(s))$. Define $V^{FB} = E_s[V^{FB}(s)]$ as the expected total payoff from first-best decision-making.

The timing of the one-shot model is as follows. First, the parties negotiate control over the decision right. Second, the parties observe s, drawn from the finite set S according to the distribution $f(s)$. Third, the party with control chooses d, from the finite set D. Fourth, the parties receive their private benefits, π_A and π_B. Note that there is no time (or opportunity for contract enforcement) between the second and third stages either to change who has control over the decision or to contract on what decision should be taken by the party with control.

In a spot setting, if party i ($i =$ A, B) holds the decision right then in state s party i will choose the decision denoted $d_i(s)$ that maximizes its private benefit $\pi_i(s,d)$, producing a total payoff $V^i(s) = \pi_A(s, d_i(s)) + \pi_B(s, d_i(s))$. Define $V^i = E_s\left[V^i(s)\right]$ as the expected total payoff from spot decision-making by party i, and $V^{spot} = \max\{V^A, V^B\}$ as the expected total payoff in a spot setting when control is allocated efficiently.

3.2.2 Relational Contracting in a Simple Example

Suppose that in any given state the only two decisions of interest are $d_A(s)$and $d_B(s)$. More specifically, the private benefit to party i is positive when $d_i(s)$ is chosen, zero when $d_j(s)$ is chosen, and very negative for all other decisions. Denote the benefit to party i when $d_i(s)$ is chosen by $\pi_i(s) = \pi_i(s, d_i(s))$.

Figure 2 illustrates a setting where $s \in [0,1]$ and the size of the parties' disagreement about the decision increases with s: the benefit curves $\pi_A(s)$ and $\pi_B(s)$ are both increasing. The first-best decision is $d_B(s)$ when $s < s^*$ and $d_A(s)$ when $s > s^*$, so neither spot governance structure achieves first-best.

We now analyze whether the first-best can be achieved via a relational contract when both parties discount at rate r. Suppose that party A controls the decision. For $s > s^*$, first-best decision-making asks party A to choose $d_A(s)$, which A is of course happy to do. For $s < s^*$, however, A is asked to choose is $d_B(s)$, which produces a benefit of zero for A when A could choose $d_A(s)$and receive $\pi_A(s) > 0$. Relative to first-best decision-making, A's maximum reneging temptation is thus $\pi_A(s^*)$.

As in all the papers discussed in this essay, we assume (a) trigger strategies and (b) reversion to efficient spot governance (with expected total payoff V^{spot}) after reneging. Analogous to the other models discussed above, we then have that first-best decision-making can be achieved via a relational contract when A controls the decision if

$$\pi_A(s^*) \leq \tfrac{1}{r}(V^{FB} - V^{spot}). \tag{7}$$

Suppose instead that party B controls the decision. Now the relational contract must induce B to choose $d_A(s)$ when $s > s^*$. Relative to first-best decision-making, B's maximum reneging temptation is thus $\pi_B(1)$. First-best decision-making thus can be achieved in a relational contract when B controls the decision if

$$\pi_B(1) \leq \tfrac{1}{r}(V^{FB} - V^{spot}). \tag{8}$$

Because $\pi_B(1) > \pi_A(s^*) > 0$, there is a range of r where (7) holds but (8) does not. For this range of r, assigning control to A can achieve the first-best, while assigning control to B cannot. Note that nothing in this analysis specifies whether V^{spot} is achieved by giving A or B control. If the latter, then we have the result that the optimal formal governance structure depends on r: A should have control for the range of r where (7) holds but (8) does not, but B should have control for high enough values of r that only spot governance is feasible.

When r is sufficiently high that the first-best cannot be achieved, the parties may still be able to implement a second-best decision rule that outperforms optimal spot governance. In the environment illustrated in Figure 2, when the second-best decision rule outperforms spot governance, the second-best takes the intuitive form depicted in Figure 3: control is again assigned to A, but now the decision rule implements $d_A(s)$ rather than $d_B(s)$ between some s' and s^*.

The logic behind Figure 3 is both simple and appealing. Given r, the first-best cannot be implemented because A's maximum reneging temptation, $\pi_A(s^*)$, is high enough that (7) fails. The second-best decision rule therefore must reduce A's maximum reneging temptation, by letting A choose $d_A(s)$ when A is most tempted to do so, starting with states s just below s^*. But allowing A to choose $d_A(s)$ when $s < s^*$ reduces the expected total payoff below V^{FB}.

Define $V(s')$ as the expected total payoff from the decision rule shown in Figure 3—namely, first-best except for $d_A(s)$ between s' and s^*. If A has control and is supposed to implement this decision rule, A's maximum reneging temptation is $\pi_A(s')$. The usual argument then implies that this decision rule can be implemented via a relational contract with A controlling the decision if

$$\pi_A(s') \leq \tfrac{1}{r}(V(s') - V^{spot}). \tag{9}$$

In summary, the form of a potential second-best relational contract (that outperforms spot governance) is clear, but whether it exists depends on a race between the need to reduce A's maximum reneging temptation (which falls as the lowest state at which A is intended to take $d_A(s)$ falls from s^* to s') and the consequent reduction in the expected total payoff from the relational contract (caused by the inefficient choice of $d_A(s)$ in these states).

3.2.3 Costly Formal Contracting in the Example

An important extension of the elemental adaptation model is to allow costly contracting after s has been observed but before d needs to be chosen. We formalize such *ex post* contracting by assuming that at finite cost $c > 0$ either party can offer the other a formal contract (d, p). In equilibrium, it is the party without control (say, B) who may choose to incur c and offer such a contract to the party with control (A). If A signs the contract then A must choose decision d and B must pay p to A.

In a one-shot setting, as shown in Figure 4, if A has control then B will incur c when $\pi_B(s) - \pi_A(s) > c$ (with total payoff indicated by the red line at the left of the figure), and vice versa if B has control (indicated by the red line at the right). Henceforth we therefore write $V^{spot}(c)$, where V^{spot} in (7), (8), and (9) is $V^{spot}(\infty)$.

The introduction of *ex post* contracting at finite cost improves the expected total payoff of spot adaptation from $V^{spot}(\infty)$ to $V^{spot}(c)$. Unfortunately, an increase in V^{spot} can harm relational contracting. For example, suppose that if costly contracting is not possible (*i.e.*, $c = \infty$) then (9) holds with equality, so the parties are just able to implement the second-best relational contract shown in Figure 3. Now suppose that improvements in contracting technology reduce c to the level shown in Figure 4. This increase in V^{spot} causes (9) to fail. As a result, the form of second-best relational contracting is now as shown in Figure 5: the lowest state at which A is intended to take $d_A(s)$ falls from s' to $s''(c)$, so decision-making is inefficient in more states.

The result that a reduction in c may harm relational contracting and reduce the parties' expected total payoff is similar to Result 4 in Section 2.1, where the parties may be better off if it were not possible to use the formal incentive contract bp at all, instead relying solely on relational contracting (without the possibility of a formal incentive contract $b^{spot}p$ after reneging). As noted in connection with Result 4, there are many settings where a seemingly useful change in the economic environment can in fact make the parties worse off.

A further reduction in c may cause a second kind of harm to the parties: now c may have to be paid on the equilibrium path, whereas in Figure 5 relational contracting allowed the parties to avoid paying c. The possibility that c will have to be paid on the equilibrium path leads to our next topic.

3.2.4 Williamson (1991) on "Forbearance"

The discussion of costly contracting in Section 3.2.3 followed Grossman and Hart (1986) and many others by assuming that the feasible set of formal contracts (including any cost of contracting) is constant across alternative governance structures. For the case of two firms writing a formal contract to allocate control of a particular decision to one of them, this assumption seems natural: the cost of *ex post* contracting should not depend on which of them controls the decision. But Williamson (1991) took a different stance in discussing courts' "forbearance" in hearing disputes within firms, arguing that contracts that are court-enforceable between firms are often not enforceable within firms.[11]

[11] To hear a contract dispute, the courts need to see a contract, which requires two contracting parties, but in a dispute between two divisions of a given firm, the court cannot see a contract if it sees only one party—the firm. Of course, this discussion pertains to interactions between units of the firm such as divisions, not to legal contracts between the firm and parties such as employees.

In our 2011 paper we explored the possibility that contracts between firms have finite c and hence produce $V^{spot}(c)$ in one-shot settings, whereas interactions within firms occur in contract-free zones and hence produce $V^{spot}(\infty)$ in one-shot settings. It therefore is at least as efficient to conduct a one-shot interaction under non-integration, but we show that in a repeated game it can be efficient to conduct the parties' interaction under integration (even though, were there to be reneging, efficient spot governance would imply that the parties would dis-integrate).

The reason that infinite c may be more efficient than finite c in a repeated game is related to the new kind of harm caused by the further reduction in c from Figure 5: c is paid on the equilibrium path under non-integration. Under integration, in contrast, c is infinite, as in Figure 3, but integration differs from Figure 3 because the parties will dis-integrate after reneging to achieve efficient spot governance, whereas Figure 3 assumed that $c = \infty$ after reneging. Put differently, Williamson's idea of forbearance separates the two roles for formal contracting discussed several times above, such as b on the equilibrium path and b^{spot} off in Section 2.1: integration affects only the former.

As we show in our 2011 paper, integration may facilitate a relational contract superior to non-integration's relational contract when the latter has c paid on the equilibrium path. The advantage of integration's relational contract is that c is not paid. The disadvantage is that A is intended to choose $d_A(s)$ in even more states where choosing $d_B(s)$ would be efficient, but we show that the former can outweigh the latter. As far as we know, this is the first formal exploration of Williamson's (1991) assertion that contract law differs within firms than between—*i.e.,* forbearance.

3.2.5 General Model and Main Result

In the general model in our 2011 paper we drop the assumptions of Section 3.2.1 (namely, two parties and one decision right) and of Section 3.2.2 (namely, that $d_A(s)$and $d_B(s)$ are the only decisions of interest and that the benefit functions are as shown in Figure 2). Instead, we suppose there are I parties denoted $i \in I = \{1, \ldots, I\}$ and K decision rights denoted $k \in K = \{1, \ldots, K\}$. If decisions $\boldsymbol{d} = (d_1, \ldots, d_K)$ are taken in state s, party i receives a private benefit of $\pi_i(\boldsymbol{d}, s)$.

Define a governance structure $g \in G$ to be an assignment of the K decision rights to the I parties, assuming there is no joint control of any decision right and there is no decision right left uncontrolled. There are then I^K feasible governance structures. For simplicity, we assume that each has a unique Nash equilibrium in the one-shot game (different for each governance structure), and V^{spot} is then the maximum over the feasible set of governance structures of the expected total payoffs from these Nash equilibria.

In the spirit of all the numbered inequalities in this essay, in our 2011 paper we showed that a relational contract implementing decision rule $\boldsymbol{d}^{RC}(s)$ exists under governance structure g if and only if

$$R_g(\boldsymbol{d}^{RC}) \leq \tfrac{1}{r}(V(\boldsymbol{d}^{RC}) - V^{spot}) \,, \tag{10}$$

where $R_g(\boldsymbol{d}^{RC})$ is the maximum (across states s) aggregate (across parties i) reneging temptation created by $\boldsymbol{d}^{RC}$under g, and $V(\boldsymbol{d}^{RC})$ is the expected total payoff from $\boldsymbol{d}^{RC}(s)$.

Inequality (10) parallels results in MacLeod and Malcomson (1989) and Levin (2003). But those models allow just one formal governance structure, whereas in Sections 2.2 and 3.1 and here we allow multiple governance structures and hence focus on how the formal might facilitate the relational. For our purposes, a key observation is that the right-hand side of (10)—the present value of the net expected payoff that the decision rule $\boldsymbol{d}^{RC}(s)$ creates relative to optimal spot governance—is independent of g. (This independence follows from our assumptions of trigger strategies and efficient spot governance after reneging.) Therefore, the governance structure that can implement the decision rule $\boldsymbol{d}^{RC}(s)$ at the highest value of r is the one that minimizes the left-hand side of (10).

To conclude this section, we note that (10) substantially generalizes the last result from our 2002 paper discussed in Section 3.1—namely, that it is impossible for a firm to mimic the spot-market outcome. Now we can see that it is impossible for governance structure g to mimic the spot decision rule from governance structure g'—because the lefthand side of (10) would then be positive and the righthand side would be zero or negative. This result suggests an image of the decision rules that relational contracting can implement (for a fixed value of r): for governance structure g, relational contracting can implement a set of decision rules surrounding the Nash equilibrium of g, perhaps overlapping with the set of decision rules that can be implemented under some governance structure g', but not including the Nash equilibrium of g'. In practical applications, we suspect that this image gives at least rough guidance for choosing a formal governance structure in hopes that a particular relational contract might then be feasible: pick the governance structure with Nash equilibrium closest to the desired relational contract.

4. Conclusion

As we described in the Introduction, our 1994 paper launched the three of us on an unexpected journey—exploring formal and relational contracting from problems of incentives, through control, to adaptation. We conclude this essay with four observations that begin with the 1994 paper and then echo throughout our journey.

First, the 1994 paper was directly motivated by the world—specifically, the discretionary bonuses at Lincoln Electric. Some of the motivations for the 1999 and 2002 papers were similar, but the best example in this spirit was the 2011 paper, where discussions with practitioners greatly influenced both our substantive focus and our analytical approach.

Second, the 1994 paper opened our eyes to the informal aspects of organizations.[12] Kreps (1990) had used repeated games to launch the study of possibly the most important informal aspect of organizations: culture. Our 1994 paper surfaced a complementary point: "Even ostensibly formal processes such as compensation, transfer pricing, internal auditing, and capital budgeting often cannot be understood without consideration of their associated informal agreements" (BGM 2002: 39). And our 1999 paper went further: some aspects of organizations, such as empowerment, are sometimes taken to be formal, but may instead be entirely informal.

Third, having opened our eyes to informal aspects *within* organizations, it was perhaps natural that the 1994 paper led us to consider informal aspects in dealings *between* organizations. This happened in two ways. First, there was an empirical analogy: if relational contracts are so widespread within organizations, is it really plausible that they are non-existent between organizations? And second, there was a parallel theoretical perspective: when we asked ourselves whether Result 3 from the 1994 paper suggested anything about firms' boundaries, it seemed

[12] Of course, others long preceded us in appreciating this aspect of organizations. As Granovetter (1985: 502) summarized decades of work in sociology, "The distinction between the 'formal' and the 'informal' organization of the firm is one of the oldest in the literature … ."

inappropriate to imagine that relational contracts were the exclusive preserve of firms—*i.e.*, that markets had access to only formal contracts.

Finally, and most specific to our work (in the sense that the three previous observations also apply to much of the relational-contracts literature more generally), the 1994 paper taught us that the formal and informal aspects of organizations not only co-exist but interact.[13] As a first example of such interactions, in the 1994 paper the formal incentive contract and the relational incentive contract can be substitutes or complements.

Much more generally, the 1994 paper sparked a sequence of lessons about how the presence, quality, and/or choice of formal instruments—such as distortionary performance measures, formal delegation of authority, asset ownership, or contracts for control—can limit or support relational contracting. In addition, the parties' ability to use relational contracts affects the optimal choice of such formal instruments. These interactions are the common theme of our papers.

In this essay we described our journey from the 1991 motivation for our first paper to inequality (10) from the general model in our 2011 paper. We repeat this inequality as a succinct summary of the journey:

$$R_g(\boldsymbol{d}^{RC}) \leq \tfrac{1}{r}(V(\boldsymbol{d}^{RC}) - V^{spot}) \ . \tag{10}$$

The left-hand side shows the role of formal governance on the equilibrium path: the formal governance structure can reduce the parties' maximum aggregate temptation to renege on the relational contract. The right-hand side shows the role of formal governance off the equilibrium path: the possibility of formal governance after reneging affects the net surplus from the relational contract. As our friend and mentor, Bengt Holmström, has told us and others, "You want your models to talk back to you." These two ways that formal governance can influence relational contract are what our models said to us.

[13] Here, again, others had considered related issues. For example, Blau and Scott (1962: 6) argued that "It is impossible to understand the nature of a formal organization without investigating the networks of informal relations and the unofficial norms … since the formally instituted and the informal emerging patterns are inextricably intertwined."

References

Abreu, Dilip. 1988. "On the Theory of Infinitely Repeated Games with Discounting." *Econometrica* 56: 383-96.

Aghion, Philippe and Jean Tirole. 1997. "Formal and Real Authority in Organizations." *Journal of Political Economy* 105:1-29.

Arrow, K. 1953. "Le role des valeurs boursieres pour la repartition la meilleure des risques." *Econometrie*, Centre National de la Recherche Scientifique, Paris, 41-48.

Baker, George. 1992. "Incentive Contracts and Performance Measurement." *Journal of Political Economy* 100: 598-614.

Baker, George. 2002. "Distortion and Risk in Optimal Incentive Contracts." *Journal of Human Resources* 37: 728-751.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109:1125-56.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1999. "Informal Authority in Organizations." *Journal of Law, Economics, and Organization* 15: 56-73.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *Quarterly Journal of Economics* 117: 39-83.

Baker, George, Robert Gibbons, and Kevin J. Murphy, "Bringing the Market Inside the Firm?" *American Economic Review Papers and Proceedings,* Vol 92(2) (May 2001): 212-18.

Baker, George, Robert Gibbons, and Kevin J. Murphy, "Relational Adaptation." (November 2011)

Blau, Peter and Richard Scott, *Formal Organizations: A Comparative Approach* (San Francisco, CA: Chandler Publishing, 1962).

Bull, Clive. 1987. "The Existence of Self-Enforcing Implicit Contracts." *Quarterly Journal of Economics* 102: 147-59.

Coase, Ronald. 1937. "The Nature of the Firm," *Economica* 4: 386-405.

Debreu, G. 1959. *Theory of Value*. New York, NY: John Wiley.

Denzau, A. and D. North. 1994. "Shared Mental Models: Ideologies and Institutions." *Kyklos* 47: 3-31.

Dhillon, A. and J. Rigolini. 2011. "Development and the interaction of enforcement institutions." *Journal of Public Economics* 95: 79-87.

Di Tella, R. and R. MacCulloch. 2002. "Informal Family Insurance and the Design of the Welfare State." *Economic Journal* 112: 481-503.

Fast, N. and N. Berg. 1975. "The Lincoln Electric Company." Harvard Business School Case #376-028.

Friedman, James W. 1971. "A Non-cooperative Equilibrium for Supergames." Review of Economic Studies, 38: 1-12.

Gibbons, Robert. 2010. "Transaction-Cost Economics: Past, Present, and Future?" *Scandinavian Journal of Economics* 112: 263-88.

Granovetter, M. 1985. "Economic action and social structure: The problem of embeddedness." *American Journal of Sociology* 91: 481-510.

Grossman, Sanford. 1981. "An Introduction to the Theory of Rational Expectations Under Asymmetric Information." *Review of Economic Studies* 48: 541-59.

Grossman, Sanford and Oliver Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Ownership." *Journal of Political Economy* 94: 691-719.

Halonen, Maija. 2002. "Reputation and the Allocation of Ownership," *Economic Journal*, 112(July): 539-558.

Hayek, F. A. 1945. "The Use of Knowledge in Society," *American Economic Review* 35(4): 519-530.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7: 24-52.

Kerr, Steven. 1975. "On the Folly of Rewarding A, While Hoping for B." *Academy of Management Journal* 18:769-83.

Klein, Benjamin and Kevin M. Murphy. 1988. "Vertical restraints as contract enforcement mechanisms." *Journal of Law and Economics* 31: 265-97.

Klein, Benjamin. 1996. "Why hold-ups occur: the self-enforcing range of contractual relationships." *Economic Inquiry* 34: 444-63.

Klein, Benjamin. 2000. "The role of incomplete contracts in self-enforcing relationships." *Revue D'Economie Industrielle* 92: 67-80.

Kranton, Rachel E. 1996. "Reciprocal Exchange: A Self-Sustaining System." *American Economic Review* 86: 830-51.

Kreps, David. 1990. "Corporate Culture and Economic Theory." In J. Alt and K. Shepsle, eds. *Perspectives on Positive Political Economy.* Cambridge University Press.

Levin, Jonathan. 2003. "Relational Incentive Contracts." *American Economic Review* 93: 835-57.

Macaulay, S. 1963. "Non-contractual relations in business: A preliminary study." *American Sociological Review* 28: 55-67.

MacLeod, Bentley, and J. Malcomson. 1989. "Implicit contracts, incentive compatibility, and involuntary unemployment." *Econometrica* 57: 447-480.

Macneil, I. 1978. "Contracts: Adjustments of Long-Term Economic Relations under Classical, Neoclassical, and Relational Contract Law." Northwestern University Law Review 72: 854-905.

Malcomson, J. 2013. "Relational Incentive Contracts." Chapter 25 (pp. 1014-65) in R. Gibbons and J. Roberts (eds.), *Handbook of Organizational Economics*. Princeton, NJ: Princeton University Press.

Prendergast, Canice and Lars Stole. 1999. "Restricting the means of exchange within organizations." *European Economic Review* 43: 1007-19.

Richardson, George. 1972. "The Organisation of Industry." *Economic Journal* 82: 883-96.

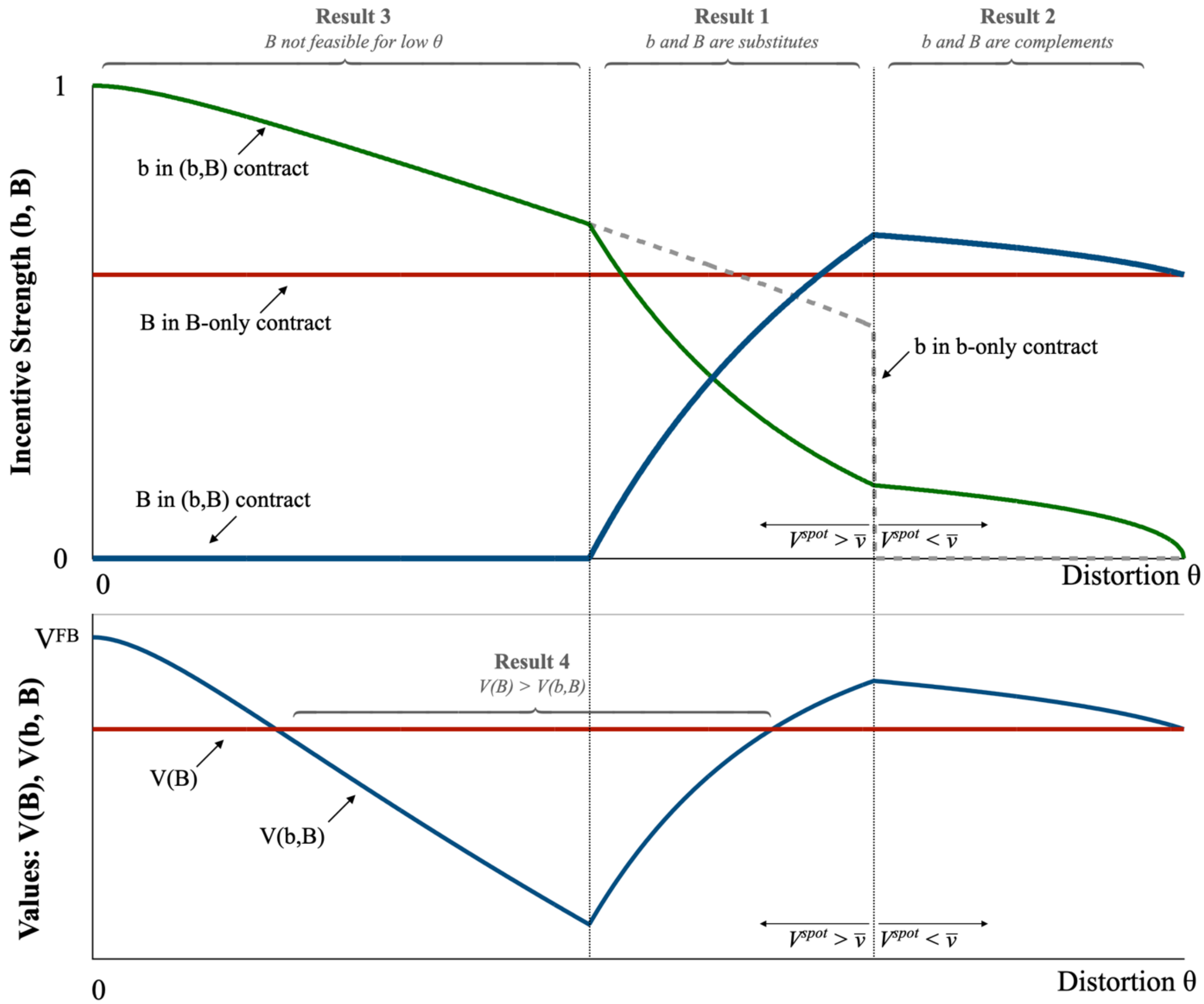
Robertson, Dennis H. 1923. *Control of Industry*.

Simon, Herbert. 1951. "A Formal Theory of the Employment Relationship." *Econometrica* 19: 293-305

Williamson, Oliver. 1979. "Transaction Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics* 22: 233-61.

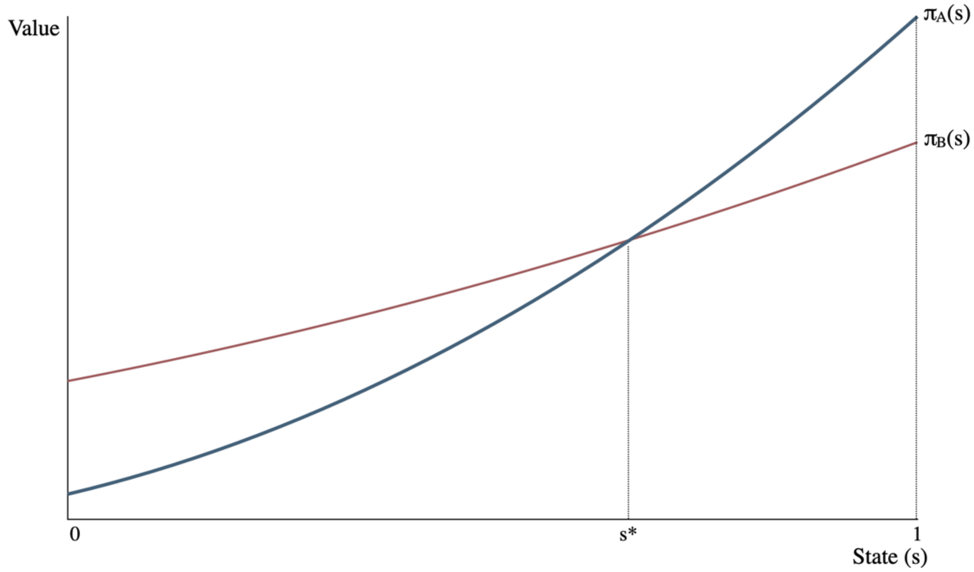
Williamson, Oliver. 1991. "Comparative Economic Organization: The Analysis of Discrete Structural Alternatives." *Administrative Science Quarterly* 36: 269-96.

Figure 1 Formal (b) and relational (B) incentive strengths in optimal contracts



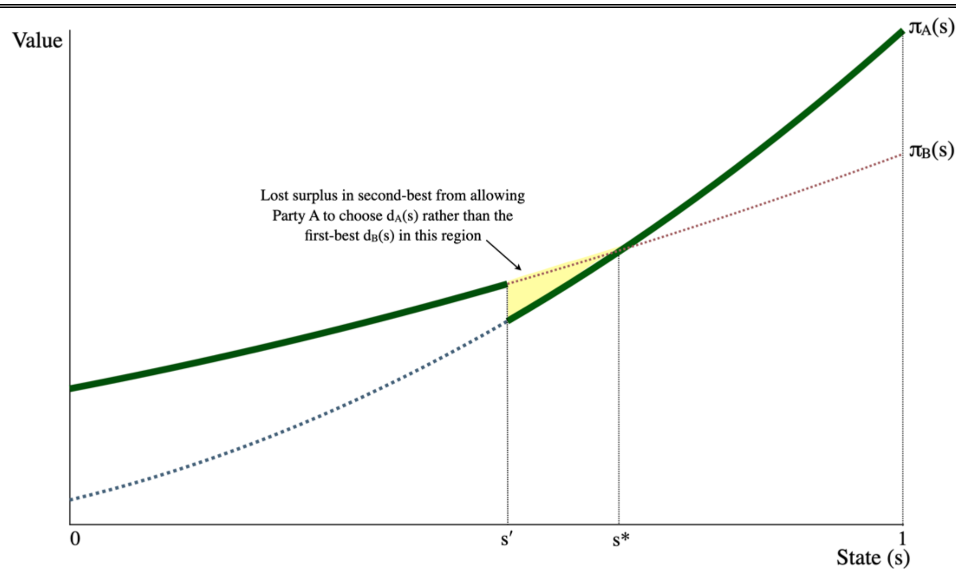
Note: The top panel shows the optimal piece rate ($b \in (0,1)$) on the objective measure p, and the optimal discretionary bonus ($B \in (0,1)$) on the non-verifiable measure y, for b-only, B-only, and contracts with both (b, B). The bottom panel shows the expected gross surplus created by the B-only contract ($V(B)$) and the (b, B) contract ($V(b, B)$). The horizontal axis is a monotonic transformation of the distortion in the objective performance measure, ranging from $\theta = 0$ (no distortion, so that b-only contracts are First Best) to $\theta = 10$ (full distortion, so the vector of efforts improving y are orthogonal to those improving p. The figure is based on the following assumptions: $f_1 = 1,\ f_2 = 0,\ h_2 \in (0,1),\ (h_1^2 + h_2^2) = 1$, and $c(a_1, a_2) = \frac{1}{2}(a_1^2 + a_2^2)$, where $r = 50\%$, $\overline{v} = .12$, and $\theta = 10^{h_2}$.

Figure 2 Illustration of private benefits for two parties with one decision right



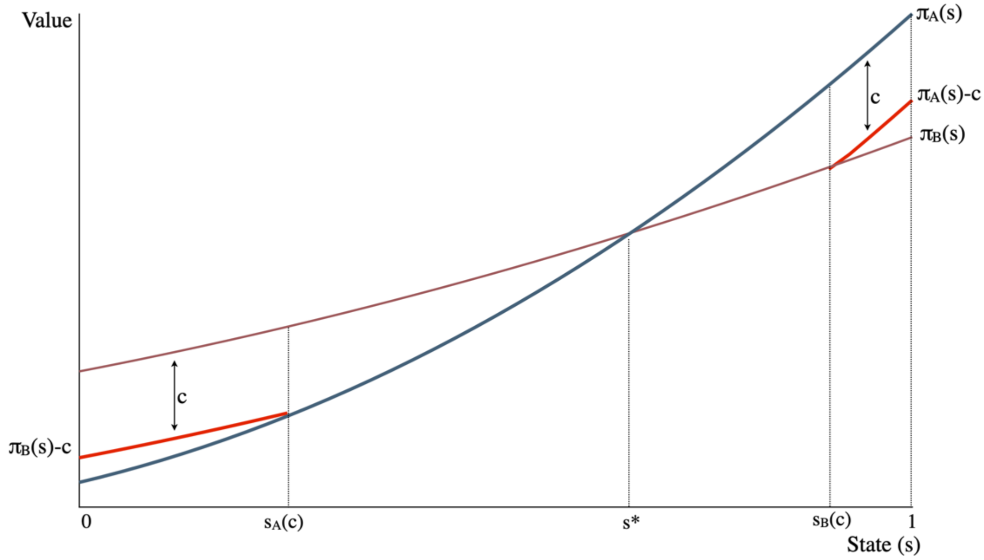
Note: The figure shows the private benefits for Parties A and B ($\pi_A(s)$ and $\pi_B(s)$, respectively) for a decision that is effectively binary (as defined in the text). The first-best decision rule is to implement the decision preferred by Party A when $s > s^*$ (that is, when ($\pi_A(s) > \pi_B(s)$)) and to implement the decision preferred by Party B when $s < s^*$.

Figure 3 Second-best decision rule for two parties with one decision right



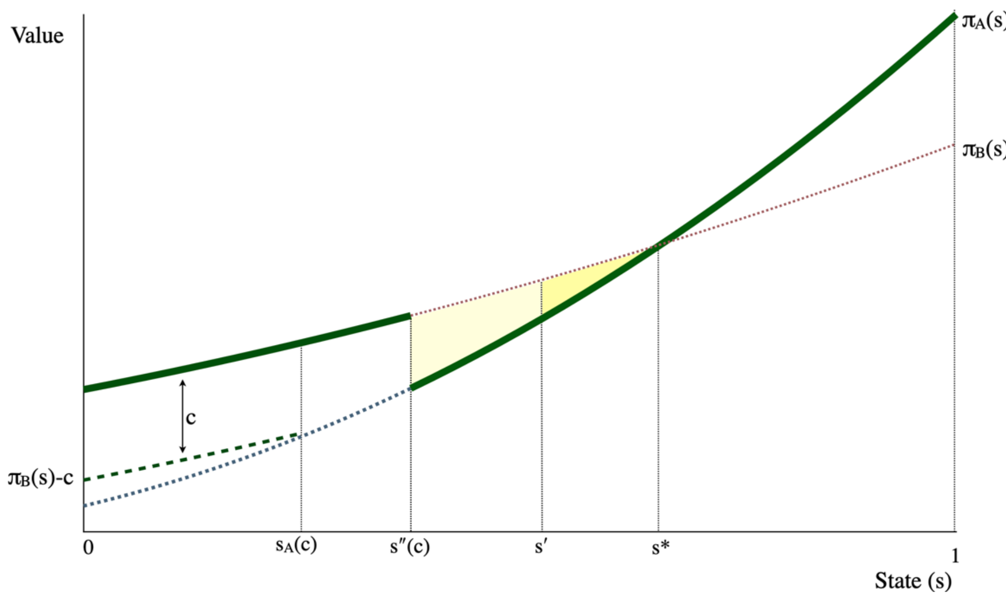
Note: The figure shows the private benefits for Parties A and B. For intermediate values of r, the second-best has Party A in control, taking the decision preferred by Party B when $s < s'$ and the decision preferred by Party A when $s > s'$, where $s' < s^*$. The critical value s' maximizes the parties' expected total payoff, subject to Party A's reneging constraint.

Figure 4 States where it is efficient to pay *c* to contract on *d* under spot adaptation



Note: The figure shows states where it is efficient to pay c to contract on d after the state is realized. Because $\pi_B(0)$ - $\pi_A(0) > c$, for states $s < s_A(c)$ if A were in control then it would be efficient to contract on $d = d_B$ rather than allow A to choose d_A (where $s_A(c)$ solves $\pi_A(s) = \pi_B(s) - c$). Because $\pi_A(1)$ - $\pi_B(1) > c$, for states $s > s_B(c)$ if B were in control then it would be efficient to contract on $d = d_A$ rather than allow B to choose d_B (where $s_B(c)$ solves $\pi_B(s) = \pi_A(s) - c$).

Figure 5 Second-best relational adaptation in the shadow of enforceable contracts



Note: The figure depicts a potential second-best relational contract where A chooses to implement $d = d_A$ for $s \geq s''(c)$ and $d = d_B$ for $s < s''(c)$. Since $s''(c) < s' < s^*$, the contract results in less surplus than first-best (where $d = d_B$ for $s < s^*$) or second best when ex post contracting is prohibited (where $d = d_B$ for $s < s'$). Also, since $s_A(c) < s''(c)$, the second-best relational contract illustrated here does not involve ex post contracting even though it is feasible at cost c.