Deals That Start When You Sign Them

Robert Gibbons

MIT Sloan School of Management, Cambridge, MA, USA
* Corresponding author. Email: rgibbons@mit.edu

Abstract
This essay explores six sentences from Oliver Williamson—five providing context and the sixth the central topic. Decades ago, Williamson asserted that: (a) “substantially the same factors” (1973: 316) create governance issues not only within organizations but also in interactions between organizations; and (b) relational contracting might be useful in addressing these issues in both domains (1979, Figure II). More recently—in an informal conversation in 2002—he suggested a perspective on relational contracting that appears valuable in both of these domains: relational contracts as “deals that start when you sign them.” The bulk of this essay explores past, present, and potential research on this perspective.

1. Introduction

Among the many pearls and contributions in Oliver Williamson’s writing and speaking, six sentences have etched themselves into my memory. For this essay, five provide background and the sixth gives the essay its title and focus.

Like many, I dove into Williamson’s work when I wanted to understand the make-or-buy decision that determines the boundary of the firm. I didn’t always find the writing easy going, but eventually seven words pierced my consciousness: “fiat is frequently … more efficient … than … haggling” (1971: 114). I find it hard to imagine a more efficient summary of a (not “the”) theory of integrated versus non-integrated control.

Of course, others also contributed to this domain, so I sought to distinguish one contribution from another. I became fond of a new sentence—this one more in Olly’s unique patois: “The most consequential difference between the TCE and GHM setups is that the former holds that maladaptation in the contract execution interval is the principal source of inefficiency, whereas GHM vaporize ex post maladaptation by their assumptions of common knowledge and costless ex post bargaining” (2000: 605).1

I am very grateful to: David Kreps for guidance that took me twenty years to begin to absorb; George Baker, Rebecca Henderson, and Kevin J. Murphy for decades of research and discussion about related issues; Charles Angelucci, Claude Ménard, and three anonymous referees for suggesting improvements to this essay; and MIT Sloan’s Program on Innovation in Markets and Organizations for intellectual and financial support.

1 For readers new to this literature, “TCE” refers to transaction-cost economics, which might be roughly defined as work written or inspired by Williamson, whereas “GHM” refers to the property-rights models by Grossman and Hart (1986) and Hart and Moore (1990); see Gibbons (2005) on similarities and differences between the two.
Even those who came to Williamson for his thinking on the boundary of the firm soon encountered a broader vision. For example, “any issue that either arises as or can be recast as a problem of contracting is usefully examined in transaction cost terms” (1985: xii). Put differently, I found Williamson because of the binary distinction between integration versus non-integration, but I learned from him to focus instead on another binary distinction: transactions that are governed by the parties themselves versus those that are cleanly priced in a market.²

Importantly, such governance occurs not only within organizations but also between them, such as in alliances, hand-in-glove supply relationships, and the like.³ Encompassing governance activities both within and between organizations, Williamson (2002: 438) wrote: “Private ordering efforts by the parties, to … embed transactions in more protective governance structures, have the purpose and effect of mitigating the contractual problems that would otherwise arise.” Both of these governance domains—within organizations and between them—appeared prominently in his 1979 paper, and as early as (1973: 316) he also compared them, concluding that “substantially the same factors that are ultimately responsible for market failures also explain failures of internal organization.”⁴

These sentences took me on a journey, teaching me a sequence of five points:

1) “fiat”—control (even dictatorship) could be a relatively good thing;
2) “vaporize ex post maladaptation”—there is more to control than threat points in efficient bargaining;
3) “arises as or can be recast as a problem of contracting”—there is more to governance than control;
4) “private ordering … in more protective governance structures”—the parties are governing themselves, at most in the shadow of a court; and
5) “substantially the same factors”—all of this applies not only within organizations but also between them.

These five points were a lot to learn from one person; I say more about this in Section 3. But Olly also made a sixth point—one that has extended this journey, but in a direction I’m not sure he intended; I say more about this in Section 4. To begin, however, I turn next to the sixth sentence.

2. “I’m interested in deals that start when you sign them.”

Much to my dismay, I had the privilege of only one sustained two-person conversation with Oliver Williamson, but over time it has had a big effect on me. It was late afternoon on a

---

² In a beautiful essay, Dixit (2009: 5) argues that “[a]s with any buzzword, everyone understands the concept [of governance] a little differently. This is unavoidable, so I will just give my definition for the purpose of this article, and leave it at that.” Grateful for this license, I follow Gibbons (2020) by seeing governance as the main thing that visible hands do, whereas adjusting prices is what the invisible hand does (and visible hands sometimes do, too).

³ For early work and examples between just two firms, see Richardson (1972), Dore (1983), and Stinchcombe (1985). For examples involving more than two parties and more recent work, see Eccles (1981), Powell (1990), Ménard (2004), and Chassagnon (2014).

⁴ See Gibbons (2010: 277) on how this similarity threatens the identification strategy of most empirical TCE studies of the boundary of the firm.
beautiful summer day in 2002, in Lucca (Italy), and Olly had nobody walking back to the hotel with him after the conference in honor of Jim March’s 75th birthday.

Lucca is surrounded by a huge wall—broad enough for trees, benches, and pedestrians on top, and tall enough for a nice breeze up there. If you look up “stroll” in the dictionary, you should see a picture of that afternoon on top of the wall. The atmosphere loosened at least my conversation.

About halfway back to the hotel, Olly told me he wanted to teach negotiations. Surprised, I asserted that social psychology seemed to be the dominant input to negotiations teaching. He replied “They study deals that are done when you sign them; I’m interested in deals that start when you sign them.”

I love this (third) binary distinction. There are indeed deals that are done when you sign them. Many of those require little or no negotiation, such as goods ordered from Amazon or services from TaskRabbit. And, of course, there are other deals that require massive negotiation but nonetheless are done when you sign them, such as some corporate acquisitions.

Then there are deals that could be said to start when you sign them but clearly are not what Williamson had in mind: complete contingent contracts, as introduced abstractly by Arrow (1953) and Debreu (1959) and as implemented more concretely by Moore and Repullo (1988) in extensive-form games. Starting with his earliest writings, Williamson clearly differentiated his approach from complete contingent contracts—see (1971: 115), (1979: 236), and as early as page 7 of his 1975 book. Accordingly, for purposes of this essay, I will count complete contingent contracts as deals that are done when you sign them, in the sense that when the deal is signed its possible futures are perfectly articulated, albeit contingent on uncertain events.

What, then, might Oliver Williamson have meant by “deals that start when you sign them”? My guess is that he meant something involving relational contracts (defined below)—both within and between organizations.

3. Relational Contracts as Deals that Start When You Sign Them?

Williamson’s most cited paper (1979) and most cited book (1985) both emphasized what he called “relational contracting.” Citing important work in the field of law & sociology (Macaulay, 1963) and especially in law itself (Macneil, 1978), the 1979 paper gave relational contracting a central role in the governance of recurrent, specific transactions both within and between firms (p. 253).

The 1985 book continued this emphasis via its subtitle, Firms, Markets, and Relational Contracts, although this list of terms risks the interpretation—contrary to Figure II of the 1979

---

5 Perhaps my memory is tricking me, but it feels like “No, no, young Gibbons” and an avuncular pat on the top of the head accompanied this sentence.
paper—that relational contracts arise only between firms, not within. Importantly, Macneil (1978) discussed relational contracting not only between firms but also in employment relations.

In this section I will use the term “relational contract” to mean a shared understanding of the parties’ roles in and rewards from collaborating together (an understanding so rooted in the details of the parties’ relationship that it cannot be enforced by a court). In economic models, such a relational contract is an equilibrium of a repeated game and hence self-enforced (whereas a formal contract has standing in a court). Such relational contracts thus might be called “informal contracts,” in contrast to formal contracts, but “relational” better connotes the ongoing aspect of a relationship. This concept and terminology have been used to interpret Williamson (1975, 1985); see Kreps (1996). More generally, this concept and terminology are now standard in organizational economics.6

Within organizations, I see such relational contracts as representing part of the informal aspects of organizations.7 Of course, the purely instrumental logic of repeated-game equilibria abstracts from non-economic considerations such as embeddedness (Granovetter, 1985), trust (Rousseau et al., 1998), and social capital (Burt, 2005).8 Without disputing the importance of these non-economic considerations, it is widely agreed that this instrumental logic—that the shadow of the future may influence behavior today—is sometimes important in ongoing relationships; for example, see Granovetter (1985: 490).

Between organizations, this purely informal conception of relational contracting is distinct (by definition) from agreements that have standing in courts. Of course, parties often use formal and relational contracts together. As Klein (2000: 68) observed, “although Macaulay and others are correct in noting that many business relationships are self-enforced, transactors are not indifferent regarding the [formal] contract terms they choose to govern their self-enforcing relationships.”9

Finally, such self-enforcing relational contracts fit well with Williamson’s notion of “private ordering,” including in settings beyond those that Williamson considered. For example, Ostrom (1990), Greif (1993), Clay (1997), and others consider settings where neither courts nor formal

---


7 As Granovetter (1985: 502) noted, “The distinction between the ‘formal’ and the ‘informal’ organization of the firm is one of the oldest in the literature, and it hardly needs repeating that observers who assume firms to be structured in fact by the official organization chart are sociological babes in the woods.”

8 Regarding non-economic considerations, Williamson (1975: 37-39) devoted two pages to the possibility of “atmosphere,” especially within organizations, linking it to “quasimoral involvements” that seem different from the pure instrumental logic of repeated games. Atmposphere did not appear in the 1985 book and made only a cameo appearance in the 1996 book. Baudry and Chassagnon (2010: 494) suggest that “atmosphere has progressively been overtaken by the idea of ‘relational contracting’” in Williamson’s work.

9 Similarly, within organizations, as Blau and Scott (1962: 6) noted, “It is impossible to understand the nature of a formal organization without investigating the networks of informal relations and the unofficial norms as well as the formal hierarchy of authority and the official body of rules, since the formally instituted and the informal emerging patterns are inextricably intertwined.” And, blending considerations internal and external to firms, see Chassagnon and Haned (2019) on complementarity between private and public normative orderings in the theory of the firm.
contracts exist, yet parties may govern their interactions via relational contracts. In all of these settings—within firms, between firms, and beyond those that Williamson considered—I therefore see such relational contracts as illustrating points (3), (4), and (5) of the journey.

In the next three sub-sections I provide informal summaries of models developed with George Baker and Kevin J. Murphy (1994, 2002, 2011) that study different interactions between formal and relational contracts. These papers converged towards Williamson—from incentives (1994) to property rights (2002) to adaptation (2011). More specifically: the incentives paper was not motivated by Williamson and does not cite him, but nonetheless pushed our thinking in his direction; the property-rights paper addressed important questions raised by Williamson but made assumptions that “vaporize ex post maladaptation;” and the adaptation paper considered how relational contracting might ameliorate “maladaptation in the contract execution interval [as] … the principal source of inefficiency.” Put differently, the property-rights and adaptation models illustrate points (1) and (2).

Incentives

George, Kevin, and I started down this path because of a case study on Lincoln Electric—a Fortune 500 firm making arc welders in Cleveland (Fast and Berg, 1975). Lincoln’s compensation scheme made heavy use of both a formal piece rate and a discretionary bonus. In the notation of agency theory, a worker’s total compensation (w) could be seen as including a salary (s), an objective weight (b) attached to an objective performance measure (p), and a discretionary bonus (B) meant to reflect aspects of the worker’s total contribution (y) not captured by the objective measure.

Knowing no work on such combinations of formal and informal contracting, we analyzed a model in which \( w = s + bp + B(y) \), where \( bp \) is enforced by a court but \( B(y) \) is at the discretion of the firm—that is, \( B(y) \) is a relational contract.\(^1\) We intended the resulting paper (1994) as a contribution to agency theory, but it led us to the make-or-buy decision, as follows.

Of course, there would be no need to rely on the discretion of the firm if the formal contract were perfect. We therefore considered the optimal use of both formal and relational incentive contracting as a function of the imperfection of the formal contract (e.g., p. 1144). Surprisingly, we found that when the best available formal contract is imperfect but not terribly so, all relational contracting may be impossible: even a small bonus is not credible because after reneging the firm could employ a new worker using only the formal contract.\(^2\)

---

\(^1\) For those new to models of relational contracts as repeated-game equilibria, consider this version of the firm’s decision about whether to pay the bonus \( B(y) \) at the end of a particular period. This decision arises after \( p \) and \( y \) have been realized, and the firm is obligated by formal contract to pay \( s \) and \( bp \) so paying the bonus produces profit \( y - s - bp - B(y) \) this period and some expected profit \( \pi_c \) in each future period, whereas not paying the bonus produces \( y - s - bp \) now and some expected profit \( \pi_d \) in each future period, where the subscripts C and D connote cooperation and defection. The expected profit \( \pi_c \) might be \( E\{y - s - bp - B(y)\} \), exceeding the expected profit \( \pi_d \) determined by using only formal contracting (with some new coefficient \( b' \)) after reneging on this period’s promised bonus. If the firm is sufficiently patient then the present value of \( \pi_c - \pi_d \) in future periods exceeds the reneging temptation \( B(y) \) today, so the firm pays the bonus.

After absorbing this surprising result, we thought we heard echoes of Williamson (1975). The imperfect formal contracting was easy to map into Williamson’s ideas: that was the “market failure” that might induce integration. And the relational contracting was not hard: in Chapter 2, Williamson had discussed “atmosphere” within firms (see footnote 8 above), and in Chapter 4 he summarized Simon’s (1951) model of the employment relationship—which, although not formally a repeated game, did inspire Simon to note (p. 302) that the parties could both be better off in a repeated employment relationship than in a static one.

It was the next step where we needed to squint: did it make any sense to interpret our model as saying that imperfect non-integration (i.e., an interaction governed solely by the formal contract) could make employment (involving relational contracts) impossible? That is, was the model suggesting that an imperfect market might crowd out what could have been a more effective firm? Eventually, we decided that (far) too much squinting was required in this interpretation. To consider these issues on a solid footing, we built our next model (2002) in terms of property rights instead of incentives.

**Property Rights**

One of our concerns with over-interpreting the 1994 model was that it would have been arbitrary to allow relational contracts within firms but only formal contracting between firms. Instead, in the 2002 model we allowed relational contracts both within and between firms; and to simplify the model, we ignored formal contracts entirely—within or between firms.

Absent formal contracts, to address formal governance we added to the model an asset that could be owned by either an upstream or a 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. As in other property-rights models, (i) we defined ownership of the modeled asset as conveying ownership of the intermediate good produced by the upstream party and used by the downstream, and (ii) we vaporized ex post maladaptation by assuming that bargaining under non-integration is efficient.

For ease of exposition, I will re-use the notation from the agency-theoretic model: we solved for the optimal relational contract, $B_I(y)$, under integration (i.e., the downstream party owns the modeled asset) and the optimal relational contract, $B_{NI}(y)$, under non-integration (i.e., the upstream party owns the asset). 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 producing strong results, as in our 1994 model. In contrast, under non-integration, the upstream party initially owns the intermediate good, so the relational contract $B_{NI}(y)$ is a sales contract, rewarding the upstream party for creating the good and transferring it to the downstream party at the anticipated price.

The main analytical result from the model (obvious in retrospect) was that a given relational contract—i.e., some promised payment function $B(y)$—typically has different reneging temptations under integration versus under non-integration. As one stark example, under
integration it is of course the downstream party who is tempted to renege on the relational
incentive contract $B_I(y)$, but under non-integration it may be the upstream party who is tempted
to renege on the relational sales contract $B_{NI}(y)$ if the outside option for the intermediate good is
unusually high.

The fact that a given relational contract has different reneging temptations under different
formal governance structures 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
optimal 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.

**Adaptation**

Once the property-rights model taught us that the integration decision could be made in the
service of the relationship, that idea seemed to have wider application than our 2002 model had
conveyed. In particular, while our property-rights model shared with Klein et al. (1978),
Williamson (1979, 1985), Grossman and Hart (1986), and Hart and Moore (1990) an emphasis
on specific investments as a source of governance problems, there was another theoretical
tradition from Simon (1951) through Williamson (1971, 1975, 1991) that instead emphasized
adaptation as a source of governance problems—even in the absence of specific investments.\(^\text{12}\)

Furthermore, and at least as importantly, there were two growing empirical literatures
studying optimal governance structures with few if any specific investments: one on contracting
for control, where parties use formal contracts to allocate decision rights across fixed firm
boundaries; and another on the classic make-or-buy problem, but again with little or no reference
to specific investments.\(^\text{13}\) Given these theoretical and empirical motivations, we wrote the 2011
paper to explore how formal governance might be chosen in the service of relational adaptation.

In an adaptation model, the issue is how well the parties choose decisions $d_t = (d_{1t}, ..., d_{nt})$
in response to $s_t$, the state of the world in period $t$. If complete contingent contracts were
feasible, the parties could solve this problem in advance; and if decisions were costlessly
contractible once the state was realized, the parties could solve this problem via spot contracting
in each period. In contrast to these idealized situations, a growing theoretical literature
recognizes both the difficulties of state-contingent contracting and the limitations of spot
contracting by imposing the stark assumption that decisions are not contractible: the party
controlling a decision can take it as they like.\(^\text{14}\)

Like the first empirical literature noted above, the simplest interpretation of formal
governance in our adaptation model was “contracting for control”—i.e., using contracts to
specify which parties control which decisions, holding firms’ boundaries fixed in the sense that

\(^{12}\) See Gibbons (2010: 270-2) for more on the latter theoretical tradition.

\(^{13}\) See our 2011 paper for citations.

\(^{14}\) Again, see the 2011 paper for citations.
the payoff to firm $i$ in period $t$ is its profit $\pi_i(d_t, s_t)$, regardless of what decisions it controls. In the spirit of Ostrom’s (2010) “polycentric governance,” we allowed for $K \geq 2$ firms. Absent relational contracting, “maladaptation” arises if there is no allocation of decision rights across firms that produces the first-best decisions $d^{FB}(s)$ as a Nash equilibrium of the one-shot game.

In this setting a relational contract specifies a decision rule $d(s)$ and bonuses $B(s, d)$ paid to and by appropriate parties based on the state that is realized and the decisions that are taken. This time we expected the model’s main result: a given relational contract has different reneging temptations under different formal governance structures (i.e., different allocations of decision rights across firms). We therefore could once again ask what relational contract would be most effective and then choose the governance structure to minimize the reneging temptation. That is, in this model, formal contract terms are chosen in service of the relationship.\footnote{See Gilson, Sabel, and Scott (2009, 2010), Bernstein (2015), Hadfield and Bozovic (2016), Frydlinger, Hart, and Vitatek (2019), and Grandori and Furlotti (2019) for work in a similar spirit (i.e., formal contracts that facilitate relationships), but based on contract law and contracting practice rather than on a formal model.}

**Assessment**

As advertised, this sequence of three models converged towards Williamson. Now, do these models capture “deals that start when you sign them”? In one sense, they do: although in each model the formal governance structure—i.e., the coefficient $b$ in the agency model, the integration decision in the property-rights model, and the allocation of decision rights in the adaptation model—is chosen afresh in each period, in a stationary equilibrium that choice is the same in each period, giving the impression of formal terms having been decided at the beginning and informal aspects then playing out over time. That is, in this sense of Williamson’s phrase, what is signed is the formal aspect, and what starts at signing is the informal aspect.

But these models (and all the relational-contract models I know, whether or not they also include formal governance) share the spirit of complete contingent contracts: in a Nash equilibrium, the parties have a shared understanding of each other’s strategies; in a dynamic game with uncertain future states, such an equilibrium involves shared understanding of complete contingent plans, if not complete contingent contracts. Thus, the relational contracts in the models above can also be seen as deals that are done when you sign them—at inception, their possible futures are perfectly articulated, albeit contingent on uncertain events.

In contrast, a second sense of “deals that start when you sign them” would be deals where possible futures are not perfectly articulated when the deal is signed, so the parties must interpret and refine the informal aspect as events unfold. In 2002 this second sense had not yet dawned on me, so in Lucca I interpreted Olly’s remark in the first of these two senses—concerning relationships where informal aspects play out over time, supported and constrained by formal governance structures, and possible futures are perfectly articulated from the beginning.

Soon after Lucca, however, discussions and research with Rebecca Henderson began to change my view from the first to the second sense of “deals that start when you sign them.” More specifically, as I describe in the next section, I still see existing models of relational contracts as illustrating only the first sense, but I now see actual relational contracts as often
illustrating also the second sense: real parties must interpret and refine the informal aspects as events unfold.

And I find myself considering in which of these two senses Williamson might have meant “deals that start when you sign them.” Since he abhorred complete contingent contracts, he might have felt similarly about complete relational contracts, thus rejecting the first sense of this phrase. But complete contingent contracts require a contract that could be adjudicated by a court, whereas complete relational contracts refer to the parties’ shared informal understanding. In the spirit of Polanyi’s (1966: 4) observation that tacit knowledge may mean that “we know more than we can tell,” parties may be able to share a more nuanced informal understanding than they can capture in a formal contract. Thus, it is possible that Williamson believed in complete relational contracts while abhorring complete contingent contracts.

4. The Clarity Problem

Rebecca came to relational contracts on a different path than mine: from a strategic interest in competitive advantage, technology adoption, and organizational change, rather than a theoretical interest in the interaction of formal and relational contracts and a substantive interest in the informal aspects of organizations. Blending our perspectives, we eventually converged on the following motivation for what we called the clarity problem.16

- There are persistent performance differences among seemingly similar enterprises (where “enterprise” might mean work group, plant, division, firm, alliance, community, or more)—see Syverson (2011) and GH (Section 2).
- Within firms, proxies for management practices are correlated with these performance differences—see Bloom and Van Reenen (2007), GH (Section 3), and Bloom et al. (2019).
- Many of these management practices (especially competitively significant ones) rely on relational contracts—see Foss (2003), GH (Section 4), and Turco (2016).

These observations led us to ask: what might slow the diffusion of competitively significant relational contracts?

Our proposed answer relates to the second sense of “deals that start when you sign them.” More specifically, players in models of relational contracts are in equilibrium from the beginning, but parties in life may not find it easy to reach (efficient) equilibria. Put differently, game-theoretic models of relational contracts have made huge progress on the credibility problem (“Should I believe the promise you are making me?”) but such equilibrium models side-step the clarity problem (“Do we have a shared understanding of the promise you are making me?”), implicitly assuming this clarity problem has already been solved.

16 The citations below are to the lengthy discussion in Gibbons and Henderson (2013)—hereafter, GH. For an efficient introduction to these ideas, especially for non-economists, see Gibbons and Henderson (2012).
Consistent with the clarity problem being a real issue, surely two firms about to launch an alliance would not only sign a formal contract but also discuss how they hoped their relationship would unfold. We imagine such discussions occurring on Sunday night, before the alliance begins its work on Monday, but these discussions do not arise in models that study equilibria—there is nothing to discuss.\(^\text{17}\)

As another indication of the clarity problem, after long shared experience, parties may have deep shared understanding of their relational contract, but (a) it may be difficult to explain it to newcomers and (b) when times change there may be disagreement even among experienced insiders about what responses would be consistent with the original understanding.\(^\text{18}\)

To sketch the clarity problem formally, consider a simplified version of the 2011 adaptation model in Section 3 (noting that parallel illustrations are easy to provide and important to consider in the other two models as well). Suppose there is only one decision right, and only two decisions are possible—say, \(d = L\) or \(d = R\). Although the decision space is simple, suppose the state space is complex: each period, the state \((s)\) is drawn from a rich space \((S)\). The clarity problem then is to build a shared understanding of which decision should be taken in which state: \(L\) in some subset of states \(S_L\) and \(R\) in the complementary subset \(S_R\).

To illustrate the potential difficulty of reaching such a shared understanding, consider psychology experiments such as Chiu (1972), where children are shown three objects and asked which two belong together. For example, if the objects are a cow, a tuft of grass, and a chicken, one might reason that the cow eats the grass or that the cow and the chicken are both animals. If the parties establish that the appropriate decision is \(d = L\) when the state of the world is a cow, does that mean it is the grass or the chicken that is also in \(S_L\)? Could members of one organization think grass while members of another think chicken?\(^\text{19}\)

In this sketch, the clarity problem concerns categorization: before a decision is taken, the parties ask themselves “Which kind of state is this?” The 2011 adaptation model implicitly assumed that the parties had a shared way to interpret states of the world; more generally, equilibrium models of relational contracts assume that the parties have a shared understanding of their strategies, specifying what to do in each possible contingency.

In short, at the start of a relationship the clarity problem is closer to “the parties can’t be clear” (about the relational contract of interest) than to “the parties should be clear” (about a relational contract that would need to be simple—perhaps simple enough to be a formal contract). And for purposes of this essay, the point is then that the need to build clarity over time may be an important part of deals that start when you sign them.

\(^{17}\) See Frydlinger et al. (2019) for extensive advice to alliance partners about such discussions—so extensive as to require a month of Sundays. It would be interesting to apply their advice to new relationships within organizations.

\(^{18}\) See GH (Section 6.1) on imperfectly shared understandings within organizations. For complementary work between organizations, see Doz (1996), Mayer and Argyres (2004), and Keller et al. (2021) on “disruptions” in alliances. And sacrificing external validity for laboratory control, see Gibbons et al. (2022) for an experiment.

\(^{19}\) Nisbett (2003) synthesizes decades of related research in discussing how Westerners and Asians might answer such questions differently; the clarity problem considers that neighboring firms might answer them differently.
To conclude this sub-section, I consider what Williamson wrote about persistent performance differences or the difficulty of building clarity over time. For example, Rebecca and I see the difficulty of building clarity over time as shedding new light on why some firms may achieve “consummate” performance but others only “perfunctory.” Williamson (1975: 69, 1996: 270-1) also used these terms, but I interpret him to mean perfunctory performance as a mere baseline (e.g., performance that might not survive competitive pressure) and consummate performance as the result of optimal governance via private ordering. Put differently (without any guesswork about what Williamson did or did not mean), I know of nothing in Williamson’s writings that connects transaction-cost economics to persistent performance differences.

Quite separately from Williamson, the early evidence and informal theory on performance heterogeneity started with Penrose (1959) and continued through decades of work in strategic management on organizational competence. Eventually, in the Strategic Management Journal, Williamson (1999) compared his “governance” perspective to strategy’s “competence” view, but more as a critique of the latter than as a reconciliation of the two.

Instead, it has been some of Williamson’s (direct and indirect) students who have pursued such a reconciliation for at least two decades. For example, building on important work in the 2000s by themselves and others, Argyres et al. (2012) wrote an introduction to a special issue of Organization Science dedicated to the “Organizational Economics of Capability and Heterogeneity” (“capability” having largely replaced “competence” as the term of art in the strategy literature). And more recently Argyres (2021) described both subsequent progress and significant opportunities, the latter including a call to “apply formal theories from organizational economics that have emerged since Williamson (1985, 1991) to understand capabilities and their development” (p. 524).

Taking all this into account—the conversation in Lucca, Williamson’s (1999) piece and other writings, and the decades of work by some of his students and their like—my assessment is that Williamson said an enormous amount about governance, including points (1)-(5) in Section 1, but essentially nothing about the sources of organizational capability and performance heterogeneity. Thus, if he did mean “deals that start when you sign them” in the second sense above—namely, that parties must interpret and refine the informal aspects of their relationship as events unfold, and hence that parties may not find it easy to reach (efficient) equilibria—he did not capture in his writing either this interpretation of “deals that start when you sign them” or the possible role of this second sense of this phrase in explaining organizational capability and performance heterogeneity. In this sense, I see his sixth sentence as a natural extension of the first five, but in a direction he may not have intended.

Organizational Culture As Shared Understanding

Another topic that received relatively little attention in Williamson’s writing is organizational culture. For example, the 1975 book discussed the related concept of “atmosphere” for two pages, the 1985 book mentioned neither atmosphere nor culture, and the 1996 book considered “corporate culture” in three places.

In my view—and, seemingly, in that of the others quoted below—part of organizational culture is closely related to the clarity issue just discussed. Since culture is related to the broader
interests of this *Journal* (i.e., beyond this special issue), I close this section by articulating the connection I see between culture and clarity, and I return to this connection in the Conclusion.


What topic were these disparate scholars discussing? Culture (and, for Pettigrew and Schein, specifically organizational culture)—the point being that such categories, frames, and mental models are shared by those within a given culture. DiMaggio (1994: 28) called this the constitutive aspect of culture, shaping how parties see and process the world around them. In this sense, relational contracts and organizational culture both involve shared understandings.

If the clarity problem is the need to build shared understandings that underlie the relational contracts of interest, and if part of organizational culture is shared categories, frames, and mental models, might building relational contracts require cultural work—i.e., building shared categories, frames, and mental models?

Backing up, Kreps (1990) long ago proposed that repeated-game equilibria capture part of corporate culture—a principle specifying “how things are done, and how they are meant to be done in the organization” (p. 93). Given the multiplicity of equilibria in a repeated game, Kreps’s original analysis is consistent with the persistent performance differences noted above, but there is a discomfiting aspect of the model: low-performing parties know there exists a better equilibrium, but the model gives the parties no way to try to reach it (and, as a result, the model does not suggest why reaching it might be hard).

Recently, Gibbons *et al.* (2021a) revisited Kreps’ approach, in a setting where parties (unknowingly) categorize the complex world around them. The parties play the best repeated-game equilibrium they can see, given their categorization; consistent with persistent performance differences, small differences in categorization can produce big differences in equilibria. Reaching (or even just perceiving) a better equilibrium would require changing the parties’ categorization, so the paper explores how exemplars, stories, and the like may induce categorization change, finding that first categorization change and then equilibrium change can be slow, risky, and path-dependent.

If changing shared categories, frames, and mental models is slow, risky, and path-dependent, then the project of building relational contracts to achieve consummate performance will be consistent with Barney’s (1986) observation that, if organizational culture is to provide competitive advantage, it must be difficult to imitate. Put differently, *cultural work on the clarity problem may be crucial in deals that start when you sign them.*

---

20 In a similar spirit, Adelstein (2010: 336) argued that not just a firm’s culture but even the firm itself is a multilateral “relational contract … [requiring] a ‘meeting of the minds’ of the individual contractors.”

21 See Gibbons *et al.* (2021b) on how this model relates to other conceptions of organizational culture.
5. Conclusion

This essay discussed three binary distinctions considered by Oliver Williamson. The first, between integrated versus non-integrated control, is illustrated by the 1975 book title *Markets and Hierarchies*. The second, between transactions that are governed by the parties themselves versus those that are priced in a market, is illustrated by the 1996 book title *Mechanisms of Governance*.

One of the important strengths of Williamson’s work was his unified approach to two domains of such governance—within and between organizations. As he noted long ago, the problems are similar: “substantially the same factors that are ultimately responsible for market failures also explain failures of internal organization” (1973: 316). And as he also argued, the instruments for addressing these problems are similar: relational contracting plays a central role in the governance of recurrent, specific interactions both within and between firms (1979: 253).

As far as I know, the third distinction, between “deals that are done when you sign them” versus “deals that start when you sign them,” was never published. Certainly, it received much less attention, from Williamson and from others. The core of this essay therefore discussed the novel half of this third distinction: past, present, and potential research on relational contracts as deals that start when you sign them.

Models where optimal governance involves a stationary formal contract and then a repeated-game equilibrium that plays out over time explore one sense of “deals that start when you sign them.” But such models ignore a conspicuous feature of the process of building real relationships: Sunday night. Put differently, the theoretical literature has developed great expertise on the credibility problem but essentially ignored the clarity problem.

In my view, we need more work on building (and refining, and updating, and repairing) relational contracts. Such work could be important for deepening our understanding of persistent performance differences, the informal aspects of organizations, and perhaps organizational culture. As far as I know, Oliver Williamson wrote relatively little about these topics, but I hope he would think that his work had facilitated productive new explorations.

To conclude this essay, I shift the focus from Williamson to institutional economics, asking whether cultural work on the clarity problem might be relevant to some of the broader concerns of this *Journal*, in three ways.

First, returning to settings beyond those that Williamson considered—such as those studied by Ostrom (1990), Greif (1993), Clay (1997), where neither courts nor formal contracts exist—the historical accounts inspiring repeated-game models of institutions such as Milgrom, North, and Weingast (1990) and Greif, Milgrom, and Weingast (1994) could be reconsidered in terms of the clarity problem, complementing the exclusive focus of these models on the credibility problem. That is, both at the institution’s inception and when it needed to be reinterpreted or reconfigured in response to long-run shocks, there is the potential for clarity issues quite separate from the credibility issues that determine the institution’s steady-state persistence.
Second, turning from institutions themselves to things that are institutionalized, the early sociological literature on institutionalized organizations—starting with Meyer and Rowan (1977) and DiMaggio and Powell (1983)—anticipates the clarity problem’s emphasis on the importance of shared (taken-for-granted) understandings, but the emphasis in this early literature is on how institutionalized organizations are so similar, the opposite of the focus on persistent performance heterogeneity above. More recently, however, an important literature on “institutional logics” has emerged as a rejoinder to the early literature’s emphasis that institutionalization might erase agency; see Thornton et al. (2012) for an introduction to existing work and open questions.

This newer literature on institutional logics may have interesting connections and interactions with the clarity perspective articulated here, but it is beyond the scope this essay to explore this in detail. As a first, tentative step, consider the following. Discussing culture rather than institutions, Gibbons et al. (2021b) distinguish between “big-C” culture (such as societal culture) that seeps into an organization via the backgrounds that members acquired before they joined the organization, versus “little-c” culture that is “invented, discovered, or developed by a given group” (Schein, 1985: 9) given whatever backgrounds the group’s members have. Just as it would be interesting to analyze how a given big-C culture influences the benefits and costs of building particular little-c cultures / relational contracts, it would likewise be interesting to ask how the big-I institutionalization pressures studied in the early sociology literature might affect attempts to build little-i institutions (even as little-i as relational contracts within and between organizations).

Third, in addition to the large literature on performance heterogeneity / organizational capability noted above, there is a large literature on managerial cognition that has developed almost independently but offers important prospects for interactions with the capability literature (Eggers and Kaplan, 2013). While much of the research on managerial cognition considers individual managers such as CEOs, there is also work on organizational cognition, including the possibility that an organization may know more than the aggregate of what its members know (Walsh, 1995).

The clarity problem emphasized in this essay illustrates another interaction between cognition and performance, but it focuses on an aspect of cognition different from both the cognition of individual managers and organizational cognition: in the clarity problem, the question is whether the parties’ understanding is intersubjective (i.e., existing between conscious minds) — as illustrated, for example, by the mutual knowledge players have of each other’s strategies in Nash equilibrium. That is, regarding the promises inherent in a relational contract, the clarity problem causes the parties to ask “Do we have a shared understanding of the promise you are making me?” and equilibrium analyses of relational contracts assume the answer is “Yes.”

Consistent with the correct answer being “No” in some circumstances, Kogut and Zander (1992: 383) argued long ago that “what firms do better than markets is the sharing and transfer of knowledge … within an organization.” Such knowledge is “held by individuals … [but] also expressed in [the] regularities by which members cooperate.” Kogut and Zander called such regularities “organizing principles” whereas here we have called them relational contracts (which involve not only the clarity problem of knowledge but also the credibility problem of incentives).
Terminology aside, we have now focused on the development of shared knowledge that is specific to the parties in question—a “specific asset,” if you will—which of course brings us full circle, back to Williamson’s central concerns. To repeat, I learned an enormous amount from the writings of Oliver Williamson, including from the journey through the five sentences highlighted in Section 1. This essay focused on a sixth, unwritten sentence, which I see as a natural extension of the first five, if perhaps in a direction that Williamson might not have entirely intended: the slow, hard work of building shared understanding implies an informal, accretive aspect of how specific knowledge assets are built, potentially resulting in organizational capability and competitive advantage—not only within organizations but also in organized interactions between them.

References


R. Gibbons
Deals That Start When You Sign Them
August, 2021


