Review of Economic Studies (2017) **84**, 816–852 doi:10.1093/restud/rdw024

Advance access publication 9 July 2016

A Theory of Contracts with Limited Enforcement

DAVID MARTIMORT

Paris School of Economics-EHESS

AGGEY SEMENOV

University of Ottawa

and

LARS STOLE

University of Chicago Booth School of Business

First version received October 2013*; final version accepted April* 2016 (*Eds.*)

We develop a theory of contracts with limited enforcement in the context of a dynamic relationship. The seller is privately informed on his persistent cost, while the buyer remains uninformed. Public enforcement relies on remedies for breaches. Private enforcement comes from terminating the relationship. We first characterize *enforcement constraints* under asymmetric information. Those constraints ensure that parties never breach contracts. In particular, a high-cost seller may be tempted to trade high volumes at high prices at the beginning of the relationship before breaching the contract later on. Such *"take-the-money-and-run"* strategy becomes less attractive as time passes. It can thus be prevented by backloading payments and increasing volumes over a *transitory phase*. In a *mature phase*, enforcement constraints are slack and the optimal contract, although keeping memory of the shadow cost of enforcement constraints binding earlier on, looks stationary. Second-best distortions depend on a *modified virtual cost* that encapsulates this shadow cost of enforcement.

Key words: Asymmetric information, Enforcement, Breach of contracts, Dynamic contracts

JEL Codes: D82

1. INTRODUCTION

1.1.1. Motivation. The theory of contracts is a major building block of our current understanding of how markets and organizations perform. Its most impressive contribution is to offer a full characterization of trading possibilities in environments where privately informed parties might have conflicting interests. Equipped with a description of the set of incentive feasible allocations, modellers look for institutions, organizations, or contractual forms that optimally balance efficiency and rent extraction. If any frictions impede contractual performances and prevent efficient trades, those frictions are supposed to come from asymmetric information.

Although quite successful, this methodology remains somewhat at odds with the predominant view of contracts that is instead offered by legal scholars. Authors in this field actually devote much effort in studying how contracts are enforced (Macaulay, 1963). Two key concerns are whether legal disputes and remedies arise to fill contractual loopholes and whether private enforcement, often referred to as *relational contracting*, could be an efficient substitute for more formal agreements (Baker and Choi, 2014).[1] By overlooking enforcement constraints in comparison with more traditional incentive compatibility constraints, the *theory of contracts* as it currently stands might bias recommendations on organizational choices in a systematic way and neglect important features of contracting. Without further inquiry, it is *a priori* unclear whether the lessons of this theory remain with limited enforcement. Indeed, a richer set of comparative statics may emerge from a theory that addresses how standard agency distortions depend on the quality of enforcement.

To bridge the aforementioned gap, this article develops a theory of contracts under limited enforcement. Parties may renege on their contractual duties at any point in time in the course of their relationship. To illustrate, observe that although the quantity that parties want to trade in any period may be verified and so subject to formal agreements, quality may hardly be so. The seller may pocket advance payments and shirk by providing a lower quality widget later on. The buyer may also delay or not fulfil her payment obligations following delivery. Such opportunistic behaviours can be jointly prevented by means of remedies enforced by courts (the public side of enforcement) but also by the threat of terminating the relationship (its private side).[2] Facilitating enforcement may call for reduced volumes and decreased prices so as to lower the opportunity cost of breaches.

A first goal of our analysis is to assess how the distortions induced by limits on enforcement modify the screening responses which are now well known from asymmetric information models. We delineate circumstances in which limits on enforcement impact on the rent/efficiency trade-off highlighted by this literature and investigate how those limits shape trading patterns. A second issue at the core of our analysis is to understand how to enforce the actual play of a contract when courts of law will only impose limited damages following a breach. Although parties can include in their contract an explicit statement about the damages that should be paid by a breaching party (*"liquidated"* or *"stipulated"* damages), courts routinely strike down contractual provisions stipulating damages for breach of contract when those damages appear excessive in terms of actual or anticipated damages or otherwise appear as a penalty rather than as a compensation. It is hard to articulate an internally consistent theory for this *"penalty doctrine."* (See Farnsworth, 1982; Posner, 2011, Chapter 8.) The practical upshot is that in any given trading environment, there is an upper bound to the amount of damages that a court will enforce following a breach, even if the damages were agreed to by both parties at an early time.

1. This hiatus between the views on contract theory held by economists and law scholars is well summarized by Masten (1999) who wrote:

> …, the literatures on contract design and contract enforcement have largely developed independently of one another. Economic theories of contracting, for the most part, give little explicit attention to enforcement issues, the presumption being that courts will see to it (subject only to verifiability constraints) that whatever terms contracting parties arrive at are fulfilled. Indeed, enforcing contracts as written is the court's only function in mainstream contract theory […] This judicial deference to contracts in economic theory contrasts with the far more intrusive role of courts in economic analyses of contract law, in which courts are called on to adjudicate disputes, fill gaps, and devise and implement default rules.

See also Kornhauser and MacLeod (2012) for a recent account.

2. In his text on contract law, Atiyah (1995, p. 6) stresses the joint use of public remedies and private devices to enforce contract when he writes: *"there are many sanctions against promise-breakers and law is not needed. The simplest sanction is not to deal with that person again."* Nevertheless, *"most systems of law have established rules which will impose sanctions on those who break their contracts."* Johnson *et al*. (2002) have also pointed out the joint use of relational contracting and legal remedies in transition economies where various enforcement costs make it difficult to rely exclusively on the judicial system.

Accepting this limitation on the *public* enforcement of contracts, we ask how parties will optimally design their contract to provide complementary *private* incentives for fulfilling their contractual obligations.

1.1.2. Set-up and main results. Consider a highly stylized model of a trading relationship between an uninformed buyer (the principal, or *she* in the sequel) and a seller (the agent, *he*) who repeatedly trade for delivery of a good or service. The seller has private information on his cost function. This type is persistent over the whole relationship. The buyer, whose preferences are common knowledge, has all bargaining power in designing a long-term contract which specifies prices and quantities over the course of the relationship.

Enforcement constraints. This relationship is subject to bilateral opportunism. The seller may fail to trade a widget with the requested quality; the buyer may fail to meet payment obligations. Contracts should now provide safeguards against opportunism so as to ensure that, any point of time, parties comply with their obligations. Even when taken in tandem, the private and public sides of enforcement may not ensure enforceability. This is so when the perspective of future trading does not suffice to motivate parties to abide to their current obligations or when court-enforceable remedies prove inadequate. A first step of our analysis thus consists in deriving enforcement constraints ensuring that both the seller and the buyer abide to the contract. Because parties have quasi-linear payoffs, individual enforcement constraints can be pooled into forward-looking *enforcement constraints*. The foregone benefits of future trades plus the total remedies paid following a breach by either party must be large enough to ensure joint compliance.

"Take-the-money-and-run" strategy. Enforcement constraints require that the value of continuing trades exceeds the benefits that parties may withdraw from not fulfilling their contractual obligations. Under asymmetric information, the value of trade is reduced because part of it is left as information rent to the privately-informed seller. Enforcement constraints are thus hardened.

New strategic possibilities arise under asymmetric information. To induce a low-cost seller to reveal his private information at the start of the relationship, the buyer raises the price paid to this type. A high-cost seller may thus find it attractive to adopt the behaviour of a low-cost one at the beginning of the relationship, pocketing large payments for a while before breaching. This *"take-the-money-and-run"* strategy shapes intertemporal incentives. Making such strategies less attractive requires backloading the low-cost seller's payments and reducing outputs below the optimal levels achieved with costless enforcement.

As a consequence, the optimal dynamic contract goes through two different phases. In the first *transitory phase*, trading volumes and prices increase over time as the high-cost seller's incentives to breach diminish. Indeed, the *"take-the-money-and-run"* strategy becomes less attractive as time passes. In the limit, a high-cost seller always mimicking a low-cost type would just violate incentive compatibility. After enough periods, the enforcement problem then looks very much as if it was taking place under complete information. In this more *mature phase*, outputs and prices become stationary. The low-cost seller produces the first-best level of output while that of the high-cost seller remains distorted below the static optimal second-best level. This new distortion is captured by a *modified virtual cost* that accounts for the overall shadow cost of binding enforcement constraints over the transitory phase. Even in the long run, the optimal contract thus keeps memory of the cost of enforcement. Output distortions are exacerbated in comparison with the static optimal second-best contract. Those distortions are magnified when the quality of enforcement (be it public or private) deteriorates.

1.1.3. Application: construction contracts. Construction contracts offer a particularly relevant application for our framework since they exhibit features that closely replicate elements in our model.[3] Construction projects, especially those involving large-scale infrastructure, are inherently multistage production processes. For each stage, the contract not only specifies an expected time for delivery but also requires that various indicators are met to check the adequacy between the buyer's needs and what the contractor delivers. Late completions, failures to meet specifications and missed payments in due time are pervasive phenomenon.

To respond to those contractual hazards, practitioners have developed devices whose goal is to deal with liquidated damages and extensions of time in a way that can be compatible with the needs of both contractors and clients (Eggleston, 2009). The theoretical concerns at the core of our analysis are also concerns on the practitioners' agenda. Most standard forms of contracts require that parties establish in advance remedies for possible breaches and those remedies are genuine estimates of possible losses incurred by the party on the other side of the transaction. Echoing the analysis we develop below, enforcement of construction contracts relies on a mixture of public and private devices. In particular, parties often rely on relational contracting and trust to continue relationships even following unforeseen events that could have triggered legal disputes (Johnson and Sohi, 2015).

Lastly, a major issue faced by practitioners is that *"take-the-money-and-run"* strategies might put contracts at risk. The *Channel Tunnel* provides a famous example. One of the main criticisms of the original arrangement came when observers recognized the hazards associated with excessive front-loaded payments (Vinter and Price, 2006, p. 100). To illustrate this point more broadly, a client may agree to make an advance payment (so called *"down payment"*) to a supplier so as to help the latter to pay start-up costs, hire subcontractors or access key resources and equipment. With the bulk of payments being paid upfront, contractors may nevertheless play a *"duck-and-run"* game and delay completion.[4] In response, the client may want to secure his payment against the contractor's default through so-called *"advance payment bonds"*. This practice certainly echoes the benefits of backloading payments as predicted by our model.

1.1.4. Literature review. Our analysis demonstrates that, even though contract enforcement might rely on public remedies, the threat of terminating relationships following a breach also plays a useful disciplinary role. This threat not only improves enforcement but it also shapes the intertemporal design of incentives and trades. This aspect of our modelling is reminiscent of the relational contracting literature[5] which has already highlighted how the benefits of a continued relationship may be a substitute for missing contracts. Because it assumes that the seller's effort/output, although observable, remains non-verifiable, this literature leaves little room for courts beyond their ability to enforce base payments.[6] This assumption is useful but also extreme. Relational contracts are rarely established in a vacuum and courts have often enough information to specify some aspects of trade. This basic framework of the relational contracting literature has thus been extended to study how explicit incentives may crowd out implicit ones (Baker *et al*.,

3. Dynamic lending relationships, sovereign debt agreements and trade with developing countries, represent other well-suited applications.

4. Such an alleged scam is sometimes referred to by practitioners as *"spiking the job"*. Its attractiveness might come from the fact that it triggers only civil rather than criminal charges if a little bit of the work has already been completed.

5. See Macaulay (1963), Bull (1987), MacLeod and Malcomson (1989, 1998) for earlier contributions and Malcomson (2012) and Gibbons (2005a) for recent surveys.

6. We do not model the role of contract enforcer involved into contractual relationship as a player. For such issue, see Maskin and Tirole (2004) and Rahman (2012).

1994; Schmidt and Schnitzer, 1995; Bernheim and Whinston, 1998; Pearce and Stachetti, 1998; Iossa and Spagnolo, 2009; Li and Matouschek, 2013; Itoh and Morita, 2015).

In this article, we consider how explicit contracts may be a complement (rather than a substitute) to implicit relationship, by delineating the repeated game to be played through relational agreements. This calls for embedding insights from the relational contracting literature into a mechanism design framework where opportunistic behaviours are jointly controlled through explicit remedies and self-enforcing agreements. This path, that we follow thereafter, has already been pointed out by in the law and economics literature (Baker and Choi, 2014). In our article, the constraints imposed by the enforceability of contracts, be it through public remedies following breaches or by the perspective of a continued relationship, are added to the more familiar incentive and participation constraints from mechanism design to fully characterize the set of feasible allocations.

Enforcement constraints affect the nature of output distortions. While the dynamic mechanism design literature emphasizes that distortions never arise for the most efficient type,[7] the interesting dynamics are here reversed: preventing the high-cost seller from *"take the money and run"* leads to decreasing distortions for the efficient, low-cost seller.[8]

To highlight the interaction between enforcement constraints and more familiar incentive compatibility constraints in its simplest form, we assume below that the seller has private information on his persistent cost. Under full commitment and costless enforcement, Baron and Besanko (1984) demonstrate that, under those circumstances, the optimal long-term contract could be implemented by the infinite replica of the optimal short-term contract, leading to stationary trades. On the other side of the spectrum, assuming that only relational contracts are feasible, Levin (2003) shows that the optimal relational contract is again stationary when types are either common knowledge or private information but independently drawn over time. In our setting, a persistent, high-cost seller may adopt the same behaviour as a low-cost one in earlier periods of the relationship before breaching the contract. A key contribution of this article is to demonstrate how such a *"take-the-money-and-run"* strategy is optimally controlled through growing trades.[9]

Horner (2002), Fong and Li (2010), and Halac (2012) have also addressed how persistence of types affects relational contracts. These authors all investigate how private information is revealed over time and how it determines time-varying stakes.[10] Malcomson (2016) shows that full separation is not possible in Levin (2003) model with persistent private information. Respecting enforcement constraints requires some pooling. This pattern of information revelation is reminiscent of that found in the dynamic contracting literature with short-term contracts (Laffont and Tirole, 1993, Chapter 9). There, the *"ratchet effect"* calls for bonuses to reward earlier information revelation by a low-cost seller but it also exacerbates a so-called *"take-the-money-and-run"* strategy which is found attractive by a high-cost seller.[11] With relational

7. See Baron and Besanko (1984) and Battaglini (2005) among others.

8. From a formal point of view, those distortions are determined by the Lagrange multipliers of the enforcement constraints. These multipliers form a decreasing sequence that eventually attains zero as the *"take-the-money-and-run"* strategy becomes less relevant over time. This analysis requires a construction of the optimal contract as the solution to an optimization problem with an infinite number of enforcement constraints; a point which necessitates a careful use of duality theory in infinite-dimensional spaces, an approach based on the work of Dechert (1982).

9. Building on Levin (2003), Kwon (2013) derives the optimal relational contract with persistent shocks and shows that it is no longer stationary. There is also a related literature on dynamic moral hazard problems with persistent shocks (Kwon, 2015; Fuchs, 2007).

10. The non-stationary of relational contracts may also come from learning persistent types as in a model of the labour market proposed by Yang (2012) or intertemporal insurance concerns as in Hemsley (2013).

11. In Laffont and Tirole (1993, Chapter 9), the expression *"take-the-money-and-run"* is, we believe, somewhat misused and this might have created some confusion in the theoretical literature. Indeed, when only spot contracts are

contracts, this tension implies some pooling. In our setting, commitment allows the buyer to better control this *"take-the-money-and-run"* strategy by means of growing stakes while still inducing information revelation upfront. The canonical model of relational contracting with persistent types in Malcomson (2016) and our model thus stand in sharp contrast in terms of how information is revealed over time. As a result, the optimal mechanism in our asymmetric information context cannot be implemented with relational contracts.

1.1.5. Organization. Section 2 presents the model. Section 3 describes the set of allocations that are incentive feasible and enforceable under asymmetric information. Section 4 characterizes the optimal contract and provides conditions for a pattern of growing trades. Section 5 discusses the results of our findings, allowing either one-sided opportunism or renegotiation. Section 6 draws some implications from our findings for organization theory. Proofs are relegated to an Appendix.

2. THE MODEL

2.1. *Basics*

2.1.1. Preferences. We consider an infinitely repeated relationship between a buyer (the principal or *she*) and a seller (the agent, *he*) who provides a service or good on her behalf. Time is indexed by $\tau \geq 0$ and we denote by $\delta < 1$ the common discount factor.

A *trade profile* is an infinite array of payments and (non-negative) outputs $(\mathbf{t},\mathbf{q}) \equiv \{(t_\tau, q_\tau)\}_{\tau=0,..,\infty}$ over this long-term relationship. Both the buyer and the seller have quasi-linear utility functions defined over trade profiles. Their discounted payoffs are, respectively, given by:

$$(1-\delta)\sum_{\tau=0}^{\infty}\delta^{\tau}(S(q_\tau)-t_\tau) \text{ and } (1-\delta)\sum_{\tau=0}^{\infty}\delta^{\tau}(t_\tau-\theta q_\tau).$$

The buyer's benefit function S is differentiable, increasing and concave ($S' > 0 > S''$) with $S(0) = 0$. To ensure positive outputs under all circumstances below, we require that $S'(0)$ is sufficiently large, though bounded. The set of feasible outputs is an interval $\mathcal{Q} = [0, \overline{Q}]$ with $\overline{Q}$ sufficiently large enough to ensure interior solutions.

2.1.2. Information. The seller has private information about his cost parameter θ which takes values in $\Theta = \{\underline{\theta}, \overline{\theta}\}$, with $\Delta\theta = \overline{\theta} - \underline{\theta} > 0$). This parameter is persistent over the whole relationship and is revealed to the seller before contracting. Let $\nu > 0$ (resp. $1-\nu > 0$) be the probability that the seller has a low (resp. high) cost. Let also $\mathbb{E}_\theta(\cdot)$ denote the expectation operator.

2.2. *Costless enforcement*

As benchmarks, we consider the simple case of costless enforcement, first with complete and then with asymmetric information.

feasible, there is by definition no *"run"* because the agent is not bound to any future contract. Not contracting in the future is a valid option that does not trigger remedies. The notion of a *"run"* is only relevant when parties may breach an agreement over its course and such breach is subject to liquidated damages and other remedies.

2.2.1. Complete information. Under complete information, it is routine to show that the first-best outcome is implemented by means of a stationary contract $(t^{fb}(\theta), q^{fb}(\theta))$ (or a stationary rent-output allocation $(U^{fb}(\theta), q^{fb}(\theta))$) which is defined as:

$$S'(q^{fb}(\theta)) = \theta \text{ and } U^{fb}(\theta) = t^{fb}(\theta) - \theta q^{fb}(\theta) = 0 \quad \forall \theta \in \Theta.$$

At the first best, the buyer's marginal benefit must equal the seller's marginal cost. Given that the buyer has all bargaining power, she can fully extract the seller's rent. This allocation is infinitely repeated over time.

2.2.2. Asymmetric information. The *Revelation Principle* applies in a dynamic trading environment when parties commit to a long-term contract (Baron and Besanko, 1984). There is no loss of generality in restricting the analysis to direct and truthful revelation mechanisms that stipulate the seller's payments and outputs in each period as a function of his upfront report on cost. Such a contract is an infinite sequence $\mathcal{C} = \{(\mathbf{t}(\theta), \mathbf{q}(\theta))\}_{\theta \in \Theta}$.

A second important insight due to Baron and Besanko (1984) is that, among all possible payments profiles that may implement the optimal allocation, one possibility is to rely on the infinite replica of the optimal static contract. This contract entails first-best production for the low-cost seller and the usual *Baron–Myerson distortion* for the high-cost seller's output $q^{bm}(\overline{\theta})$ (which remains positive provided that $S'(0)$ is large enough):

$$S'(q^{bm}(\overline{\theta})) = \overline{\theta} + \frac{\nu}{1-\nu} \Delta\theta.$$

Under asymmetric information, the buyer's marginal benefit must equal the seller's *virtual cost* $\overline{\theta} + \frac{\nu}{1-\nu}\Delta\theta$. A downward output distortion for the high-cost seller reduces the information rent left to the low-cost type. As familiar in such screening environments, the rents for both types are expressed as:

$$U^{bm}(\underline{\theta}) = \Delta\theta q^{bm}(\overline{\theta}) > U^{bm}(\overline{\theta}) = 0.$$

Among other possible payments profiles that may implement this allocation, the following stationary prices offer a convenient benchmark for the rest of the analysis:

$$t^{bm}(\underline{\theta}) = \underline{\theta} q^{fb}(\underline{\theta}) + \Delta\theta q^{bm}(\overline{\theta}) \text{ and } t^{bm}(\overline{\theta}) = \overline{\theta} q^{bm}(\overline{\theta}).$$

2.3. *Costly enforcement: setting the stage*

Parties can fully commit to a long-term contract stipulating a trade profile. Yet, at any date, the mere play of this mechanism may be subject to opportunistic behaviours. A party breaches the contract if his current benefit of doing so exceeds the cost. This cost includes the foregone opportunities for future trades (the private side of enforcement) but also the legal remedies that this party has to pay for not fulfilling obligations (the public side). A contract stipulates prices and quantities in all periods, but trade itself is at risk.

To control such bilateral opportunistic behaviours, the parties will find it useful to contract on two prices for each trading period τ—a *pre-production* payment $t_{1,\tau}$ (required by the seller before producing the good) and a *post-delivery* payment $t_{2,\tau}$, required by the seller after the delivery of the good.[12] The pre-production payment helps control the buyer's incentives to pay for delivery.

12. For simplicity, we assume that there is no discounting between those sub-periods.

After having pocketed the pre-delivery payment $t_{1,\tau}$, the seller may choose not to deliver the quantity q_τ. This deviation is attractive if $t_{1,\tau}$ is large enough. The post-trade payment helps control the seller's incentives to deliver the good as required. After the delivery of a quantity q_τ, the buyer may not pay the post-delivery price $t_{2,\tau}$ if this payment is now too large.

A direct mechanism is now an infinite triplet $\mathcal{C}=\{(\mathbf{t}_1(\theta),\mathbf{t}_2(\theta),\mathbf{q}(\theta))\}_{\theta\in\Theta}$ stipulating pre- and post-delivery payments as well as outputs in each period. The total payment to a seller reporting type θ is denoted $t_\tau(\theta)=t_{1,\tau}(\theta)+t_{2,\tau}(\theta)$ at date τ.

Denote by $K\geq 0$ (resp. $L\geq 0$) the remedy paid by the buyer (resp. seller) in case she (resp. he) breaches the agreement.[13] When remedies are infinite ($K=L=+\infty$), enforcement is perfect and the optimal contract is the infinite replica of the Baron–Myerson allocation. To focus on less trivial cases, we will thus assume finite penalties ($K,L<+\infty$).[14] As we will show below, the availability of both pre-production and post-delivery payments allows the parties to shift the force of a penalty from one side of the transaction to the other. The critical constraint on public enforcement is thus the sum of the available penalties, that we define as $M=K+L$.

2.3.1. **Timing.** The contracting game unfolds as follows.

(1) Prior to any trading, at date $\tau=0^-$, the seller privately learns his cost parameter θ. The buyer then offers a mechanism $\mathcal{C}$. The seller, in turn, accepts or rejects the offer. If he accepts, the seller reports a type $\hat{\theta}$; if he rejects, both parties receive their reservation values, normalized to zero.
(2) Each trading period at date $\tau\geq 0$ is as follows.

- The buyer offers the pre-delivery payment $t_{1,\tau}(\hat{\theta})$.
- The seller produces $q_\tau(\hat{\theta})$ or breaches the contract and pays the remedy L.
- If $q_\tau(\hat{\theta})$ is produced as required, the post-delivery payment $t_{2,\tau}(\hat{\theta})$ is paid by the buyer or she breaches the contract and pays the remedy K. Following breach on either side of the transaction, the contract is terminated.[15]

2.3.2. **Notation.** In order to express various feasibility constraints in compact form, we now define the per-period value of the seller's output, his rent, and the buyer's payoff going *forward* from date τ on as, respectively,

$$q_\tau^+(\theta)=(1-\delta)\sum_{s=0}^{\infty}\delta^s q_{\tau+s}(\theta),$$

$$U_\tau^+(\theta)=(1-\delta)\sum_{s=0}^{\infty}\delta^s(t_{\tau+s}(\theta)-\theta q_{\tau+s}(\theta)),\quad V_\tau^+(\theta)=(1-\delta)\sum_{s=0}^{\infty}\delta^s\left(S(q_{\tau+s}(\theta))-t_{\tau+s}(\theta)\right).$$

13. Such remedies can be viewed as expected remedies that incorporate the probability that certain terms are not enforced by courts, as might be the case, for example, under the *"speculative loss"* doctrine. This penalty limitation is assumed to be exogenously set by the court.

14. Our formulation and results below work the same way if penalties are liquidated damages paid by the breaching party to its partner, or instead are paid to a third party (or destroyed).

15. We thus follow here Abreu (1988) and the literature on relational contracts (Levin, 2003; Halac, 2012 among others) in specifying that the worst equilibrium is played following a breach, assuming as we have that there is no renegotiation on or off the equilibrium path.

We may also define the seller's *backward* output and per-period average rent leading up to date τ, assuming the seller breaches the contract at τ after accepting the pre-production payment, $t_{1,\tau}$. These *backward* output and rent functions are, respectively,[16]

$$q_\tau^-(\theta)=(1-\delta)\sum_{s=0}^{\tau-1}\delta^s q_s(\theta),$$

$$U_\tau^-(\theta)=(1-\delta)\sum_{s=0}^{\tau-1}\delta^s(t_s(\theta)-\theta q_s(\theta))+\delta^\tau(1-\delta)(t_{1,\tau}(\theta)-L).^{17}$$

To simplify the formulation of the enforcement constraints, we also define the *enforcement surplus* (an expression that we will make clear below) as

$$\Psi(\theta,\mathbf{q}_\tau(\theta))=(1-\delta)\left(\sum_{s=1}^{\infty}\delta^s(S(q_{\tau+s}(\theta)-\theta q_{\tau+s}(\theta))-\theta q_\tau(\theta)+M\right), \qquad (2.1)$$

where $\mathbf{q}_\tau(\theta)=\{q_s(\theta)\}_{s\geq\tau}$ is an intertemporal trade profile starting at date τ.

2.4. *Reinterpretation of the model/discussions of our assumptions*

Instead of viewing quantity and price as being imperfectly enforceable, we could have assumed that it is costless to write a contract stipulating how much quantity is traded and at what price but that quality is instead non-verifiable. Suppose, for example, that the good may either be of a high or low quality, and quality, although observable, remains non-verifiable unless the court is called upon for inspection. A high quality good yields surplus to the buyer and is costly (but efficient) to produce; a worthless quality good is costless to produce. The seller now breaches the contract when supplying the low quality. The buyer breaches the agreement when not paying following his claim that the supplied good has low quality. Disputes arise whenever the buyer and the seller disagree on their assessments of quality. The court is called upon to solve this dispute. Following the court's inspection, the seller pays a remedy if the quality is low, while the buyer does so if the quality is high but he pretended the opposite to avoid due payment.

This interpretation of the model is much in lines with the existing relational contracting literature. As in this literature, the seller's incentives to produce and the buyer's incentives to pay for high quality cannot be contractually specified. It must be equilibrium behaviour for the seller to offer a high quality and for the buyer not to argue about quality. Again, the continuation value of the relationship helps to curb bilateral opportunism; the private side of enforcement matters.

Compared with the relational contracting literature, two novel ingredients are added when quantities are contractible. First, since a court is already present to enforce such contracts, we posit that it can also assess a disputed quality by inspecting. Remedies are now also useful tools to curb opportunism; the public side of enforcement also matters.[18] Our assumption of bounded penalties for breach is then consistent with the legal doctrine and with existing evidence. Our

16. The following identities hold: $U_\infty^-(\theta)\equiv U_0^+(\theta)$ and $q_\infty^-(\theta)\equiv q_0^+(\theta)$ $\forall\theta\in\Theta$ where $U_0^+(\theta)$ (resp. $q_0^+(\theta)$) is the seller's intertemporal rent (resp. output) over the whole relationship.

17. We adopt the convention that $\sum_{s=0}^{-1}y_s\equiv 0$ for any sequence $\mathbf{y}$.

18. There is a vast literature on remedies and their role on litigation (Shavell, 1980; Bebchuk, 1984; Reinganum and Wilde, 1986 among others). In contrast, we view dispute resolution as a *black-box*.

goal is not to endogenize the levels of remedies. We take these levels as given and derive the consequences of such limited remedies on contracting patterns.

Secondly, the possibility of choosing contractual stakes in each period allows parties to *ex ante* control the game they will be playing. This new feature of the modelling transforms the study of relational contracts into a mechanism design exercise.

In contrast with the relational contracting literature, and much in spirit of the dynamic mechanism design literature,[19] we start with the natural assumption that parties can fully commit to a long-term contract stipulating quantities and prices at all trading periods. Our article thus shares with the dynamic mechanism design literature a common concern for the intertemporal design of incentives. This literature has assumed that drafting a new contract at any point of time is infinitely costly, while transactions in each period are costlessly enforced. By instead assuming that transactions are subject to opportunistic behaviour and that remedies for breach are limited, we put on the front stage the enforcement of transactions in any given period. This focus offers a better account of the contractual environment depicted by law scholars.[20]

The main focus of this article is the setting in which renegotiation is not possible.[21] This assumption has several justifications that talk both to its theoretical and to its empirical relevance. On the theory side, this assumption allows us to describe an upper bound on the possible gains from trade that can be achieved under asymmetric information and costly enforcement. The consequences of renegotiation are already well known from the work of Dewatripont (1989), Hart and Tirole (1988), Laffont and Tirole (1993, Chapter 10), Rey and Salanié (1996), and Maestri (2011). There, a long-term contract can be breached at no cost to reach a Pareto-improving new agreement as information is revealed through earlier performance. This issue is orthogonal to the enforcement of transactions in any given period that is the focus of our analysis. There is indeed no conflict in assuming simultaneously that prices and quantities traded on path are committed to *ex ante* and that those prices and quantities cannot be delivered if parties do not abide to the contract. Additionally, assuming commitment avoids the intricacies arising with gradual information revelation, since all information is now revealed upfront in the first-period of the relationship.

Commitment is also a standard assumption in the law and economics literature that serves as a background of our study. For instance, Edlin (1998) assumes that *"… renegotiation is impossible prior to breach decision",* while Shavell (2004), p.315) stresses that renegotiation fails because of time constraints and asymmetric information. From practical viewpoint, commitment is a way to avoid reputation and legal costs or to mitigate the consequences of inefficient *ex ante* investments that would be caused otherwise. Again, the construction sector shows ample evidence that contractual breaches involve termination of the relationship; renegotiation being rather a rare event or almost absent. Practitioners in the field have spent much effort writing guides on how to behave in a termination.[22]

In practice, competition on both sides of the market may also provide an effective commitment to terminate a relationship once it has been breached, thereby hindering renegotiation. To illustrate, suppose that the seller has not complied and that remedies have been paid; the buyer could immediately turn to another producer to complete the project in a market context, a situation

19. Baron and Besanko (1984), Battaglini (2005), Pavan *et al.* (2014), Escobar and Toikka (2013).

20. Some authors have nevertheless stressed the limited ability of courts to enforce obligations (Laffont and Martimort (2002, Chapter 9); Guasch, Laffont and Straub, 2003; Schwartz and Watson, 2004; Kvaløy and Olsen, 2009; Doornik, 2010).

21. For completeness, Section 5.2 below discusses the possibility of renegotiation.

22. *Termination and suspension of construction contracts. A Guide.* Available at http://www.out-law.com/en/topics/projects–construction/construction-contracts/termination-and-suspension-of-construction-contracts.

illustrated by Macchiavello and Morjaria (2015) in their analysis of the Kenyan market for roses where oversupply is pervasive and punishments take the form of trade disruption. In the construction sector that motivates our study, competition prevails among both contractors and subcontractors. When a subcontractor fails to deliver, renegotiation does not occur. Instead, the contractor typically hires another firm to complete the work. Legal cases illustrating such situations abound. In *RDP Royal Palm Hotel vs. Clark Construction Group* (Callahan, 2009, p. 199–200)), *RDP* hired another contractor and sued *Clark* for breach. In *Saxon Construction vs. Masterclean of North Carolina* (Hinze, 2001, p. 246), a subcontractor, *Masterclean*, failed to complete the work which led *Saxon Construction* to hire another subcontractor.

3. IMPLEMENTABLE ALLOCATIONS

This section characterizes the set of feasible allocations when enforcement is costly. This set is constrained by the usual agent's incentive compatibility and participation constraints but also, and this is the novelty of our framework compared with more standard mechanism design environments, by a new set of dynamic enforcement constraints.

3.1.1. Seller's participation constraints. A seller with type θ finds the mechanism C *individually rational* when the following *interim* participation constraint holds:

$$U_0^+(\theta) \geq 0, \quad \forall \theta \in \Theta. \tag{3.1}$$

3.1.2. Buyer's enforcement constraints. The buyer should pay in each period what is due to the seller. If she does not pay for delivery, she incurs a penalty K (public side of enforcement) and the relationship ends (the private side).

Definition 1. *The mechanism $\mathcal{C}$ is **buyer-enforceable** if and only if:*

$$\delta V_{\tau+1}^+(\theta) \geq (1-\delta)(t_{2,\tau}(\theta) - K) \quad \forall \theta \in \Theta \quad \forall \tau. \tag{3.2}$$

The left-hand side represents the buyer's discounted payoff from period $\tau+1$ onwards on the equilibrium path. The right-hand side is her deviation payoff. It takes into account the fact that trade never occurs from date τ on following breach by the buyer.

3.1.3. Seller's enforcement constraints. The seller's enforcement constraints are complex because they interact with incentive compatibility requirements.

Definition 2. *The mechanism $\mathcal{C}$ is **seller-enforceable** if and only if:*

$$U_0^+(\theta) \geq \max_{\hat{\theta} \in \Theta} U_\tau^-(\hat{\theta}) + (\hat{\theta} - \theta) q_\tau^-(\hat{\theta}) \quad \forall \theta \in \Theta, \quad \forall \tau. \tag{3.3}$$

The enforcement constraints (3.3) say that a seller with type θ prefers to choose his targeted contract rather than adopting a *"take-the-money-and-run"* strategy. This strategy consists in mimicking a type $\hat{\theta}$ at all dates $0 \dots \tau-1$, delivering the corresponding output, but breaching the contract at date τ, being only punished from that date onwards.

3.1.4. Incentive compatibility. Taking $\tau=\infty$, the enforceability conditions (3.3) imply the standard, static notion of incentive compatibility:

$$U_0^+(\theta)=\max_{\hat{\theta}\in\Theta} U_0^+(\hat{\theta})+(\hat{\theta}-\theta)q_0^+(\hat{\theta}), \quad \forall\theta\in\Theta. \tag{3.4}$$

From this incentive compatibility condition, it follows that the discounted output over the relationship satisfies a well-known monotonicity condition:

$$q_0^+(\theta) \text{ weakly decreasing.} \tag{3.5}$$

3.1.5. Pooling enforcement constraints. The enforcement surplus (2.1) represents the parties' net gain from enforcing the contract from date τ on. It takes into account future gains from trade but also the foregone penalties from not deviating at date τ.

Definition 3. *The mechanism $\mathcal{C}$ is* ***enforceable*** *if and only if it is both buyer- and seller-enforceable.*

Pooling together the individual enforcement constraints (3.2) and (3.3), we obtain a new set of feasibility conditions.

Lemma 1. *An incentive compatible and individually rational mechanism $\mathcal{C}$ is* ***enforceable*** *if and only if*

$$\Psi(\theta,\mathbf{q}_\tau(\theta))\geq\delta^{-\tau}\max_{\hat{\theta}\in\Theta}\left\{U_0^+(\theta)-U_0^+(\hat{\theta})+(\theta-\hat{\theta})q_\tau^-(\theta)\right\}, \quad \forall\theta\in\Theta, \quad \forall\tau\geq 0. \tag{3.6}$$

Note that Ψ on the left-hand side of equation (3.6) depends only on the aggregate penalty, $M=K+L$. An immediate consequence of Lemma 1 is thus that the distribution of remedies is irrelevant. The buyer, who has all bargaining power, can undo any initial distribution and structure payments so as to internalize fully the consequences of a breach even when it might originate from the seller. Everything thus happens as if the only threat was a breach of contract by the buyer herself and the remedy that would apply for such breach is M.

Although the procedure of pooling enforcement constraints on both sides of the market is reminiscent of Levin (2003)'s approach in his study of bilateral opportunism in relational contracts, the details differ in interesting ways. First, the aggregate pooled constraint in this article must account for the fact that parties are asymmetrically informed. This explains the presence on the right-hand side of equation (3.6) of the possible manipulation of the seller's reports on his type.

3.1.6. "Take-the-money-and-run". It should be intuitive that the most salient enforcement constraints on the seller's side are those of a high-cost type. Indeed, a low-cost seller, if he chooses to lie on his type and produce a quantity $q_\tau(\overline{\theta})$ at a low marginal cost gets an information rent $\Delta\theta q_\tau(\overline{\theta})$ in period τ. This seller has no incentives for early breaches if he wants to pocket these rents over the whole relationship.

Instead, because asymmetric information requires an increase in the price paid to a low-cost type to induce information revelation, a high-cost seller may now find a *"take-the-money-and-run"* strategy particularly attractive. That type may want to mimic a low-cost seller for a few initial periods so as to pocket these large prices before breaching the contract. These incentives are captured in the right-hand side of equation (3.6). Mimicking the low-cost seller forever is not profitable, however, given the standard incentive compatibility condition, so the incentives for breach culminate early in the relationship.

3.1.7. Enforcement under complete information. To deepen our understanding of condition (3.6), suppose for the moment that the seller's cost is common knowledge and, thus, there is no possibility of non-truthful reports on the right-hand side of equation (3.6). In this complete information setting, the enforcement constraints would thus become

$$\Psi(\theta, \mathbf{q}_\tau(\theta)) \geq 0, \quad \forall \theta \in \Theta, \quad \forall \tau \geq 0. \tag{3.7}$$

Feasibility would now require simply that the discounted value of future trades covers the sum of the individual costs for breaching the contract.

Condition (3.6) is similar to (3.7) with the only difference being that, under asymmetric information, the prices paid to the low-cost seller must also account for his information rent. The enforceability constraint is hardened under asymmetric information.

3.1.8. Time-dependence. The existence of a *"take-the-money-and-run"* strategy evoked above already stressed the difference between the earlier periods and the rest of the relationship. To sharpen intuition on this issue and analyse its consequences on contractual dynamics, consider the aggregate enforcement constraint (3.6) in state $\underline{\theta}$. Suppose also that the incentive compatibility constraint (3.4) that prevents a low-cost seller from pretending being a high cost one is binding, a property that will hold at the optimal contract:

$$U_0^+(\underline{\theta}) = U_0^+(\bar{\theta}) + \Delta\theta q_0^+(\bar{\theta}). \tag{3.8}$$

The enforcement constraint (3.6) written in state $\underline{\theta}$ then becomes:

$$\Psi(\underline{\theta}, \mathbf{q}_\tau(\underline{\theta})) \geq \delta^{-\tau} \max\left\{\Delta\theta(q_0^+(\bar{\theta}) - q_\tau^-(\underline{\theta})), 0\right\}. \tag{3.9}$$

By construction, $q_\tau^-(\underline{\theta})$ converges towards $q_0^+(\underline{\theta})$ as τ grows large. If the monotonicity condition (3.5) is strict, the maximum on the right-hand side of equation (3.9) is thus zero. The enforcement constraint (3.9) then reduces to its complete information expression (3.7). In the long run, asymmetric information has no impact on enforcement constraints.

In the short run, however, asymmetric information impacts enforcement. Taking instead $\tau = 0$, the enforcement constraint (3.9) becomes more stringent than (3.7):

$$\Psi(\underline{\theta}, \mathbf{q}_0^+(\underline{\theta})) \geq \Delta\theta q_0^+(\bar{\theta}). \tag{3.10}$$

Reinforcing an intuition given above, the *"take-the-money-and-run"* strategy only matters early in the relationship. In contrast with the case of complete information, enforcement constraints now depend explicitly on time. The beginning and the tail of the relationship do not look the same for the high-cost seller. This feature explains that the optimal contract may not be stationary and provides an intuition for the nature of the trading dynamics under the optimal contract.

4. OPTIMAL DYNAMIC CONTRACT

The buyer's objective is to maximize the discounted net surplus he obtains from trade, subject to the seller's participation, incentive compatibility and the new enforcement constraints:

$$(\mathcal{P}): \quad \max_{(\mathbf{q}(\theta), U_0^+(\theta))} \mathbb{E}_\theta \left((1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \left(S(q_\tau(\theta)) - \theta q_\tau(\theta) \right) - U_0^+(\theta) \right)$$

subject to equations (3.1), (3.4), and (3.6).

4.1. *Complete information benchmark*

To build intuition and provide further comparison with the case of asymmetric information, let us briefly analyze the case of complete information. To this end, suppose *a priori* that the buyer offers a stationary output profile $\mathbf{q} = (q, q, \ldots.)$ in state θ. The enforcement surplus would become:

$$\psi(\theta, q) = \Psi(\theta, \mathbf{q}) = \delta S(q) - \theta q + (1-\delta) M.$$

Observe that ψ is strictly concave in q, achieves a maximum at an output $q_\psi(\theta)$ which is defined as $\delta S'(q_\psi(\theta)) = \theta$ and which is less than $q^{fb}(\theta)$. Furthermore, it can be checked that being concave, positive at $q = 0$ and negative for q large enough, it admits a zero at some positive $q^e(\theta) > q_\psi(\theta)$:

$$\delta S(q^e(\theta)) - \theta q^e(\theta) + (1-\delta) M = 0. \tag{4.1}$$

It follows that $q^e(\theta)$ is the largest stationary output that could be enforced under complete information.

The optimal contract under complete information and limited enforcement can be shown to be stationary but not necessarily first-best.[23] The intuition for stationarity is straightforward. The enforcement constraints (3.7) look the same from any date onward, so optimal distortions are the same in every period. The corresponding optimal output under complete information $q^{ci}(\theta)$ is thus given by the following expression:

$$q^{ci}(\theta) = \min\{q^e(\theta), q^{fb}(\theta)\}.$$

Observe that, for a fixed level of M, the inequality $q^e(\theta) \geq q^{fb}(\theta)$ always holds when δ is close to 1. Enforcement is not an issue if parties care enough about the future, even with zero penalties. In contrast, when the discount factor and the available penalties are sufficiently small, the enforcement constraint is binding. The optimal output in any period is then the greatest output compatible with enforcement, namely $q^e(\theta)$. Production must be reduced below the first best to reduce payments and incentives for breaches.

The optimal contract under complete information can be implemented with a pure relational contract. Seeing how requires a slight extension of the relational contracting literature. Suppose that parties can include into the stage game a *"no trade"* action that would trigger liquidated damages. Doing so expands punishments beyond the foregone value of trade. Following Levin (2003), we can derive the optimal relational contract in such a context. This contract again implements the stationary output $q^{ci}(\theta)$ as above. In other words, the ability to parties to commit does not matter under complete information. Relational contracts suffice. As the analysis below will show us, this is no longer the case under asymmetric information.

23. The proof of existence and stationarity of an optimal contract is available upon request.

4.2. *Implementing the Baron–Myerson outcome*

Turning now to the characterization of the optimal contract under asymmetric information, we first highlight simple conditions that ensure that the infinite replica of the Baron–Myerson's outcome is dynamically enforceable.

Assumption 1.

$$\Delta\theta q^{bm}(\overline{\theta}) \leq \psi(\underline{\theta}, q^{fb}(\underline{\theta})) \text{ and } 0 \leq \psi(\overline{\theta}, q^{bm}(\overline{\theta})).$$

Proposition 1. *Under the optimal contract, the Baron–Myerson outputs* $(q^{fb}(\underline{\theta}), q^{bm}(\overline{\theta}))$ *are implemented in each period if and only if Assumption 1 holds.*

When the gains from trade suffice to prevent breaches, the stationary Baron–Myerson allocation is feasible under limited enforcement. This is so even if the price paid to this low-cost seller is greater than under complete information so as to pay for his information rent. To give more intuition, observe that the first condition in Assumption 1 implies that the enforcement constraint (3.9) holds at date 0 if all payments to the low-cost seller are paid upfront. The most attractive deviation for the buyer consists of not paying that amount at date 0 and immediately breaching the contract; Assumption 1, however, ensures that such a deviation is not valuable. The second condition in Assumption 1 says that the Baron–Myerson output could be enforced if the seller was known to have a high cost parameter.

4.3. *Stationary contracts*

When Assumption 1 fails, the repeated Baron–Myerson outcome is no longer implementable under asymmetric information. In this case, a useful starting point is to consider a stationary contract $\mathbf{q}(\theta) = (q(\theta), q(\theta), \ldots.)$, but with outputs that are modified to satisfy enforcement constraints.[24] The enforcement constraint in state $\underline{\theta}$, namely (3.9), becomes

$$\psi(\underline{\theta}, q(\underline{\theta})) \geq \delta^{-\tau} \max\left\{\Delta\theta(q(\overline{\theta}) - \left(1-\delta^{\tau}\right) q(\underline{\theta})), 0\right\} \quad \forall \tau \geq 0.$$

For any weakly decreasing output profile, this constraint holds at $\tau \geq 1$ if it already holds at $\tau = 0$. Intuitively, with a stationary contract, if any breach were to happen, it should arise as soon as possible. This leads to the simpler requirement

$$\psi(\underline{\theta}, q(\underline{\theta})) \geq \Delta\theta q(\overline{\theta}). \tag{4.2}$$

Denote by Λ the non-negative Lagrange multiplier of the binding enforcement constraint (4.2) in the buyer's optimization programme. This multiplier measures the shadow cost of enforcement. It is important to stress that, because this multiplier determines the optimal screening distortions, the quality of enforcement and contract performances are now linked altogether. Indeed, maximizing the buyer's payoff under the feasibility constraints (4.2) and the restriction to stationary contracts yields the following expressions of the downward output distortions:

$$S'(q^{st}(\underline{\theta})) = \underline{\theta} + \frac{(1-\delta)\Lambda}{\nu + \Lambda\delta}\Delta\theta \text{ and } S'(q^{st}(\overline{\theta})) = \overline{\theta} + \frac{\nu + \Lambda}{1-\nu}\Delta\theta.$$

24. A rationale for this stationarity restriction in the set of feasible contracts is that the buyer is involved in a series of bilateral relationships, facing a population of sellers whose arrivals follow a Poisson process and bilateral contracts remain anonymous and thus independent on the first date at which such bilateral trade occurs.

Both outputs are reduced below the Baron–Myerson levels when the multiplier Λ is positive. To relax the binding enforceability constraint (4.2), the buyer would like to reduce the price paid to a low-cost seller so as to make breach less attractive. Two instruments are used in tandem. First, the buyer procures even less from a high-cost seller than in the Baron–Myerson scenario. This reduces the low-cost seller's information rent and thus his payment. Secondly, the buyer also asks for less output from a low-cost seller which is another strategy to reduce his payment.

With stationary contracts, reducing the low-cost seller's output has nevertheless two conflicting effects. First, as just claimed, it decreases the benefits of a current breach. Secondly, it also reduces surpluses in future trading rounds which, on the contrary, harms enforceability. Those two conflicting roles are disentangled with non-stationary contracts.

4.4. *Growing dynamics*

To characterize the dynamics of optimal contracts, we first introduce some conditions.

Assumption 2.

$$0 < \psi(\underline{\theta}, q^{fb}(\underline{\theta})) < \Delta\theta q^{bm}(\overline{\theta}) \text{ and } 0 \leq \psi(\overline{\theta}, q^{bm}(\overline{\theta})).$$

These conditions preclude the positive results of Proposition 1. From the first inequality, enforcement would not be an issue if the seller was known to have a low cost, *i.e.* $q^{fb}(\underline{\theta}) < q^{e}(\underline{\theta})$. Instead, the second inequality implies that the enforcement surplus in state $\underline{\theta}$ does not suffice to ensure enforceability if the buyer has to pay for the extra rent $\Delta\theta q^{bm}(\overline{\theta})$ that a low-cost seller earns under asymmetric information. Finally, the third inequality means that there would *not* be any enforcement problem if the high-cost seller was asked to produce the Baron–Myerson output under complete information. Taken together, Assumption 2 ensures that only the enforcement constraint (3.9) in state $\underline{\theta}$ may matter at the optimum.

Recall that pooling enforcement constraints for both the seller and the buyer require us to take into account asymmetric information. In particular, the enforcement constraint (3.9) is obtained by summing up the buyer's enforcement constraint for state $\underline{\theta}$ with the seller's enforcement constraint (*"take-the-money-and-run"* strategy) in state $\overline{\theta}$. The first of these constraints is certainly relaxed by reducing the price paid to a low-cost seller. As with stationary contracts, it *a priori* means reducing not only the output of a low-cost seller but also, by incentive compatibility, the output of a high-cost one. Relaxing the second of these constraints requires making the *"take-the-money-and-run"* strategy less attractive. This not only calls for reducing outputs but also for a careful design of payments over time.

Equipped with Assumption 2, we can now characterize contractual dynamics. Theorem 1 shows that the pattern of trades with a low-cost seller entails two distinct phases. In the earlier periods, output continuously increases while remaining below efficiency for both types. Later, in a more mature phase, trade with a low-cost seller entails first-best production and the sole distortion concerns the high-cost seller's output. The optimal contract in the long run exhibits features which are similar to those found under the standard Baron–Myerson scenario *modulo* a modification of the virtual cost that now reflects the magnitude of the enforcement problem.

Theorem 1. *Suppose that Assumption 2 holds. There exists $\tau^* \geq 1$ such that the optimal contract passes through two different phases.*

(1) TRANSITORY PHASE. *For $\tau \leq \tau^*$, the optimal output $q_\tau^{sb}(\underline{\theta})$ of the low-cost seller strictly increases over time but remains below its first-best value:*

$$q^{e}(\overline{\theta}) < q_\tau^{sb}(\underline{\theta}) \leq q^{fb}(\underline{\theta}). \tag{4.3}$$

The enforcement constraint (3.9) is binding at all dates $\tau \le \tau^*$ *and the sequence* $q_\tau^{sb}(\underline{\theta})$ *obeys the recursive condition* $q_{\tau+1}(\underline{\theta}) = \Phi(q_\tau(\underline{\theta}))$ *where the function* $\Phi(q) = S^{-1}\left(\frac{1}{\delta}\left(\overline{\theta}q - (1-\delta)M\right)\right)$ *defined over the interval* $\left[\frac{(1-\delta)M}{\overline{\theta}}, +\infty\right)$*, is increasing, convex and has a unique fixed point* $q^e(\overline{\theta})$.

(2) MATURE PHASE. *For* $\tau > \tau^*$*, the optimal output* $q_\tau^{sb}(\underline{\theta})$ *of a low-cost seller is set at its first-best level:*

$$q_\tau^{sb}(\underline{\theta}) = q^{fb}(\underline{\theta}). \tag{4.4}$$

The enforcement constraint (3.9) is slack.

The high-cost seller always produces the same quantity $q^{sb}(\overline{\theta})$ *which remains below the Baron–Myerson level. Specifically, there exists* $\Lambda_\infty > 0$ *such that*

$$S'(q^{sb}(\overline{\theta})) = \overline{\theta} + \frac{\nu + \Lambda_\infty}{1-\nu}\Delta\theta. \tag{4.5}$$

Only the low-cost seller receives a positive information rent:

$$U_0^{+sb}(\underline{\theta}) = \Delta\theta q^{sb}(\overline{\theta}) > 0 = U_0^{+sb}(\overline{\theta}).$$

The brief analysis of the stationary contracts made in Section 4.3 showed how the buyer is torn between two objectives when he wants to ease contract enforcement. On the one hand, he would like to compress payments and reduce production today, especially from a low-cost seller. On the other hand, keeping a high output from this seller also increases future gains from trade, making it more attractive not to breach the relationship which, in turn, relaxes the current enforcement constraint. With non-stationary contracts, the buyer benefits from the fact that the *"take-the-money-and-run"* strategy becomes less attractive over time to separate those two objectives. In the early periods, the buyer distorts production for the low-cost seller much as what was needed with stationary contracts. This compresses current payments and eases earlier enforcement constraints. Yet, the buyer can also use his commitment power to delay payments for the low-cost seller's rent. There is less need to distort production, up to the point where efficient quantities are traded in the mature phase of contracting.

4.4.1. Intuition for the dynamics. To better understand output distortions, we now construct a *"third-best"* contract which also exhibits a growing dynamics as the optimal contract of Theorem 1 but in a crude way using only two output steps. This construction conveys the main intuition behind the existence of two different contracting phases. Consider the following output profile:[25]

$$\mathbf{q}_{\hat{\tau}^*}(\underline{\theta}) = (\underbrace{q^e(\overline{\theta}), \ldots, q^e(\overline{\theta})}_{\text{for dates } \tau \le \hat{\tau}^*}, \quad \underbrace{q^{fb}(\underline{\theta}), \ldots}_{\text{for dates } \tau > \hat{\tau}^*}) \text{ and } \mathbf{q}_{\hat{\tau}^*}(\overline{\theta}) = (q^{bm}(\overline{\theta}), \ldots, q^{bm}(\overline{\theta})). \tag{4.6}$$

Although the high-cost seller's output is stationary and fixed at the Baron–Myerson level, the low-cost seller's production goes through two phases. For the first $\hat{\tau}^* + 1$ periods, the low-cost

25. Payments are determined by the binding incentive and participation constraints for the low- and high-cost seller, respectively.

seller produces the maximal quantity that can be enforced had the seller been known of a high cost, namely $q^e(\overline{\theta})$. After that, trade is efficient.

For the sake of the argument, we strengthen the first condition in Assumption 2 as:

$$0 < \psi(\underline{\theta}, q^{fb}(\underline{\theta})) + (1-\delta)\underline{\theta}(q^{fb}(\underline{\theta}) - q^e(\overline{\theta})) < \Delta\theta q^{bm}(\overline{\theta}).$$

The right hand side inequality ensures that offering

$$\mathbf{q}_0(\underline{\theta}) = (\underbrace{q^e(\overline{\theta})}_{\text{at } \tau=0}, \underbrace{q^{fb}(\underline{\theta}), \ldots}_{\text{for dates } \tau>0})$$

together with $\mathbf{q}_0(\overline{\theta}) = (q^{bm}(\overline{\theta}), \ldots, q^{bm}(\overline{\theta}))$ would violate the enforcement constraint (3.9) at date 0. In other words, a contract of the form (4.6) is necessarily such that the transitory phase has more than one period, *i.e.* $\hat{\tau}^* > 0$. As we will see below, $\hat{\tau}^*$ is constructed so that the buyer can still enjoy efficient trades with a low-cost seller from date $\hat{\tau}^*+1$ onwards if he is ready to maintain a low output over the first τ^* periods of the transitory phase. Such a distortion allows one to keep the low-cost seller's payments small and make the *"take-the-money-and-run"* strategy of the high-cost seller less tempting.

To make this simple point more formally, observe that Assumption 2 (and thus its strengthening) also implies that $q^{bm}(\overline{\theta}) < q^e(\overline{\theta})$ so that the sequence

$$\delta^{-\tau}\Delta\theta\left(q^{bm}(\overline{\theta}) - (1-\delta^{\tau})q^e(\overline{\theta})\right)$$

is actually decreasing. This monotonicity implies that there exists a first date $\hat{\tau}^* \geq 1$ at which the enforcement constraint (3.9) holds for the profile $(\mathbf{q}_{\hat{\tau}^*}(\underline{\theta}), \mathbf{q}_{\hat{\tau}^*}(\overline{\theta}))$. This τ^* is the first integer such that:

$$\delta(S(q^{fb}(\underline{\theta})) - \underline{\theta}q^{fb}(\underline{\theta})) - (1-\delta)\underline{\theta}q^e(\overline{\theta}) + (1-\delta)M \geq \delta^{-\hat{\tau}^*}\Delta\theta\left(q^{bm}(\overline{\theta}) - (1-\delta^{\hat{\tau}^*})q^e(\overline{\theta})\right). \quad (4.7)$$

Intuitively, if the buyer waits long enough before requesting efficient trades, he will eventually prevent the *"take-the-money-and-run"* strategy. Delaying efficient trades till the end of the transitory phase is thus a first cost of limited enforcement. The second cost of limited enforcement is that output is less than first best over this transitory phase.

With this *"third-best"* contract, the transitory phase is constructed so that the output profile (4.6) also satisfies the enforcement constraints (3.9) at all dates $\tau \leq \hat{\tau}^*$.[26] Intuitively, a contract that offers $q^e(\overline{\theta})$ to a low-cost seller in the earlier periods is certainly immune to the possibility that the high-cost seller adopts the *"take-the-money-and-run"* strategy at those dates. Indeed, even if the seller was known to have a high cost, the benefits of breaching a contract that would request $q^e(\overline{\theta})$ from such seller would just cover the cost of the breach.

26. To see how, just multiply equation equation (4.7) by $\delta^{\hat{\tau}^*-\tau}$ and equation (4.1) (taken for $\theta = \overline{\theta}$) by $1-\delta^{\hat{\tau}^*-\tau}$ and sum the two conditions so obtained to get, after simplifications, that the enforcement constraint (3.9) at all dates τ such that $\tau \leq \hat{\tau}^*$ also holds:

$$(1-\delta)\left(\sum_{s=1}^{\hat{\tau}^*-\tau}\delta^s(S(q^e(\overline{\theta})) - \theta q^e(\overline{\theta})) + \sum_{\hat{\tau}^*-\tau+1}^{\infty}\delta^s(S(q^{fb}(\underline{\theta})) - \underline{\theta}q^{fb}(\underline{\theta}))\right) - (1-\delta)\underline{\theta}q^e(\overline{\theta}) + (1-\delta)M$$
$$\geq \delta^{-\tau}\Delta\theta\left(q^{bm}(\overline{\theta}) - (1-\delta^{\tau})q^e(\overline{\theta})\right).$$

4.4.2. Output distortions. The optimal contract differs from the *"third-best"* contract just constructed along two dimensions. First, the buyer compresses payments to the low-cost seller by reducing his information rent. This means implementing an output $q^{sb}(\overline{\theta})$ lower than the Baron–Myerson level. Since, under Assumption 2, the only concern is to prevent the high-cost seller's *"take-the-money-and-run"* strategy, and since all periods are the same, an equal incentive distortion is imposed on his own output at all dates. By imposing downward distortions below the Baron–Myerson level, the buyer facilitates enforcement in earlier periods and shortens the transitory phase. This minimizes the distortions of not implementing efficient trades earlier on. The term on the right hand side of equation (4.5) can now be interpreted as a *modified virtual cost* in the parlance of Myerson. This modified virtual cost encapsulates the cost of enforcement through its dependency on Λ_∞ which is the cumulative multiplier of all enforcement constraints over the transitory phase.

Secondly, the buyer also relaxes enforcement constraints over the transitory phase with less extreme downward distortions of the low-cost seller's output. At the optimum, the buyer implements outputs $q_\tau^{sb}(\overline{\theta})$ which is close to $q^e(\overline{\theta})$ for the transitory phase but remains above. As times passes over the transitory phase, those distortions are less significant in response to enforcement issues of decreasing magnitude.[27]

Compared with the case of stationary contracts, optimal distortions are spread over the first $\hat{\tau}^*$ periods. Indeed, an optimal stationary contract would cause the enforcement constraint (3.9) to bind with a very high shadow cost at date $\tau=0$, but to be slack for all other periods. It is more profitable for the buyer to spread the cost of enforceability across the first periods, but this requires a non-stationary allocation. The construction thus resembles the output profile (4.6). There also, enforcement constraints are binding at all dates $\tau \leq \hat{\tau}^*$.

In the mature phase of contracting, output distortions with the low-cost seller are no longer needed. The contract becomes stationary. In contrast with what arises when restricting *a priori* to stationary contracts, a low-cost seller now produces efficiently. Our model thus predicts an increasing dispersion of outputs over time, contrary to what happens with stationary contracts.

The parameter Λ_∞ that characterizes output distortions is actually the sum of the Lagrange multipliers for all binding enforcement constraints over the transitory phase. Echoing our findings with stationary contracts, this parameter again links altogether the nature of the screening distortions and the quality of enforcement. Remarkably, the output distortion for a high-cost seller obeys a *modified Baron–Myerson* formula (4.5) that illustrates how virtual costs must now be modified with limited enforcement. A greater value of Λ_∞ translates into greater output distortions.

4.4.3. Backloaded payments. To ease enforceability, the buyer supplements output distortions with payments that make the *"take-the-money-and-run"* strategy less attractive. To this end, the buyer backloads payments to the low-cost seller while still keeping an overall price large enough to pay for the latter's information rent.

That the enforcement constraints (3.9) bind over the transitory phase puts some structure on the intertemporal profile of payments. This stands in sharp contrast with the case of costless enforcement (Baron and Besanko, 1984). There, only the values of the overall intertemporal payments to both types are known from the binding incentive and participation constraints and, although payments can be chosen to be stationary, there is much leeway beyond that specific

27. The growing phase of the optimal contract is somewhat reminiscent of the reputation literature (Sobel, 1985; Ghosh and Ray, 1996; Watson, 1999, 2002; Halac, 2012). There, relationships might start *"small"* to ease reputations building when there is uncertainty on traders' degree of opportunism.

choice. The next proposition characterizes payments that implement the optimal allocation described in Theorem 1.

Proposition 2. *Suppose that Assumption 2 holds. The following payments implement the optimal contract.*

(1) Pre-delivery payments cover the seller's penalty for breach:

$$t^{sb}_{1,\tau}(\theta) = L \quad \forall \theta \in \Theta, \quad \forall \tau \geq 0. \tag{4.8}$$

(2) The high-cost seller's payment is constant over time:

$$t^{sb}_{\tau}(\overline{\theta}) = \overline{\theta} q^{sb}(\overline{\theta}) \quad \forall \tau \geq 0. \tag{4.9}$$

(3) The low-cost seller's payment is increasing over the transitory phase and constant over the mature phase:

$$t^{sb}_{\tau}(\underline{\theta}) = \begin{cases} \overline{\theta} q_{\tau}(\underline{\theta}) & \forall \tau \leq \tau^* - 1, \\ \underline{\theta} q_{\tau}(\underline{\theta}) + \frac{\delta^{-\tau} \Delta\theta}{1-\delta} \left(q^{sb}(\overline{\theta}) - q^{-}_{\tau}(\underline{\theta}) \right) & \tau = \tau^*, \\ \underline{\theta} q^{fb}(\underline{\theta}) & \forall \tau > \tau^*. \end{cases} \tag{4.10}$$

Choosing a pre-delivery payment that just covers the seller's penalty for breach (condition (4.8)) is akin to redistributing remedies between the buyer and the seller. Everything happens as if the former now pays all remedies ($K' = M$) in case she breaches the agreement, while the latter pays nothing ($L' = 0$) and is only subject to the private side of enforcement if he breaches. In other words, the buyer who has all bargaining power can undo any initial allocation of remedies through a convenient design of payments without modifying the nature of the enforcement constraints. With such choice, the post-delivery price becomes:

$$t^{sb}_{2,\tau}(\theta) = -L + t^{sb}_{\tau}(\theta) \quad \forall \theta \in \Theta, \quad \forall \tau.$$

The contract must also extract the high-cost seller's rent so that his interim participation constraint is binding at the optimum. The payments (4.9) achieve this goal by imposing the stricter requirement of a binding participation constraint in every single period. With such scheme, a high-cost seller who revealed his type at the start of the relationship certainly will not breach the contract at any future date because he is just indifferent between abiding to the terms of the contract and breaching in each period.

Preventing the high-cost seller's *"take-the-money-and-run"* strategy requires one to keep the high-cost type indifferent between telling the truth (and making zero profit each period) and pretending to be more efficient. In particular, the payment for the low-cost seller must be increased to cover the high-type's cost of producing the high output over the transitory phase. In other words, the high-cost seller makes also zero profit in each period of the transitory phase if he pretends to be efficient. Countervailing incentives are neutralized with this scheme.[28]

28. Contrary to the earlier literature (Lewis and Sappington, 1989), countervailing incentives here apply to different types and full separation remains possible.

The final step consists in checking that the payments in (4.10) offered over the mature phase, together with those given over the transitory phase, ensure that the low-cost seller receives enough information rent to reveal his type truthfully at date 0.

From Proposition 2, the dynamics of the low-type seller's *current* payoff are given by

$$U_\tau(\underline{\theta})=t_\tau(\underline{\theta})-\underline{\theta}q_\tau(\underline{\theta})=\begin{cases}\Delta\theta q_\tau(\underline{\theta}) & \forall\tau\leq\tau^*-1,\\ \frac{\delta^{-\tau}\Delta\theta}{1-\delta}\left(q^{sb}(\overline{\theta})-q_\tau^-(\underline{\theta})\right) & \tau=\tau^*,\\ 0 & \forall\tau>\tau^*.\end{cases} \tag{4.11}$$

Because output increases over the transitory phase, the low-cost seller's per-period payoff thus also grows over this phase. The main purpose of reducing payoffs for this type is indeed to prevent the high-cost type from *"taking the money and run"*. Because of discounting, it is more efficient to reduce payments to a low-cost seller earlier on. Later, that payoff must increase to provide enough rent over the whole relationship to induce information revelation.[29] The last non-zero payoff $U_{\tau^*}(\underline{\theta})$ corrects for this effect. After date τ^*, both types enjoy no rent in each period.

4.4.4. Length of the transitory phase. This length trades off two competing effects that can be best seen by again considering the third-best contract that uses two output steps which we discussed above. First, increasing the number of periods where the enforcement constraints (3.9) are binding allows the buyer to keep the high-cost seller's output close to the Baron–Myerson level. However, doing so also forces her to keep inefficient trades with a low-cost seller for too long. We now provide bounds on the length of this transitory phase.

Proposition 3. *Suppose that Assumption 2 holds. The length of the transitory phase τ^* satisfies the following bounds:*

$$\frac{\ln\left(\frac{\Delta\theta\left(q^{bm}(\overline{\theta})-q^e(\overline{\theta})\right)}{\psi(\underline{\theta},q^{fb}(\underline{\theta}))-\Delta\theta q^e(\overline{\theta})}\right)}{\ln(\delta)}>\tau^*\geq\frac{\ln\left(\frac{\Delta\theta\left(q^{bm}(\overline{\theta})-q^{fb}(\overline{\theta})\right)}{\psi(\underline{\theta},q^{fb}(\underline{\theta}))-\Delta\theta q^{fb}(\overline{\theta})}\right)}{\ln(\delta)}-1. \tag{4.12}$$

When $q^{fb}(\underline{\theta})-q^e(\overline{\theta})$ is small enough, both the right hand and the lefthand sides of (4.12) grow large. Intuitively, when $q^{fb}(\underline{\theta})$ is *"almost"* enforceable, the optimal contract remains close to $q^e(\overline{\theta})$ for a very long time before moving towards the nearby first-best level.

When δ goes to 0, only the public side of enforcement matters. Future gains from trade are of no help to prevent opportunism. In that limiting case, the numerators on both sides of equation (4.12) remain bounded while the denominator goes to infinity. The transitory phase lasts only one period when Assumption 2 holds. The enforcement constraint (3.9) at date 0 then almost reduces to a simple constraint on current output and forward rent:

$$M-\underline{\theta}q_0(\underline{\theta})\approx_{\delta\approx 0}\Delta\theta q^{sb}(\overline{\theta}).$$

When the future no longer matters, the relationship is almost static with the sole possibilities for enforcement coming from public remedies. Output distortions for the low-cost seller are then concentrated on the first period of the relationship.

29. This point is somewhat related to Board's (2011) findings that, to prevent agent's opportunism, a principal must give to his agent a sufficiently large rent which comes both in terms of payments today and promised future payments.

When instead δ goes to 1, the private side of enforcement becomes *de facto* the best vehicle to sustain the relationship. The bounds in equation (4.12) give much less information. This is precisely the scenario where enforcement constraints might be slack. For δ close to 1, Assumption 1 simplifies to:

$$S(q^{fb}(\underline{\theta}))-\underline{\theta}q^{fb}(\underline{\theta})\geq\Delta\theta q^{bm}(\overline{\theta}), \tag{4.13}$$

implying that the gains from trade with a low-cost seller have to only exceed the rent left to that type to ensure enforcement.

5. ROBUSTNESS

5.1. *One-sided breach*

When opportunism is only one sided, there exist pre- and post-delivery payments that implement the Baron–Myerson outcome at no extra cost. One set of instruments is used to curb opportunism by the relevant party while the other helps implementing the Baron–Myerson payments. To illustrate this point, we first consider the case where only the buyer is opportunistic. In each period, pre- and post-delivery payments can be constructed so as to induce the buyer to perform. More precisely, consider the following payments:

$$t_{2,\tau}(\theta)=K+\frac{\delta}{1-\delta}(S(q^{bm}(\theta))-t^{bm}(\theta)) \text{ and } t_{1,\tau}(\theta)+t_{2,\tau}(\theta)=t^{bm}(\theta), \quad \forall\tau\geq 0, \quad \forall\theta\in\Theta.$$

Post-delivery payments are such that the buyer is always indifferent between performance and breaching, in the latter case paying the corresponding remedies and losing his future gains from trade at the Baron–Myerson allocation. In other words, the mechanism so constructed is buyer-enforceable and constraints (3.2) always hold. Pre-delivery payments are such that the seller receives the corresponding Baron–Myerson payments.

The analysis of the case where the informed seller misbehaves is left to the Appendix. Details differ but the main idea remains. With one-sided opportunism, there is enough freedom in designing an intertemporal profile of payments to prevent one-sided breaches at no cost. We can thus conclude:

Theorem 2. *One-sided opportunism is costless.*

5.2. *Renegotiation*

In our main analysis, we assumed that parties might not be able to perfectly enforce a transaction in any given period. At the same time, parties can instead commit not to renegotiate their *ex ante* agreement, an assumption that was also discussed at length in Section 2.4.[30] In some circumstances, a commitment not to renegotiate may be less reasonable. To illustrate, large infrastructure projects often require specific investments so that parties are *ex post* locked into bilateral monopoly relationships. A contract when breached might still be renegotiated towards another deal with the same partner if completing the project has a significant joint value.[31] With

30. Although there is no consensus on the most practical notion, renegotiation has not only been a concern for contract theory but also for repeated games (Bergin and MacLeod, 1993) and, as such, this concern has percolated to the analysis of relational contracts. In Levin (2003), an optimal relational contract is renegotiation-proof assuming that renegotiation is only possible before payments. Fong and Surti (2009) and Goldlücke and Kranz (2013) also investigate renegotiation before actions are taken and payment are made.

31. Other circumstances, referred to as *"change orders"* in the parlance of construction contracts, may also require that parties agree to tear up their initial agreement and change it for another deal which might be more responsive to changing conditions if any unexpected contingency arises.

such renegotiation, the court also takes a more active stance, not only enforcing remedies for breaches of the old agreement but also enforcing the new contract.

To model a simple renegotiation protocol, we suppose that parties can write any long-term contract they wish, making full use of the court's penalties L and K when drafting a new agreement. In this scenario, the court has a simple role. It will enforce any long-term contract using penalties up to L and K, but it will allow parties to tear up their initial contract when there is mutual agreement to do so. Parties may also continue their relationship following an earlier breach with a pure relational contract; the court being irrelevant in this case. A last option is for parties to continue to perform under the terms of the original contract.

In his study of pure relational contracting with persistent types and with the objective of demonstrating the non-existence of separating allocations in a context with a continuum of types, Malcomson (2016, p. 324) defines an equilibrium with optimal continuation as a perfect Bayesian equilibrium in pure strategies for which equilibrium-path continuation equilibria following full revelation of the seller's type have payoffs on the Pareto frontier of subgame perfect equilibria of the full information game for that seller type. We adopt this definition in our set-up where parties can write and commit to formal contracts, in particular contracts that induce full separation at date 0. A renegotiation-proof contract with full separation is thus such that the corresponding continuation payoffs always lie on the Pareto frontier of the complete information game from date 1 onwards.

This definition has several implications. First, the seller truthfully reveals his cost at date 0 and any deviation from equilibrium play from date 1 onwards is believed to come from that particular type only. This standard and quite natural beliefs refinement, due to Osborne and Rubinstein (1990, p. 96), is known as the *"never dissuaded once convinced" (NDOC) condition*. The literature on contract renegotiation (Hart and Tirole, 1988; Dewatripont, 1989; Laffont and Tirole, 1993, Chapter 10; Maestri, 2011) suggests that randomization and a gradual revelation of information helps relaxing renegotiation constraints. We leave for further research the analysis of the interaction between limited enforcement and gradual revelation of information that would arise under limited commitment. We just notice that the buyer could also offer a pooling contract for date 0 provided that it is enforceable, induce no information revelation at that date, and let the continuation from date 1 onwards be the one-period-delayed renegotiation-proof contract that we shall now characterize. When δ is small enough, inducing such earlier pooling is certainly not optimal[32] and we will concentrate on that scenario in this subsection.

Secondly, we showed in in Section 4.1 that the optimal enforceable contract under complete information is stationary and can be implemented with a relational contract. The stationary contract with outputs $q_\tau(\theta)=q^{ci}(\theta)=\min\{q^e(\theta),q^{fb}(\theta)\}$ for all $\tau\geq 1$ and payments $t_\tau(\theta)=\theta q^{ci}(\theta)$ thus lies on the Pareto frontier of subgame-perfect equilibria of the complete information continuation game. Observe that the buyer promises nothing from date 1 on with such a renegotiation-proof contract which means that all payments to induce information revelation have to be done at date 0.

To get a simple characterization of the optimal renegotiation-proof contract which is comparable with that obtained in Theorem 1, we still suppose that Assumption 2 holds. This assumption implies that the renegotiation-proof outputs from date 1 are given by $q^{ci}(\underline{\theta})=q^{fb}(\underline{\theta})$ and $q^{ci}(\bar{\theta})\geq q^{bm}(\bar{\theta})$.

Given this continuation, we now want to characterize date 0-outputs at the optimal renegotiation-proof contract. To this end, we first observe that, being given the expression of the outputs implemented in the complete information continuation just described, the seller's

32. See Laffont and Martimort (2001, Chapter 9) for the comparison between fully pooling and fully separating contracts when enforcement is costless.

intertemporal information rents at an optimal contract (which are obtained when the usual incentive and participation constraints are binding[33]) can be expressed as:

$$U_0^+(\underline{\theta}) = \Delta\theta\left((1-\delta)q_0(\overline{\theta}) + \delta q^{ci}(\overline{\theta})\right) \text{ and } U_0^+(\overline{\theta}) = 0.$$

Taking into account the expressions of these rent and output profiles, the date 0-enforcement constraint (3.6) in state $\underline{\theta}$ becomes:

$$\delta\left(S(q^{fb}(\underline{\theta})) - \underline{\theta}q^{fb}(\underline{\theta})\right) - (1-\delta)\underline{\theta}q_0(\underline{\theta}) + (1-\delta)M \geq \Delta\theta((1-\delta)q_0(\overline{\theta}) + \delta q^{ci}(\overline{\theta})). \tag{5.1}$$

The optimal date 0-outputs $\left(q_0^r(\underline{\theta}), q_0^r(\overline{\theta})\right)$ maximize the buyer's expected profit subject to constraint (5.1).

Proposition 4. *Suppose that the Assumption 2 holds. The output profiles* $\mathbf{q}_0(\underline{\theta}) = (q_0^r(\underline{\theta}), q^{fb}(\underline{\theta}), \ldots, q^{fb}(\underline{\theta}), \ldots)$ *and* $\mathbf{q}_0(\overline{\theta}) = (q_0^r(\overline{\theta}), q^{ci}(\overline{\theta}), \ldots, q^{ci}(\overline{\theta}), \ldots)$ *implemented at the optimal renegotiation-proof contract with full separation at date 0 are such that* $q_0^r(\underline{\theta}) < q^{fb}(\underline{\theta})$ *and* $q_0^r(\overline{\theta}) < q^{ci}(\overline{\theta})$ *with:*

$$S'(q_0^r(\underline{\theta})) = \left(1 + \frac{\lambda_r}{\nu}\right)\underline{\theta} \text{ and } S'(q_0^r(\overline{\theta})) = \overline{\theta} + \frac{\nu + \lambda_r}{1-\nu}\Delta\theta, \tag{5.2}$$

where $\lambda_r > 0$ *is the Lagrange multiplier for date 0-enforcement constraint (5.1).*

Thus, the renegotiation-proof output profiles exhibit growing dynamics which bear strong similarities with the commitment case. This shows the robustness of our earlier findings to the possibility of renegotiation. However, some differences remain. Renegotiation-proofness requires that payments and outputs from date 1 onwards are the best ones that can be implemented with relational contracts under complete information. In particular, the seller makes zero profit in all those periods. To induce information revelation, the buyer is thus forced to pay all the low-cost seller's information rent at date 0. This makes the high-cost seller's *"take-the-money-and-run"* strategy particularly attractive at that date. Unfortunately, the buyer can no longer push the cost of date 0-enforcement to future periods as she would do under full commitment. Date 0-enforcement constraint (5.1) can now only be relaxed by reducing payments and outputs at this date. This explains that date 0-outputs are set below their Baron–Myerson levels for both types.

6. APPLICATIONS

Our model is useful to address a number of important questions in organization theory.

6.1. *Asset specificity and contract enforcement*

Our contractual setting can be viewed as a stylized modelling of an ongoing relationship between a contractor and his long-term supplier for an essential input. *Transaction Costs Economics* has already discussed at length how opportunism and asset specificity shape such relationships, especially in terms of their impacts on the optimal degree of vertical integration and more generally

33. The proof is standard and thus omitted.

on contract duration.[34] In our context, and using the language of this paradigm, parties must build safeguards against bilateral opportunism over the growing phase whenever enforcement constraints are binding. Instead, in the long run, more mature relationships are no longer subject to such threat.

Even though specific investments are not present *per se* in our baseline model, our framework can readily be extended to link asset specificity and the quality of enforcement. Making assets more specific to the relationship may thus act as a safeguard against the possibility of breaches. As a simple extension along these lines, suppose that, prior to contracting, the buyer makes a relation-specific investment whose cost is i. This investment enhances surplus in each period τ by an amount $B(i)$ (with $B'(i)>0$ and $B''(i)<0$) so that the net gains from trade in period τ can be written as:

$$S(q_\tau)-\theta q_\tau+B(i)-i.$$

Assuming, for simplicity, an interior solution, the efficient level of investment i^{fb} satisfies:

$$B'(i^{fb})=1.$$

Denote by $C(i)$ the buyer's opportunity cost for the foregone use of dedicated assets if the contract is breached. This cost also increases with asset specificity ($C'(i)>0$ with $C''(i)>0$). This cost can be counted as an implicit penalty for breach that would be borne by the buyer if she does not fulfil her obligations. Keeping the same expression as above for the enforcement surplus $\Psi(\underline{\theta},\overline{\mathbf{q}}_\tau(\underline{\theta}))$, the enforcement constraints become:

$$\Psi(\underline{\theta},\overline{\mathbf{q}}_\tau(\underline{\theta}))+\delta B(i)+(1-\delta)C(i)\geq\delta^{-\tau}\max\left\{\Delta\theta(q_0^+(\overline{\theta})-q_\tau^-(\underline{\theta})),0\right\} \quad \forall\tau. \tag{6.1}$$

Assuming that $C''(i)$ is small enough so that the left-hand side of Constraint (6.1) remains concave in i , the optimal investment level i^e must now take into account the impact of such investment on contracts enforceability. Increasing investment has a first direct effect in relaxing enforcement constraints and secondly, an indirect effect, that comes from changing optimal trade profiles in response to the fact that enforcement constraints are easier to satisfy. By the Envelope Theorem, this indirect effect vanishes and i^e simply solves

$$B'(i^e)=1-\Lambda_\infty\left(\delta B'(i^e)+(1-\delta)C'(i^e)\right)<1.$$

From this, we immediately conclude that $i^e>i^{fb}$. The buyer is now eager to increase her investment as a commitment device to facilitate enforcement. Moreover, those incentives are more pronounced as Λ_∞ is bigger, *i.e.* when enforcement is more difficult.

6.2. *Relational contracting and firm's boundaries*

Gibbons (2005b) and Baker *et al.* (2001) argue that one of the key research questions in the property rights literature is to understand how relational contracts are affected by firm's boundaries. Our article contributes to this important debate. Suppose now that the informed seller may perform some specific investment i_s prior to contracting. That investment improves the value of trade by increasing the probability $\nu(i_s)$ of being efficient. In a vertically integrated firm owned by the buyer, the seller becomes an employee of the firm and, at any point in time,

34. Joskow (1987, 1988), Crocker and Masten (1988, 1996), Ramey and Watson (2001), and Halac (2015).

this employee has the right to leave the firm.[35] Following Riordan (1990), we may also assume that ownership gives access to information. The seller thus enjoys no information rent and has no incentives to make any *ex ante* investment; $i_s=0$.

Under vertical separation instead, the seller owns the assets, retains private information and enjoys an expected informational rent worth $\nu(i_s)q^{sb}(\bar{\theta})$. This rent acts as an engine for investment but it also hardens the enforcement problem. Our model first predicts that *"take-the-money-and-run"* strategy will only arise in market relationships between firms that remain vertically separated. Secondly, market relationships come with greater volumes over time whereas intrafirm exchanges may exhibit more stable patterns. Finally, output distortions being greater when enforcement problems are more acute, vertical integration becomes more attractive when enforcement is more difficult.

APPENDIX

Proof of Lemma 1. Necessity. Observe that

$$U_0^+(\hat{\theta})=U_\tau^-(\hat{\theta})-\delta^\tau(1-\delta)(t_{1,\tau}(\hat{\theta})-L)+\delta^\tau U_\tau^+(\hat{\theta}) \quad \forall\hat{\theta}\in\Theta.$$

Using this condition, we rewrite Constraint (3.3) as:

$$U_0^+(\theta)\geq U_0^+(\hat{\theta})+\delta^\tau(1-\delta)(t_{1,\tau}(\hat{\theta})-L)-\delta^\tau U_\tau^+(\hat{\theta})+(\hat{\theta}-\theta)q_\tau^-(\hat{\theta}) \quad \forall(\theta,\hat{\theta})^2\in\Theta^2, \quad \forall\tau\geq 0.$$

Permuting the roles of θ and $\hat{\theta}$ and manipulating the latter condition yields:

$$\delta^\tau\left(U_\tau^+(\theta)-(1-\delta)(t_{1,\tau}(\theta)-L)\right)\geq U_0^+(\theta)-U_0^+(\hat{\theta})+(\theta-\hat{\theta})q_\tau^-(\theta) \quad \forall(\theta,\hat{\theta})^2\in\Theta^2, \quad \forall\tau\geq 0. \tag{A.1}$$

We can now rewrite equation (3.2) in a more explicit form as:

$$\sum_{s=0}^{\infty}\delta^s(S(q_{\tau+s}(\theta))-t_{\tau+s}(\theta))\geq S(q_\tau(\theta))-t_{1,\tau}(\theta)-K \quad \forall\tau\geq 0.$$

Developing, we get:

$$(1-\delta)\sum_{s=1}^{\infty}\delta^s(S(q_{\tau+s}(\theta))-\theta q_{\tau+s}(\theta))\geq(1-\delta)\theta q_\tau(\theta)+U_\tau^+(\theta)-(1-\delta)(t_{1,\tau}(\theta)+K) \quad \forall\tau\geq 0.$$

Multiplying by δ^τ yields:

$$(1-\delta)\delta^\tau\left(\sum_{s=1}^{\infty}\delta^s(S(q_{\tau+s}(\theta))-\theta q_{\tau+s}(\theta))-\theta q_\tau(\theta)+M\right)\geq\delta^\tau\left(U_\tau^+(\theta)-(1-\delta)(t_{1,\tau}(\theta)-L)\right). \tag{A.2}$$

Taken together, equations (A.1) and (A.2) are compatible if and only if:

$$(1-\delta)\delta^\tau\left(\sum_{s=1}^{\infty}\delta^s(S(q_{\tau+s}(\theta))-\theta q_{\tau+s}(\theta))-\theta q_\tau(\theta)+M\right)\geq\max_{\hat{\theta}\in\Theta}\{U_0^+(\theta)-U_0^+(\hat{\theta})+(\theta-\hat{\theta})q_\tau^-(\theta)\}$$

(where equation (3.4) holds) which can be rewritten as (3.6).

A.1.1. Sufficiency. Suppose that equation (3.6) holds for a quantity profile $\{\mathbf{q}(\theta)\}_{\theta\in\Theta}$. Consider the profile of payments $\{\mathbf{t}(\theta)\}_{\theta\in\Theta}$ (and thus forward rents $U_\tau^+(\theta)$) defined as:

$$\delta^\tau\left(U_\tau^+(\theta)-(1-\delta)(t_{1,\tau}(\theta)-L)\right)=\max_{\hat{\theta}\in\Theta}\{U_0^+(\theta)-U_0^+(\hat{\theta})+(\theta-\hat{\theta})q_\tau^-(\theta)\}. \tag{A.3}$$

By construction, both equations (A.1) and (A.2) hold with those payments. In particular, equation (A.1) is an equality. If we add the requirement

$$t_{1,\tau}(\theta)=L \quad \forall\theta\in\Theta,\forall\tau\geq 0 \tag{A.4}$$

then equation (A.3) fully determines the profile of forward looking rents for the given values of $U_0^+(\overline{\theta})$ and $U_0^+(\underline{\theta})$ that respect equation (3.4).

35. Presumably, $L=0$ in the case of non-alienable employment relationships.

Consider an allocation such that the incentive compatibility of a low-cost seller and the participation constraint of a high-cost one are both binding. These conditions altogether determine the values of $U_0^+(\underline{\theta})$ and $U_0^+(\overline{\theta})$ as:

$$U_0^+(\underline{\theta}) = \Delta\theta q_0^+(\overline{\theta}) \text{ and } U_0^+(\overline{\theta}) = 0. \tag{A.5}$$

From this, we obtain:

$$\max_{\hat{\theta}\in\Theta}\{U_0^+(\underline{\theta}) - U_0^+(\hat{\theta}) + (\underline{\theta} - \hat{\theta})q_\tau^-(\underline{\theta})\} = \max\left\{\Delta\theta(q_0^+(\overline{\theta}) - q_\tau^-(\underline{\theta})), 0\right\} \tag{A.6}$$

and

$$\max_{\hat{\theta}\in\Theta}\{U_0^+(\overline{\theta}) - U_0^+(\hat{\theta}) + (\overline{\theta} - \hat{\theta})q_\tau^-(\overline{\theta})\} = \max\left\{\Delta\theta(-q_0^+(\underline{\theta}) + q_\tau^-(\overline{\theta})), 0\right\} = 0, \tag{A.7}$$

where the last inequality follows from $q_\tau^-(\overline{\theta}) \leq q_0^+(\overline{\theta}) \leq q_0^+(\underline{\theta})$ (since equation (3.5) necessarily holds from incentive compatibility).

From equation(A.3) taken for $\theta = \overline{\theta}$, (A.4) and (A.6), we deduce that:

$$U_\tau^+(\overline{\theta}) = 0 \quad \forall\tau \tag{A.8}$$

and thus

$$t_\tau(\overline{\theta}) - \overline{\theta}q_\tau(\overline{\theta}) = 0 \quad \forall\tau. \tag{A.9}$$

From equation (A.3) taken for $\theta = \underline{\theta}$, (A.4) and (A.6), we also deduce that:

$$U_\tau^+(\underline{\theta}) = \delta^{-\tau}\max\left\{\Delta\theta(q_0^+(\overline{\theta}) - q_\tau^-(\underline{\theta})), 0\right\} \quad \forall\tau. \tag{A.10}$$

||

For future references, we may rewrite Lemma 1 by developing the enforcement constraints (3.6) as:

Lemma A.1. *An incentive compatible mechanism $\mathcal{C}$ is* ***enforceable*** *if and only if the following enforcement constraints hold at all dates $\tau \geq 0$:*

$$\Psi(\underline{\theta}, \mathbf{q}_\tau^+(\underline{\theta})) \geq \delta^{-\tau}\max\left\{\Delta\theta(q_0^+(\overline{\theta}) - q_\tau^-(\underline{\theta})), 0\right\}, \tag{A.11}$$

$$\Psi(\overline{\theta}, \mathbf{q}_\tau^+(\overline{\theta})) \geq \delta^{-\tau}\max\left\{\Delta\theta(q_\tau^-(\overline{\theta}) - q_0^+(\underline{\theta})), 0\right\}, \tag{A.12}$$

$$\min\left\{\delta^\tau\Psi(\underline{\theta}, \mathbf{q}_\tau^+(\underline{\theta})) + \Delta\theta q_\tau^-(\underline{\theta}); \Delta\theta q_0^+(\underline{\theta})\right\} \geq \max\left\{-\delta^\tau\Psi(\overline{\theta}, \mathbf{q}_\tau^+(\overline{\theta})) + \Delta\theta q_\tau^-(\overline{\theta}); \Delta\theta q_0^+(\overline{\theta})\right\}. \tag{A.13}$$

Proof of Lemma A.1. The incentive compatibility conditions (3.4) imply:

$$\Delta\theta q_0^+(\underline{\theta}) \geq U_0^+(\underline{\theta}) - U_0^+(\overline{\theta}) \geq \Delta\theta q_0^+(\overline{\theta}). \tag{A.14}$$

Inserting the second (resp. first) of these inequalities into equation (3.6) taken for $\theta = \underline{\theta}$ (resp. taken for $\theta = \overline{\theta}$) yields (A.11) (resp. (A.12)).

There exist values of $U_0^+(\underline{\theta}) - U_0^+(\overline{\theta})$ that satisfy equations (A.14) and (3.6) if and only if the following condition holds:

$$\min\left\{\delta^\tau\Psi(\underline{\theta}, \mathbf{q}_\tau^+(\underline{\theta})) + \Delta\theta q_\tau^-(\underline{\theta}); \Delta\theta q_0^+(\underline{\theta})\right\} \geq U_0^+(\underline{\theta}) - U_0^+(\overline{\theta}) \geq \max\left\{-\delta^\tau\Psi(\overline{\theta}, \mathbf{q}_\tau^+(\overline{\theta})) + \Delta\theta q_\tau^-(\overline{\theta}); \Delta\theta q_0^+(\overline{\theta})\right\}. \tag{A.15}$$

This finally gives us condition (A.13). Observe that the participation constraints (3.1) can be satisfied for both types by conveniently choosing non-negative values for $U_0^+(\underline{\theta})$ and $U_0^+(\overline{\theta})$ while keeping $U_0^+(\underline{\theta}) - U_0^+(\overline{\theta})$ that satisfies equation (A.15). ||

Proof of Proposition 1. Consider problem $(\mathcal{P})$ written with the enforcement constraints (A.11), (A.12), and (A.13). We first neglect these constraints and consider the participation constraint (3.1) for type $\overline{\theta}$ and the incentive constraint (3.4) for type $\underline{\theta}$. Notice that the enforcement constraints no longer contain $U_0^+(\underline{\theta})$ and $U_0^+(\overline{\theta})$ thanks to Lemma A.1. Thus, equations (3.1) and (3.4) are both binding at the optimum of the so relaxed problem. The corresponding optimal outputs are stationary and, respectively, given by $q^{fb}(\underline{\theta})$ and $q^{bm}(\overline{\theta}) < q^{fb}(\underline{\theta})$. It is routine to check the remaining participation and incentive constraints. Turning now to equation (A.11) written with those stationary outputs, it becomes:

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) \geq \delta^{-\tau}\max\left\{\Delta\theta(q^{bm}(\overline{\theta}) - (1-\delta^\tau)q^{fb}(\underline{\theta})); 0\right\}. \tag{A.16}$$

Because $q^{bm}(\overline{\theta}) < q^{fb}(\underline{\theta})$, the right hand side of equation (A.16) is maximum at $\tau = 0$. Manipulating leads to the first inequality in Assumption 1.

Observe that (A.12) now becomes:

$$\psi(\overline{\theta}, q^{bm}(\overline{\theta})) \geq \delta^{-\tau} \max\left\{\Delta\theta((1-\delta^\tau)q^{bm}(\overline{\theta}) - q^{fb}(\underline{\theta})); 0\right\}. \tag{A.17}$$

Since $(1-\delta^\tau)q^{bm}(\overline{\theta}) < q^{bm}(\overline{\theta}) < q^{fb}(\underline{\theta})$, the right hand side above is 0 giving us the second inequality in Assumption 1.

Turning now to (A.13), this condition becomes:

$$\min\left\{\delta^\tau \psi(\underline{\theta}, q^{fb}(\underline{\theta})) + \Delta\theta(1-\delta^\tau)q^{fb}(\underline{\theta}); \Delta\theta q^{fb}(\underline{\theta})\right\} \geq \max\left\{-\delta^\tau \psi(\overline{\theta}, q^{bm}(\overline{\theta})) + \Delta\theta(1-\delta^\tau)q^{bm}(\overline{\theta}); \Delta\theta q^{bm}(\overline{\theta})\right\}.$$

The latter condition immediately follows from (A.16), (A.17) and $q^{fb}(\underline{\theta}) \geq q^{bm}(\overline{\theta})$. ||

Proof of Theorem 1. Preliminaries. Denote by l_∞ the Banach space of all bounded sequences $\mathbf{x}$ such that $\|\mathbf{x}\|_\infty \equiv \sup|x_\tau| < \infty$.

Given that $\mathcal{Q}$ is a bounded interval, the set $\mathbf{Q}_\infty$ of all non-negative output sequences $\mathbf{x} = \{x_\tau\}_{\tau=0}^\infty$ on $\mathcal{Q}$ is a subset of l_∞. Let also l_1 denote the space of all sequences $\mathbf{x}$ such that $\|\mathbf{x}\| \equiv \sum_{\tau=0}^\infty |x_\tau| < \infty$. The dual space of l_∞ is $l_\infty^* = l_1 \bigoplus l_s$, where l_s is the set of bounded linear functional generated by the purely additive measures on the integers. (Theorem A.1 below shows that the sequence of Lagrange multipliers that characterizes optimal contracts belongs in fact to l_1.) We denote by $\mathcal{A}$ the closed and convex subset of $\mathbf{Q}_\infty \times \mathbf{Q}_\infty \subseteq l_\infty \times l_\infty$ such that the monotonicity condition (3.5) holds.

A.1.2. Simplifying the objective function. We first consider a relaxed problem $(\mathcal{P})$ with, on top of the enforcement constraints, only the incentive compatibility constraint (3.4) of type $\underline{\theta}$ and the participation constraint (3.1) of type $\theta = \overline{\theta}$ which are binding at the optimum. We are thus neglecting the incentive compatibility constraint (3.4) for type $\overline{\theta}$ and the participation constraint (3.1) for type $\theta = \underline{\theta}$. (Notice again that the enforcement constraints do not contain $U_0^+(\theta)$ and thus $U_0^+(\theta)$ can be decreased without affecting these constraints.) Constraint (3.4) for type $\overline{\theta}$ holds when (3.4) for type $\theta = \underline{\theta}$ is binding and the allocation satisfies equation (3.5). Constraint (3.1) holds for type $\theta = \underline{\theta}$ if equation (3.1) for $\theta = \overline{\theta}$ and equation (3.4) for type $\theta = \underline{\theta}$ are both binding since output is non-negative.

Secondly, we neglect equations (A.12) and (A.13) which are both checked *ex post*. Inserting $\overline{U}_0(\underline{\theta}) = \Delta\theta q_0^+(\overline{\theta})$ and $\overline{U}_0(\overline{\theta}) = 0$ into the maximand simplifies the objective function that becomes:

$$f(\mathbf{q}) = E_\theta\left((1-\delta)\sum_{\tau=0}^\infty \delta^\tau (S(q_\tau(\theta)) - m(\theta)q_\tau(\theta))\right) \text{ where } m(\theta) = \begin{cases} \underline{\theta} & \text{if } \theta = \underline{\theta}, \\ \overline{\theta} + \frac{\nu}{1-\nu}\Delta\theta & \text{if } \theta = \overline{\theta} \end{cases}.$$

The function $f(\mathbf{q})$ maps $\mathcal{A}$ into $\mathbb{R}$ and is strictly concave. It thus admits a (single-valued) superdifferential $\partial f(\mathbf{q})$ given by:

$$\partial f(\mathbf{q}) = (1-\delta)\delta^\tau \left\{\left(\nu\left(S'(q_\tau(\underline{\theta})) - \underline{\theta}\right), (1-\nu)\left(S'(q_\tau(\overline{\theta})) - \overline{\theta}\right) - \frac{\nu}{1-\nu}\Delta\theta\right)\right\}_{\tau \geq 0}.$$

It can be easily checked that $\partial f(\mathbf{q})$ belongs to $l_1 \times l_1$.

A.1.3. Constrained set. We rewrite equation (3.9) as:

$$g_\tau(\mathbf{q}) = \delta^\tau \Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta})) - \max\left\{\Delta\theta(q_0^+(\overline{\theta}) - q_\tau^-(\underline{\theta})), 0\right\} \geq 0. \tag{A.18}$$

The function $g_\tau(\mathbf{q})$ maps $\mathcal{A}$ into $\mathbb{R}$ and is strictly concave in $\mathbf{q}$ for all $\tau \geq 0$. It thus admits a (single-valued) superdifferential $\partial g_\tau(\mathbf{q})$. Let also denote $g(\mathbf{q}) = \{g_\tau(\mathbf{q})\}_{\tau \geq 0}$.

A.1.4. Formulation. We rewrite the maximization problem as:

$$(\mathcal{P}): \quad \max_{\mathbf{q} \in \mathcal{A}} f(\mathbf{q}) \text{ subject to } g(\mathbf{q}) \geq 0.$$

$(\mathcal{P})$ is an optimization problems with infinitely many constraints, a feature that requires careful use of duality arguments. The corresponding Lagrangian can be written as:

$$\mathcal{L}(\mathbf{q}, \lambda) = f(\mathbf{q}) + \lambda g(\mathbf{q})$$

$$= E_\theta\left((1-\delta)\left(\sum_{\tau=0}^\infty \delta^\tau (S(q_\tau(\theta)) - m(\theta)q_\tau(\theta))\right)\right) + \sum_{\tau=0}^\infty \lambda_\tau \left(\delta^\tau \Psi(\underline{\theta}, \mathbf{q}_\tau^+(\underline{\theta})) - \max\left\{\Delta\theta(q_0^+(\overline{\theta}) - q_\tau^-(\underline{\theta})), 0\right\}\right).$$

Next theorem reminds an important result due to Dechert (1982) that ensures the existence of a sequence of non-negative Lagrange multipliers $\lambda = \{\lambda_\tau\}_{\tau=0}^\infty \in l_1$ for this problem.

Theorem A.1. *(Dechert, 1982). Suppose f and g are concave and Fréchet differentiable with $\partial f(\mathbf{q})\in l_1\times l_1$. Let $\mathbf{q}^*$ be a solution to $(\mathcal{P})$. Suppose that the following conditions hold.*

(1) There exists $\tilde{\mathbf{q}}\in\mathcal{A}$ such that (Slater condition):

$$\sup_{\tau} g_\tau(\tilde{\mathbf{q}})>0. \tag{A.19}$$

(2) g is asymptotically insensitive if $\forall\mathbf{x}\in\mathcal{A}$, and $\mathbf{y}$ such that $y_\tau\neq 0$ for finitely many τ and $\mathbf{x}+\mathbf{y}\in\mathcal{A}$:

$$\textit{(AI)}\ \lim_{\tau\to+\infty} g_\tau(\mathbf{x}+\mathbf{y})-g_\tau(\mathbf{x})=0. \tag{A.20}$$

(3) g is asymptotically non-anticipatory if $\forall\mathbf{x}\in\mathcal{A}$, and $\mathbf{y}$ such that $\mathbf{x}+\mathbf{y}\in\mathcal{A}$ and $y^T_\tau=\begin{cases}0 & \text{if } \tau\leq T\\ y_\tau & \text{if } \tau>T\end{cases}$:

$$\textit{(ANA)}\ \lim_{T\to+\infty} g_\tau(\mathbf{x}+\mathbf{y}^T)=g_\tau(\mathbf{x}),\quad \forall\tau\geq 0. \tag{A.21}$$

Then there exists $\lambda\in l_1$ such that:

$$\lambda g(\mathbf{q}^*)=0,^{36} \tag{A.22}$$

$$\mathcal{L}(\mathbf{q}^*,\lambda)\geq\mathcal{L}(\mathbf{q},\lambda)\quad\forall\mathbf{q}\in\mathcal{A}. \tag{A.23}$$

We already noticed that $\partial f(\mathbf{q})\in l_1\times l_1$. It remains to check that conditions (A.19), (A.20) and (A.21) hold. First, the Slater condition (A.19) is satisfied when $M>0$ by $\tilde{\mathbf{q}}\equiv 0$. Secondly, by a remark in the text, for τ large enough, $\max\left\{\Delta\theta(q_0^+(\overline{\theta})-q_\tau^-(\underline{\theta})),0\right\}=0$ for any $\mathbf{q}\in\mathcal{A}$. Thus, we get:

$$|g_\tau(\mathbf{x}+\mathbf{y})-g_\tau(\mathbf{x})|=\delta^\tau\,|\Psi(\underline{\theta},\mathbf{q}_\tau^+(\underline{\theta})+\mathbf{y}_\tau^+(\underline{\theta}))-\Psi(\underline{\theta},\mathbf{q}_\tau^+(\underline{\theta}))|\leq K\delta^\tau\to_{\tau\to+\infty}0 \tag{A.24}$$

for some $K>0$ since $\Psi(\underline{\theta},\mathbf{q}^+)$ is continuous and $\mathcal{Q}$ is bounded. Henceforth, condition (A.21) holds. Similarly, condition (A.21) trivially holds.

A.1.5. Optimization. First, we rewrite the optimality condition by means of superdifferentials, assuming that the optimal output profile is in the interior of $\mathcal{A}$ (*i.e.* the monotonicity condition (3.5) is strict). This gives us:

$$0\in\partial f(\mathbf{q}^*)+\lambda\partial g(\mathbf{q}^*). \tag{A.25}$$

We can now explore the implications of the optimality conditions (A.25).

(1) Optimality w.r.t. $q_\tau(\underline{\theta})$:

$$S'(q_\tau(\underline{\theta}))-\underline{\theta}=\frac{\lambda_\tau\underline{\theta}-\left(\sum_{s=\tau+1}^{\infty}\lambda_s 1_s\right)\Delta\theta}{\nu+\sum_{s=0}^{\tau-1}\lambda_s}, \tag{A.26}$$

where $1_s=\begin{cases}1 & \text{if } \Delta\theta(q_0^+(\overline{\theta})-q_s^-(\underline{\theta}))>0\\ \in[0,1] & \text{if } \Delta\theta(q_0^+(\overline{\theta})-q_s^-(\underline{\theta}))=0\\ 0 & \text{if } \Delta\theta(q_0^+(\overline{\theta})-q_s^-(\underline{\theta}))<0.\end{cases}$

(2) Optimality w.r.t. $q_\tau(\overline{\theta})$:

$$S'(q_\tau(\overline{\theta}))-\overline{\theta}=\frac{\nu+\sum_{s=0}^{\infty}\lambda_s 1_s}{1-\nu}\Delta\theta, \tag{A.27}$$

where the assumption $S'(0)$ sufficiently large ensures that $q_\tau(\overline{\theta})$ remains positive.

36. We also use the convention that the product equality $\mathbf{xy}=0$ should be understood coordinate wise as $x_\tau y_\tau=0$ for all $\tau\geq 0$.

A.1.6. Output distortions. From (A.27), we necessarily have $q_\tau(\overline{\theta}) \leq q^{bm}(\overline{\theta})$ and the inequality is strict provided one multiplier at least is positive, a fact which is known to be true when Assumption 2 holds since this assumption means that the Baron–Myerson allocation (obtained when all multipliers are zero) is no longer implementable.

Turning now to the low-cost seller's output, the optimal output of the low-cost seller is first-best far enough in the future. To show that, we first prove a first lemma.

Lemma A.2. *Suppose that Assumption 2 holds. There exists* $\tau^* \geq 0$ *such that:*

$$\lambda_\tau = 0 \quad \forall \tau \geq \tau^*, \tag{A.28}$$

and

$$S'(q_\tau(\underline{\theta})) = \underline{\theta} \quad \forall \tau > \tau^*. \tag{A.29}$$

Proof of Lemma A.2. Because $\lambda_\tau \geq 0$ and $\lambda \in l_1$ (*i.e.* $\sum_{s=0}^{\infty} \lambda_s < +\infty$), we have $\lim_{\tau \to +\infty} \lambda_\tau = 0$ and $\sum_{s=0}^{\infty} \lambda_s 1_s < +\infty$. Inserting into equation (A.26), yields:

$$\lim_{\tau \to +\infty} S'(q_\tau(\underline{\theta})) = \underline{\theta}. \tag{A.30}$$

The first-best output for a low-cost seller is always implemented in the limit. For τ large enough, an allocation in the interior of $\mathcal{C}$ is strictly monotonic and thus equation (3.9) writes as in equation (3.7). But passing to the limit and using equations (A.30), (3.7) becomes:

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) \geq 0.$$

By Assumption 2, this latter inequality is actually strict and thus equation (3.9) cannot be binding for τ large enough so that (A.28) holds. ||

A.1.7. Binding enforcement constraints. Because $(\mathcal{P})$ is a concave problem, the necessary conditions for optimality (36) and (A.23) are also sufficient. From Lemma A.2, the solution is such that equation (A.18) is binding at all dates $\tau \leq \tau^*$. In that case, we conjecture that $\Delta\theta(q_0^+(\overline{\theta}) - q_\tau^-(\underline{\theta})) > 0$ for all such dates.

Let us now define the sequence Λ of cumulative multipliers as:

$$\Lambda_\tau = \sum_{s=0}^{\tau-1} \lambda_s$$

with the convention $\Lambda_0 = 0$. Because all multipliers λ_s are non-negative, Λ is a non-decreasing and non-negative sequence with terminal value $\Lambda_{\tau^*+1} = \Lambda_\infty$. From the optimality condition (A.26) and given our conjecture, the sequence Λ satisfies the recursive equation:

$$(S'(q_\tau(\underline{\theta})) - \underline{\theta})(\nu + \Lambda_\tau) = (\Lambda_{\tau+1} - \Lambda_\tau)\underline{\theta} - (\Lambda_\infty - \Lambda_{\tau+1})\Delta\theta.$$

After manipulations, we get:

$$\overline{\theta}\Lambda_{\tau+1} - \Delta\theta\Lambda_\infty = \nu(S'(q_\tau(\underline{\theta})) - \underline{\theta}) + S'(q_\tau)\Lambda_\tau. \tag{A.31}$$

Observe also that (A.26) implies

$$S'(q_{\tau^*}(\underline{\theta})) = \underline{\theta} + \frac{\underline{\theta}\lambda_{\tau^*}}{\nu + \Lambda_\infty - \lambda_{\tau^*}} \geq \underline{\theta} \Rightarrow q_{\tau^*}(\underline{\theta}) \leq q^{fb}(\underline{\theta}). \tag{A.32}$$

Consider thus a non-decreasing sequence $q_\tau(\underline{\theta})$ (and strictly so for $\tau \leq \tau^*$) such that $q_{\tau^*}(\underline{\theta}) \leq q^{fb}(\underline{\theta})$. (An argument below will show that the optimal outputs satisfy this monotonicity property.) We can rewrite equation (A.31) as:

$$\Lambda_{\tau+1} = \alpha_\tau \Lambda_\tau + \beta_\tau \tag{A.33}$$

with

$$\alpha_\tau = \frac{S'(q_\tau(\underline{\theta}))}{\overline{\theta}} \text{ and } \beta_\tau = \frac{\Delta\theta\Lambda_\infty + \nu(S'(q_\tau(\underline{\theta})) - \underline{\theta})}{\overline{\theta}}. \tag{A.34}$$

Lemma A.3. *Suppose that Assumption 2 holds. There exists* $\tau^* \geq 0$ *such that:*

$$\lambda_\tau > 0 \quad \forall \tau \leq \tau^* \text{ and } \lambda_\tau = 0 \quad \forall \tau > \tau^*. \tag{A.35}$$

Proof of Lemma A.3. From Lemma A.2, we know that there exists a maximal date $\tau^* \geq 0$ such that the multiplier λ_τ is positive only for $\tau \leq \tau^*$. We want to show that indeed $\lambda_\tau > 0$ for all $\tau \leq \tau^*$. To this end, observe that the sequence Λ is increasing at all $\tau \leq \tau^* - 1$ (so that all corresponding multipliers λ_τ remain positive). We have:

$$\Lambda_{\tau+1} > \Lambda_\tau \Leftrightarrow \Lambda_\tau < \frac{\beta_\tau}{1-\alpha_\tau} = \frac{\Delta\theta\Lambda_\infty + \nu(S'(q_\tau(\underline{\theta})) - \underline{\theta})}{\overline{\theta} - S'(q_\tau(\underline{\theta}))}. \tag{A.36}$$

Since the right hand side of (A.36) is decreasing in $q_\tau(\underline{\theta})$, the sequence $\frac{\beta_\tau}{1-\alpha_\tau}$ is itself decreasing. Moreover, the following string of conditions holds:

$$\Lambda_{\tau^*} = \Lambda_\infty - \lambda_{\tau^*} \leq \Lambda_\infty \leq \frac{\beta_{\tau^*}}{1-\alpha_{\tau^*}} = \frac{\Delta\theta\Lambda_\infty + \nu(S'(q_{\tau^*}(\underline{\theta})) - \underline{\theta})}{\overline{\theta} - S'(q_{\tau^*}(\underline{\theta}))},$$

where the first inequality follows from $\lambda_{\tau^*} \geq 0$ and the last one from equation (A.32). Now, we can write:

$$\Lambda_{\tau^*-1} = \frac{\Lambda_{\tau^*} - \beta_{\tau^*-1}}{\alpha_{\tau^*-1}} \leq \frac{\frac{\beta_{\tau^*}}{1-\alpha_{\tau^*}} - \beta_{\tau^*-1}}{\alpha_{\tau^*-1}} < \frac{\frac{\beta_{\tau^*-1}}{1-\alpha_{\tau^*-1}} - \beta_{\tau^*-1}}{\alpha_{\tau^*-1}} = \frac{\beta_{\tau^*-1}}{1-\alpha_{\tau^*-1}},$$

where the last right hand side inequality uses the fact that the sequence $\frac{\beta_\tau}{1-\alpha_\tau}$ is decreasing. Proceeding recursively, we obtain:

$$\Lambda_\tau < \frac{\beta_\tau}{1-\alpha_\tau} \quad \forall \tau \leq \tau^* - 1.$$

Hence, $\lambda_\tau = \Lambda_{\tau+1} - \Lambda_\tau > 0$ for all $\tau \leq \tau^*$ and $\lambda_\tau = \Lambda_{\tau+1} - \Lambda_\tau = 0$ for $\tau > \tau^*$. Henceforth, when equation (A.18) is binding at date τ^*, it is also so at all dates $\tau < \tau^*$. ‖

We now set up the stage for a sharp characterization of contractual dynamics. To this end, consider any sequence $\mathbf{q}$, starting with an arbitrary output level $q_0 \in [q^e(\overline{\theta}), q^{fb}(\underline{\theta})]$ and constructed recursively as:

$$\begin{cases} q_0 \in [q^e(\overline{\theta}), q^{fb}(\underline{\theta})], \\ q_{\tau+1} = \Phi(q_\tau). \end{cases} \tag{A.37}$$

The function $\Phi(q) = S^{-1}\left(\frac{1}{\delta}\left(\overline{\theta}q - (1-\delta)M\right)\right)$ is defined over the interval $\left[\frac{(1-\delta)M}{\overline{\theta}}, +\infty\right)$. It is increasing, convex and has a unique fixed point $q^e(\overline{\theta})$. For future references, we also note that the inverse function $\Gamma(q) = \Phi^{-1}(q) = \frac{1}{\overline{\theta}}(\delta S(q) + (1-\delta)M)$ is increasing and concave.

Equipped with the characterization of such sequences, we now explore the consequences of Lemma A.3 for the optimal outputs produced by a low-cost seller. When $\tau^* \geq 2$, we may indeed rewrite equation (A.18) when binding at two subsequent dates τ and $\tau+1$ for all τ such that $\tau + 1 \leq \tau^*$ respectively, as:

$$\Psi(\underline{\theta}, \mathbf{q}^+_\tau(\underline{\theta})) = \delta^{-\tau}\Delta\theta(q_0^+(\overline{\theta}) - q_\tau^-(\underline{\theta})), \tag{A.38}$$

$$\delta\Psi(\underline{\theta}, \mathbf{q}^+_{\tau+1}(\underline{\theta})) = \delta^{-\tau}\Delta\theta(q_0^+(\overline{\theta}) - q_{\tau+1}^-(\underline{\theta})). \tag{A.39}$$

By subtracting equation (A.38) from (A.39), we obtain:

$$\delta S(q_{\tau+1}(\underline{\theta})) - \underline{\theta}q_\tau(\underline{\theta}) + (1-\delta)M = \Delta\theta q_\tau(\underline{\theta}).$$

Simplifying, the sequence $\mathbf{q}(\underline{\theta}) = \{q_\tau(\underline{\theta})\}_{\tau\geq 0}$ satisfies the recursive condition (A.37) for all $\tau \geq 0$ such that $\tau \leq \tau^* - 1$. Starting thus from $q_{\tau^*}(\underline{\theta}) \in [q^e(\overline{\theta}), q^{fb}(\underline{\theta}))$, we may then construct the following (backward) recursive sequence of outputs $\Gamma(q_{\tau^*}(\underline{\theta})) = q_{\tau^*-1}(\underline{\theta})$, and thus $\Gamma^s(q_{\tau^*}(\underline{\theta})) = q_{\tau^*-s}(\underline{\theta})$ (or $\Gamma^{\tau^*-s}(q_{\tau^*}(\underline{\theta})) = q_s(\underline{\theta})$) for $s \leq \tau^*$ where Γ^k denotes the k-th iteration of the mapping Γ). By construction, $q_\tau(\underline{\theta})$ for all $\tau \leq \tau^*$ is increasing in τ for all $\tau \leq \tau^*$. Moreover, that $q^e(\overline{\theta}) < q^{fb}(\underline{\theta})$ (which is implied by Assumption 2 since $\psi(\underline{\theta}, q^{fb}(\underline{\theta})) < \Delta\theta q^{bm}(\overline{\theta}) < \Delta\theta q^{fb}(\underline{\theta})$) also implies $q^e(\overline{\theta}) = \Gamma(q^e(\overline{\theta})) \leq \Gamma(q^{fb}(\underline{\theta})) < q^{fb}(\underline{\theta})$. Therefore, we get:

$$q^e(\overline{\theta}) \leq q_\tau(\underline{\theta}) \quad \forall \tau \leq \tau^* - 1 \tag{A.40}$$

and thus equation (4.3) holds.

A.1.8. Checking the omitted constraints. It is routine to check that equation (3.4) for $\overline{\theta}$ and equation (3.1) for $\underline{\theta}$ are both satisfied.

We now check that the remaining enforcement constraints hold. First, equation (A.12) amounts to

$$\delta^\tau \psi(\overline{\theta}, q^{sb}(\overline{\theta})) \geq \max\left\{\Delta\theta((1-\delta^\tau)q^{sb}(\overline{\theta}) - q_0^{+sb}(\underline{\theta})), 0\right\}. \tag{A.41}$$

Observe that $(1-\delta^\tau)q^{sb}(\overline{\theta}) \leq q^{sb}(\overline{\theta}) < q^{bm}(\overline{\theta}) < q^e(\overline{\theta}) < q_\tau^{sb}(\underline{\theta})$ for all $\tau \geq 0$. Thus, $(1-\delta^\tau)q^{sb}(\overline{\theta}) < q_0^{+sb}(\underline{\theta})$ and (A.41) is implied by $\psi(\overline{\theta}, q^{sb}(\overline{\theta})) > \psi(\overline{\theta}, q^{bm}(\overline{\theta})) \geq 0$ where the last inequality follows from Assumption 2 and $q^{sb}(\overline{\theta}) < q^{bm}(\overline{\theta})$.

Secondly, equation (A.13) now amounts to

$$\min\left\{\delta^\tau \Psi(\underline{\theta},\mathbf{q}_\tau^{+sb}(\underline{\theta}))+\Delta\theta q_\tau^-(\underline{\theta});\,\Delta\theta q_0^{+sb}(\underline{\theta})\right\}\geq \max\left\{-\delta^\tau\psi(\overline{\theta},q^{sb}(\overline{\theta}))+\Delta\theta(1-\delta^\tau)q^{sb}(\overline{\theta});\,\Delta\theta q^{sb}(\overline{\theta})\right\}. \tag{A.42}$$

From equation (A.18) being satisfied at all date τ, the right hand side above can be bounded below by $\min\left\{\Delta\theta q^{sb}(\overline{\theta});\,\Delta\theta q_0^{+sb}(\underline{\theta})\right\}=\Delta\theta q^{sb}(\overline{\theta})$. Because $q^{sb}(\overline{\theta})<q^{bm}(\overline{\theta})<q^{fb}(\underline{\theta})<q^e(\underline{\theta})$ (where the last inequality follows from Assumption 2), $\psi(\overline{\theta},q^{sb}(\overline{\theta}))+\Delta\theta q^{sb}(\overline{\theta})=\psi(\underline{\theta},q^{sb}(\overline{\theta}))>0$ and the right hand side of (A.42) amounts to $\Delta\theta q^{sb}(\overline{\theta})$ which is thus lower than the left hand side found above. ‖

Proof of Proposition 2. Condition (A.4) amounts to (4.8). Condition (4.9) follows from the fact that the high-cost seller's output is constant over time and from (A.9). To prove equation (4.10) we need to show first that $t_\tau^{sb}(\underline{\theta})=\overline{\theta}q_\tau^{sb}(\underline{\theta})$ holds over a transitory phase. Then, Lemma A.4 below shows that this phase coincides with the transitory phase described in Theorem 1. Note that $q_0^+(\overline{\theta})=q^{sb}(\overline{\theta})$ and that $q_0^+(\overline{\theta})-q_\tau^-(\underline{\theta})=q^{sb}(\overline{\theta})-q_\tau^-(\underline{\theta})$ is decreasing in τ (with by definition $q^{sb}(\overline{\theta})-q_0^-(\underline{\theta})=q^{sb}(\overline{\theta})>0$). We have $q_{\tau^*+1}(\underline{\theta})=q_{\tau^*+2}(\underline{\theta})=\ldots=q^{fb}(\underline{\theta})$. Now observe that equation (4.3) implies that

$$q_\tau^-(\underline{\theta})=(1-\delta)\sum_{s=0}^{\tau-1}\delta^\tau q_s(\underline{\theta})>(1-\delta)\,\frac{1-\delta^\tau}{1-\delta}q^e(\overline{\theta})=\left(1-\delta^\tau\right)q^e(\overline{\theta}).$$

Since $q^e(\overline{\theta})>q^{bm}(\overline{\theta})$ (from the second condition in Assumption 2) and $q^{bm}(\overline{\theta})\geq q^{sb}(\overline{\theta})$ (from (4.5)), we have $q^e(\overline{\theta})>q^{sb}(\overline{\theta})$. Moreover, $q_\tau^-(\underline{\theta})$ is an increasing sequence, bounded above by $q^{fb}(\underline{\theta})$ (from (4.3)); so it converges towards a finite limit $q_\infty^-(\underline{\theta})$ such that $q_\infty^-(\underline{\theta})\geq q^e(\overline{\theta})>q^{sb}(\overline{\theta})$. Finally, we conclude on the existence of a date τ' such that:

$$q^{sb}(\overline{\theta})-q_\tau^-(\underline{\theta})\geq 0\quad \forall\tau\leq\tau' \text{ and } q^{sb}(\overline{\theta})-q_\tau^-(\underline{\theta})<0\quad\forall\tau>\tau'. \tag{A.43}$$

Inserting into equation (A.10), we obtain:

$$U_\tau^+(\underline{\theta})=\begin{cases}\delta^{-\tau}\Delta\theta\left(q^{sb}(\overline{\theta})-q_\tau^-(\underline{\theta})\right) & \forall\tau\leq\tau',\\ 0 & \text{otherwise.}\end{cases} \tag{A.44}$$

(1) Consider a date τ such that $\tau+1\leq\tau'$. For such τ, we have

$$U_\tau^+(\underline{\theta})=(1-\delta)\left(t_\tau(\underline{\theta})-\underline{\theta}q_\tau(\underline{\theta})\right)+\delta U_{\tau+1}^+(\underline{\theta}).$$

Or, using equation (A.44),

$$U_\tau^+(\underline{\theta})=(1-\delta)\left(t_\tau(\underline{\theta})-\underline{\theta}q_\tau(\underline{\theta})\right)+\delta\delta^{-\tau-1}\Delta\theta\left(q^{sb}(\overline{\theta})-q_{\tau+1}^-(\underline{\theta})\right).$$

Using again equation (A.44) to express the left hand side yields:

$$\delta^{-\tau}\Delta\theta(q^{sb}(\overline{\theta})-q_\tau^-(\underline{\theta}))=(1-\delta)\left(t_\tau(\underline{\theta})-\underline{\theta}q_\tau(\underline{\theta})\right)+\delta^{-\tau}\Delta\theta\left(q^{sb}(\overline{\theta})-q_{\tau+1}^-(\underline{\theta})\right).$$

Simplifying, we can write:

$$0=(1-\delta)\left(t_\tau(\underline{\theta})-\underline{\theta}q_\tau(\underline{\theta})\right)-\delta^{-\tau}\Delta\theta(1-\delta)\delta^\tau q_\tau(\underline{\theta})=(1-\delta)\left(t_\tau(\underline{\theta})-\overline{\theta}q_\tau(\underline{\theta})\right).$$

Henceforth, equation (4.10) holds for all τ such that $\tau+1\leq\tau'$.

(2) At date τ', we use equation (4.10) to rewrite (equation A.44) as:

$$U_{\tau'}^+(\underline{\theta})=(1-\delta)\left(t_{\tau'}(\underline{\theta})-\underline{\theta}q_{\tau'}(\underline{\theta})\right)=\delta^{-\tau'}\Delta\theta\left(q^{sb}(\overline{\theta})-q_{\tau'}^-(\underline{\theta})\right)$$

or, equivalently,

$$t_{\tau'}(\underline{\theta})=\underline{\theta}q_{\tau'}(\underline{\theta})+\frac{\delta^{-\tau'}}{1-\delta}\Delta\theta\left(q^{sb}(\overline{\theta})-q_{\tau'}^-(\underline{\theta})\right).$$

(3) Finally, consider $U_\tau^+(\underline{\theta})$ and $U_{\tau+1}^+(\underline{\theta})$ for $\forall\tau>\tau'$. We have

$$0=U_\tau^+(\underline{\theta})=(1-\delta)\left(t_\tau(\underline{\theta})-\underline{\theta}q_\tau(\underline{\theta})\right)+\delta U_{\tau+1}^+(\underline{\theta})=(1-\delta)\left(t_\tau(\underline{\theta})-\underline{\theta}q_\tau(\underline{\theta})\right).$$

Summarizing items (1)–(3), we obtain:

$$t_\tau(\underline{\theta})=\begin{cases}\overline{\theta}q_\tau(\underline{\theta}) & \forall\tau\leq\tau'-1,\\ \underline{\theta}q_\tau(\underline{\theta})+\frac{\delta^{-\tau}}{1-\delta}\Delta\theta\left(q^{sb}(\overline{\theta})-q_\tau^-(\underline{\theta})\right) & \tau=\tau',\\ \underline{\theta}q_\tau(\underline{\theta}) & \forall\tau>\tau'.\end{cases} \tag{A.45}$$

We now prove that the transitory phase has length τ^*.

Lemma A.4. $\tau^* = \tau'$.

Proof of Lemma A.4. Note that the enforcement constraint is binding for $\tau \leq \tau^*$. Thus we have

$$\delta(S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta})) - (1-\delta)\underline{\theta} q_{\tau^*}(\underline{\theta}) + (1-\delta)M = \delta^{-\tau^*} \max\left\{\Delta\theta(q_0^+(\overline{\theta}) - q_{\tau^*}^-(\underline{\theta})), 0\right\}.$$

But we can find a lower bound for the left hand side as:

$$\delta(S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta})) - (1-\delta)\underline{\theta} q_{\tau^*}(\underline{\theta}) + (1-\delta)M \geq \psi(\underline{\theta}, q^{fb}(\underline{\theta})) > 0$$

(where the last inequality follows from the first condition in Assumption 2). Thus, we have

$$\max\left\{\Delta\theta(q_0^+(\overline{\theta}) - q_{\tau^*}^-(\underline{\theta})), 0\right\} = \Delta\theta(q_0^+(\overline{\theta}) - q_{\tau^*}^-(\underline{\theta})) > 0$$

and finally

$$\tau^* \leq \tau'.$$

Assume now that $\tau' \geq \tau^* + 1$. Then, both constraints (A.18) for τ' and $\tau' + 1$ are binding so that:

$$\Psi(\underline{\theta}, \mathbf{q}_{\tau'}^+(\underline{\theta})) = \delta^{-\tau} \Delta\theta(q_0^+(\overline{\theta}) - q_{\tau'}^-(\underline{\theta})),$$

and

$$\delta\Psi(\underline{\theta}, \mathbf{q}_{\tau'+1}^+(\underline{\theta})) = \delta^{-\tau} \Delta\theta(q_0^+(\overline{\theta}) - q_{\tau'+1}^-(\underline{\theta})).$$

By subtracting one equation from the other, we obtain:

$$\delta S(q_{\tau'+1}(\underline{\theta})) - \underline{\theta} q_{\tau'}(\underline{\theta}) + (1-\delta)M = \Delta\theta q_{\tau'}(\underline{\theta}).$$

Because $q_{\tau'+1}(\underline{\theta}) = q_{\tau'}(\underline{\theta}) = q^{fb}(\underline{\theta})$ we obtain

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) = 0$$

which yields a contradiction with the first condition of Assumption 2. We can thus conclude:

$$\tau^* \geq \tau'$$

which ends the proof of Lemma A.4. ||

Gathering Lemma A.4 and equation (A.45) yields (4.10) and ends the proof of the proposition. ||

Proof of Proposition 3. The structure of the solution to $(\mathcal{P})$ given by our earlier findings (with (A.18) being binding at all dates $\tau \leq \tau^*$) implies that the enforcement constraint (A.18) at date τ^* can be written as:

$$\delta^{\tau^*}\left(\delta(S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta})) - (1-\delta)\underline{\theta} q_{\tau^*}(\underline{\theta}) + (1-\delta)M\right) \geq \Delta\theta\left(q(\overline{\theta}) - (1-\delta)\sum_{\tau=0}^{\tau^*-1} \delta^\tau \Gamma^{\tau^*-\tau}(q_{\tau^*}(\underline{\theta}))\right). \tag{A.46}$$

Taking into account this structure of the solution, the optimal contract that solves $(\mathcal{P})$ must also solve the following problem:

$$(\mathcal{R}): \max_{(q_{\tau^*}(\underline{\theta}), q(\overline{\theta}), \tau^*)} (1-\delta)\left(\sum_{\tau=0}^{\tau^*} \delta^\tau (S(\Gamma^{\tau^*-\tau}(q_{\tau^*}(\underline{\theta}))) - \underline{\theta}\Gamma^{\tau^*-\tau}(q_{\tau^*}(\underline{\theta}))) + \sum_{\tau=\tau^*+1}^{\infty} \delta^\tau (S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta}))\right)$$

$$+ (1-\nu)\left(S(q(\overline{\theta})) - \overline{\theta} q(\overline{\theta})\right) - \nu\Delta\theta q(\overline{\theta})$$

subject to (A.46).

Problem $(\mathcal{R})$ gives deeper results on the nature of the solution to $(\mathcal{P})$. It allows to decompose the optimization into two phases. The first transitory phase over the first τ^* periods has a growing output $\Gamma^{\tau^*-\tau}(q_{\tau^*}(\underline{\theta}))$ for the low-cost seller till one reaches a value $q_{\tau^*}(\underline{\theta}) \leq q^{fb}(\underline{\theta})$ to be found. The second phase has a fixed output $q^{fb}(\underline{\theta})$. In both phases, the high-cost seller's output remains constant. Yet, we already know from Lemma A.3, that the solution to $(\mathcal{R})$ cannot have either $\tau^* = \infty$ or $\tau^* = -1$ (with the convention that $\sum_{\tau=0}^{\tau=-1} y_\tau = 0$) when Assumption 2 holds. Fixing τ^* in the maximand above defines a collection of programmes $(\mathcal{R}_{\tau^*})$. Whenever the corresponding constraint (A.46) is slack in $(\mathcal{R}_{\tau^*})$, the solution to $(\mathcal{R}_{\tau^*})$ entails an output profile such that $q_{\tau^*}(\underline{\theta}) = q^{fb}(\underline{\theta})$ and $q(\overline{\theta}) = q^{bm}(\overline{\theta})$. Therefore, it is not the solution to $(\mathcal{R})$ (and thus to $(\mathcal{P})$). Next lemma shows the existence of a first integer τ^* such that equation (A.46) taken at that date can no longer be slack. This is at such τ^* that $(\mathcal{R})$ achieves its maximum. Lemma A.5 provides a characterization of τ^*.

Lemma A.5. *Suppose that Assumption 2 holds. There exists a unique $\tau^* \geq 0$ such that:*

$$\delta^{\tau^*}\psi(\underline{\theta}, q^{fb}(\underline{\theta})) < \Delta\theta\left(q^{bm}(\overline{\theta}) - (1-\delta)\sum_{\tau=0}^{\tau^*-1}\delta^\tau \Gamma^{\tau^*-\tau}(q^{fb}(\underline{\theta}))\right), \tag{A.47}$$

$$\delta^{\tau^*+1}\psi(\underline{\theta}, q^{fb}(\underline{\theta})) \geq \Delta\theta\left(q^{bm}(\overline{\theta}) - (1-\delta)\sum_{\tau=0}^{\tau^*}\delta^\tau \Gamma^{\tau^*+1-\tau}(q^{fb}(\underline{\theta}))\right). \tag{A.48}$$

Proof of Lemma A.5. Denote

$$\vartheta(\tau) = \delta^{-\tau}\Delta\theta\left(q^{bm}(\overline{\theta}) - (1-\delta)\sum_{s=0}^{\tau-1}\delta^s \Gamma^{\tau-s}(q^{fb}(\underline{\theta}))\right).$$

Actually, $\vartheta(\tau)$ is a decreasing sequence. Indeed, $\vartheta(\tau+1) < \vartheta(\tau)$ amounts to $q^{bm}(\overline{\theta}) < \Gamma^{\tau+1}(q^{fb}(\underline{\theta}))$ which holds since, from Assumption 2, we have $q^{bm}(\overline{\theta}) < q^e(\overline{\theta}) < \Gamma^{\tau-s}(q^{fb}(\underline{\theta})) < q^{fb}(\underline{\theta})$ for all $\tau - 1 \geq s \geq 0$. From Assumption 2, we also know that $\psi(\underline{\theta}, q^{fb}(\underline{\theta})) < \vartheta(0)$. Moreover, we have $\lim_{\tau\to+\infty}(1-\delta)\sum_{s=0}^{\tau-1}\delta^s\Gamma^{\tau-s}(q^{fb}(\underline{\theta})) = q^e(\overline{\theta}) > q^{bm}(\overline{\theta})$ where the last inequality also follows from Assumption 2. Hence, for τ large enough, we also have $\vartheta(\tau) < 0$. Gathering these findings, there exists a unique $\tau^* \geq 0$ such that both inequalities (A.47) and (A.48) hold together. ‖

We can now use (A.47), (A.48), and the inequalities $q^e(\overline{\theta}) < \Gamma^{\tau^*-\tau}(q^{fb}(\underline{\theta})) < q^{fb}(\underline{\theta})$ to find the following bounds on τ^*

$$\delta^{\tau^*}\psi(\underline{\theta}, q^{fb}(\underline{\theta})) < \Delta\theta\left(q^{bm}(\overline{\theta}) - (1-\delta^*)q^e(\overline{\theta})\right)$$

and

$$\delta^{\tau^*+1}\psi(\underline{\theta}, q^{fb}(\underline{\theta})) \geq \Delta\theta\left(q^{bm}(\overline{\theta}) - (1-\delta^{\tau^*+1})q^{fb}(\underline{\theta})\right).$$

Taken together those inequalities give us (4.12). ‖

Proof of Theorem 2. Suppose now that only the privately informed seller might behave opportunistically. To construct a mechanism that implements the Baron–Myerson allocation and is seller-enforceable, it must be that Constraint (3.3) always holds at any date and for all types. First, the low-cost seller's enforcement constraint is not an issue if

$$U_0^+(\underline{\theta}) = \Delta\theta q^{bm}(\overline{\theta}) \geq U_\tau^-(\overline{\theta}) + (1-\delta^\tau)\Delta\theta q^{bm}(\overline{\theta}), \quad \forall \tau \geq 0.$$

The following stationary payment extracts the high-cost seller's surplus in each period and ensures that the latter constraints always hold:

$$t_{1\tau}(\overline{\theta}) + t_{2\tau}(\overline{\theta}) = t^{bm}(\overline{\theta}) = \overline{\theta}q^{bm}(\overline{\theta}) \text{ with } t_{1,\tau}(\overline{\theta}) = L \quad \forall \tau \geq 0.\text{[37]}$$

Turning now to the payments given to a high-cost seller to prevent the *take-the-money-and-run* strategy, equation (3.3) implies that backward payoffs $U_\tau^-(\underline{\theta})$ must satisfy:

$$U_0^+(\overline{\theta}) = 0 \geq U_\tau^-(\underline{\theta}) - (1-\delta^\tau)\Delta\theta q^{fb}(\underline{\theta}), \quad \forall \tau \geq 0.$$

Finding such payoffs (and thus the payments to a low-cost seller) is now easy. Define now τ^* as the highest integer such that $(1-\delta^\tau)\Delta\theta q^{fb}(\underline{\theta}) < \Delta\theta q^{bm}(\overline{\theta})$. Such integer exists and is unique because $q^{fb}(\underline{\theta}) > q^{bm}(\overline{\theta})$. Over the first τ^* periods, pre-delivery payments are adjusted so that the high-cost seller remains indifferent between breaching or not in each period:

$$U_\tau^-(\underline{\theta}) = (1-\delta^\tau)\Delta\theta q^{bm}(\overline{\theta}) + \delta^\tau(1-\delta)(t_{1,\tau}(\underline{\theta}) - L) = (1-\delta^\tau)\Delta\theta q^{fb}(\underline{\theta}), \quad \forall \tau < \tau^*.$$

After those τ^* earlier periods, pre-delivery payments implement a constant backward rent equal to the low-cost seller's Baron–Myerson information rent:

$$U_\tau^-(\underline{\theta}) = (1-\delta^\tau)\Delta\theta q^{bm}(\overline{\theta}) + \delta^\tau(1-\delta)(t_{1,\tau}(\underline{\theta}) - L) = \Delta\theta q^{bm}(\overline{\theta}), \quad \forall \tau \geq \tau^*.$$

Post-delivery payments are then adjusted to implement Baron–Myerson payments:

$$t_{1,\tau}(\underline{\theta}) + t_{2,\tau}(\underline{\theta}) = t^{bm}(\underline{\theta}) = \underline{\theta}q^{bm}(\underline{\theta}) + \Delta\theta q^{bm}(\overline{\theta}) \quad \forall \tau \geq 0.$$

‖

Proof of Proposition 4. Taking into account the expression of the renegotiation-proof output profiles, date 0-enforcement constraint (3.6) in state $\overline{\theta}$ can now be written as:

$$\delta\left(S(q^{ci}(\overline{\theta})) - \overline{\theta}q^{ci}(\overline{\theta})\right) - (1-\delta)\overline{\theta}q_0(\overline{\theta}) + (1-\delta)M \geq 0 \tag{A.49}$$

where $q^{ci}(\overline{\theta}) = \min\{q^e(\overline{\theta}), q^{fb}(\overline{\theta})\}$.

Using previous notations to express virtual costs, the buyer's intertemporal payoff at a renegotiation-proof contract becomes:

$$\mathbb{E}_\theta((1-\delta)(S(q_0(\theta))-m(\theta)q_0(\theta))+\delta(S(q^{ci}(\theta))-m(\theta)q^{ci}(\theta))). \tag{A.50}$$

The optimal renegotiation-proof date-0 outputs are thus obtained by maximizing this expression subject to the enforcement constraints (5.1) and (A.49). We first neglect (A.50) and optimize with (5.1) as the sole constraint. Denoting by λ_r the non-negative Lagrange multiplier pour (5.1), and optimizing yields the first-order conditions (5.2).

Suppose now that $\lambda_r=0$. Then equation (5.1) would become:

$$\delta S(q^{fb}(\underline{\theta}))-\underline{\theta}q^{fb}(\underline{\theta})+(1-\delta)M \geq \Delta\theta((1-\delta)q^{bm}(\overline{\theta})+\delta q^{ci}(\overline{\theta}))\geq \Delta\theta q^{bm}(\overline{\theta})$$

where the last inequality follows from the second condition in Assumption 2 which amounts to $q^{bm}(\overline{\theta})\leq q^{ci}(\overline{\theta})$. A contradiction with the first condition in Assumption 2. Hence, $\lambda_r>0$.

We now check that equation (A.49) is satisfied for the optimal output $q_0^r(\overline{\theta})$ which means:

$$\delta\left(S(q^{ci}(\overline{\theta}))-\overline{\theta}q^{ci}(\overline{\theta})\right)-(1-\delta)\overline{\theta}q_0^r(\overline{\theta})+(1-\delta)M\geq 0.$$

When $q^{ci}(\overline{\theta})=q^e(\overline{\theta})\leq q^{fb}(\overline{\theta})$, this latter condition holds since $q_0^r(\overline{\theta})<q^{bm}(\overline{\theta})$, with Assumption 2 holding, implies:

$$\delta\left(S(q^e(\overline{\theta}))-\overline{\theta}q^e(\overline{\theta})\right)-(1-\delta)\overline{\theta}q_0^r(\overline{\theta})+(1-\delta)M\geq\delta\left(S(q^e(\overline{\theta}))-\overline{\theta}q^e(\overline{\theta})\right)-(1-\delta)\overline{\theta}q^{bm}(\overline{\theta})+(1-\delta)M$$

$$\geq\psi(\overline{\theta},q^{bm}(\overline{\theta}))\geq 0.$$

||

Acknowledgments. For helpful comments on earlier versions of this work, we thank seminar and conference participants at the Canadian Economic Theory 2011 Meeting in Toronto, the Canadian Economic Association 2012 in Calgary, the 2013 Workshop on Conflict and Cooperation in Bergamo, the 2013 conference on "Contracts, Procurement, and Public-Private Arrangements" in Florence, Sciences Po in Paris, The 2013 IO Workshop in Berlin, REST-ENSAE in Paris, RECODE in Ottawa, the First Brazilian Conference Series on Public Procurement and Concession Design in Sao Paulo, the 2014 Workshop on Public Procurement in Mannheim, Oslo University, the 2014 Latin-American Econometric Society Meeting in Rio de Janeiro, and 2015 Workshop on Law Enforcement in Nanterre, France. All remaining errors are our own. Special thanks to Sandeep Baliga, Sergei Guriev, Henrik Lando, Rocco Macchiavello, Lucas Maestri, Ameet Morjaria, Marco Pagnozzi, Nicola Pavoni, and Peter Vida for helpful discussions on earlier versions. We also thank the Editor Marco Ottaviani and three referees whose comments have lead to substantive changes and improvements to this manuscript. The usual disclaimer applies. Martimort wishes to thank nancial support from the ANR Polycripe and the Program Investissements d'Avenir of the French government (ANR-10-LABX-93-01).

REFERENCES

ABREU, D. (1988), "On theory of Infinitely Repeated Games with Discounting", *Econometrica*, **56**, 383–396.

ATIYAH, P. (1995), *An Introduction to the Law of Contract* (Princeton: Princeton University Press).

BAKER, G., GIBBONS, R. and MURPHY, K. (1994), "Subjective Performance Measures in Optimal Incentive Contracts", *Quarterly Journal of Economics*, **109**, 1125–1156.

BAKER, G., GIBBONS, R. and MURPHY, K. (2001), "Bringing the Market Inside the Firm?" *American Economic Review*, **91**, 212–218.

BAKER, S. and CHOI, A. (2014), "Contract's Role in Relational Contract" (Virginia Law and Economics Research Paper No. 2014-01).

BARON, D. and BESANKO, D. (1984), "Regulation and Information in a Continuing Relationship", *Information Economics and Policy*, **1**, 267–302.

BARON, D. and MYERSON, R. (1982), "Regulating a Monopolist with Unknown Costs", *Econometrica*, **50**, 911–930.

BATTAGLINI, M. (2005), "Long-Term Contracting with Markovian Consumers", *American Economic Review*, **95**, 637–658.

BEBCHUK, L. (1984), "Litigation and Settlement under Imperfect Information", *The RAND Journal of Economics*, **15**, 404–415.

BERGIN, J. and MACLEOD, W. (1993), "Efficiency and Renegotiation in Repeated Games", *Journal of Economic Theory*, **61**, 42–73.

BERNHEIM, B. and WHINSTON, M. (1998), "Incomplete Contracts and Strategic Ambiguity", *American Economic Review*, **88**, 902–932.

BOARD, S. (2011), "Relational Contracts and the Value of Loyalty", *American Economic Review*, **101**, 3349-3367.

BULL, C. (1987), "The Existence of Self-Enforcing Implicit Contracts", *Quarterly Journal of Economics*, **102**, 147–159.

CALLAHAN, M. (2009), *Termination of Construction and Design Contracts* (New York: Aspen Law and Business).

CROCKER, K. and MASTEN, S. (1988), "Mitigating Contractual Hazards: Unilateral Options and Contract Length", *The RAND Journal of Economics*, **19**, 327–343.

CROCKER, K. and MASTEN, S. (1996), "Regulation and Administered Contracts Revisited: Lessons from Transaction-Cost Economics for Public Utility Regulation", *Journal of Regulatory Economics*, **9**, 5–39.

DECHERT, W. (1982), "Lagrange Multipliers in Infinite Horizon Discrete Time Optimal Control Models", *Journal of Mathematical Economics*, **9**, 285–302.

DEWATRIPONT, D. (1989), "Renegotiation and Information Revelation over Time: The Case of Optimal Labor Contracts," *Quarterly Journal of Economics*, **104**, 589–619.

DOORNIK, K. (2010), "Incentive Contracts with Enforcement Costs", *Journal of Law, Economics, and Organization*, **26**, 115–143.

EDLIN, A. (1998), Breach Remedies. In Newman, P. (ed.) *The New Palgrave Dictionary of Economics and the Law*, Vol. 1, pp. 174–179 (London: Macmillan).

EGGLESTON, B. (2009), *Liquidated Damages and Extensions of Time in Construction Contracts*, 3rd edn (Hoboken, New Jersey: Wiley-Blackwell).

ESCOBAR, J. F. and TOIKKA, J. (2013), "Efficiency in Games with Markovian Private Information", *Econometrica*, **81**, 1887–1934.

FARNSWORTH, E. A. (1982), *Contracts* (New York: Little Brown Publishers).

FONG, Y.-F. and LI, J. (2010), "Information Revelation in Relational Contracts", (Mimeo, Northwestern University, Kellogg School of Management).

FONG, Y.-F. and SURTI, J. (2009), "On the Optimal Degree of Cooperation in the Repeated Prisoner's Dilemma with Side-Payments", *Games and Economic Behavior*, **67**, 277–291.

FUCHS, W. (2007), "Contracting with Repeated Moral Hazard and Private Evaluations", *American Economic Review*, **97**, 1432–1448.

GHOSH, P., and RAY, D. (1996), "Cooperation in Community Interaction without Information Flows", *Review of Economic Studies*, **63**, 491–519.

GIBBONS, R. (2005a), "Four Formal(izable) Theories of the Firm", *Journal of Economic Behavior and Organization*, **58**, 200–245.

GIBBONS, R. (2005b), "Incentives Between Firms (and Within)", *Management Science*, **51**, 2–17.

GOLDLÜCKE, S. and KRANZ, S. (2013), "Renegotiation-Proof Relational Contracts", *Games and Economic Behavior*, **80**, 157–178.

GUASCH, J. L., LAFFONT, J. J. and STRAUB, S. (2003), "Renegotiation of Concession Contracts in Latin America", (Vol. 3011), World Bank Publications.

HALAC, M. (2012), "Relational Contracts and the Value of Relationships", *American Economic Review*, **102**, 750–79.

HALAC, M. (2015), "Investing in a Relationship", *The RAND Journal of Economics*, **46**, 165–185.

HART, O. and TIROLE, J. (1988), "Contract Renegotiation and Coasian Dynamics", *Review of Economic Studies*, **55**, 509–540.

HEMSLEY, P. (2013), "Dynamic Moral Hazard with Self-Enforceable Incentive Payments" (Mimeo Toulouse School of Economics).

HINZE, J. (2001), *Construction Contracts*, 2nd ed (New York: McGraw-Hill).

HORNER, J. (2002), "Reputation and Competition," *American Economic Review*, **92**, 644–663.

IOSSA, E. and SPAGNOLO, G. (2009), "Contracts as Threats: On a Rationale for Rewarding A while Hoping for B" (Mimeo, Brunel University).

ITOH, H. and MORITA, H. (2015), "Formal Contracts, Relational Contracts and the Threat-Point Effect," *American Economic Journal: Microeconomics*, **7**, 318–346.

JOHNSON, J. and SOHI, R. (2015), "Understanding and Resolving Major Contractual Breaches in Buyer-Seller Relationships: A Grounded Theory Approach", *Journal of the Academy of Marketing Science*, 1–21.

JOHNSON, S., MCMILLAN, J. and WOODRUFF, C. (2002), "Courts and Relational Contracts", *Journal of Law, Economics, and Organization*, **18**, 221–277.

JOSKOW, P. (1987), "Contract Duration and Relationship and Relationship-Specific Investments: Evidence from Coal Markets", *American Economic Review*, **77**, 168–185.

JOSKOW, P. (1988), "Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence", *Journal of Law, Economics, and Organization*, **4**, 98–115.

KORNHAUSER, L. and MACLEOD, B. (2012), "Contracts Between Legal Persons", In Gibbons, R. and Roberts, J. (eds), *The Handbook of Organizational Economics* (Amsterdam: North-Holland).

KVALØY, O. and OLSEN, T. (2009), "Endogenous Verifiability and Relational Contracting", *American Economic Review*, **99**, 2193–2208.

KWON, S. (2013), "Relational Contracts in a Persistent Environment" (Working Paper, University College London).

KWON, S. (2015), "Dynamic Moral Hazard with Persistent States" (Working Paper, University College London).

LAFFONT, J. J. and MARTIMORT, D. (2002), *theory of Incentives: The Principal-Agent Model* (Princeton: Princeton University Press).

LAFFONT, J. J. and TIROLE, J. (1993), *A Theory of Incentives in Procurement and Regulation* (Cambridge: MIT Press).

LEVIN, J. (2003), "Relational Incentive Contracts", *American Economic Review*, **93**, 835–857.

LEWIS, T. and SAPPINGTON, D. (1989), "Countervailing Incentives in Agency Problems", *Journal of Economic Theory*, **49**, 294–313.

LI, J. and MATOUSCHEK, N. (2013), "Managing Conflicts in Relational Contracts", *American Economic Review*, **103**, 2328–2351.

MACAULAY, S. (1963), "Non-Contractual Relations in Business: A Preliminary Study", *American Sociological Review*, **28**, 55–67.

MACCHIAVELLO, R. and MORJARIA, A. (2015), "The Value of Relationships: Evidence from a Supply Shock to Kenya Rose Exports", *American Economic Review*, **105**, 2911–2945.

MACLEOD, B. and MALCOMSON, J. (1989), "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment", *Econometrica*, **57**, 447–480.

MACLEOD, B. and MALCOMSON, J. (1998), "Motivation and Markets", *American Economic Review*, **88**, 388–411.

MALCOMSON, J. (2012), "Relational Incentive Contracts", In Gibbons, R. and Roberts, J. (eds), *Handbook of Organizational Economics* (Princeton: Princeton University Press).

MALCOMSON, J. (2016), "Relational Incentive Contracts with Persistent Private Information", *Econometrica*, **84**, 317–346.

MASKIN, E. and TIROLE, J. (2004), "The Politician and the Judge: Accountability in Government", *American Economic Review*, **94**, 1034–1054.

MASTEN, S. (1999), "Contractual Choice", In Boukaert, B. and DeGeest, G. (eds), *Encyclopedia of Law and Economics* (Edward Elgar Publishing).

MAESTRI, L. (2011), "Dynamic Contracting under Adverse-Selection and Renegotiation" (Mimeo).

OSBORNE, M. J. and RUBINSTEIN, A. (1990), *Bargaining and Markets* (San Diego, CA: Academic Press).

PAVAN, A., SEGAL, I. and TOIKKA, J. (2014), "Dynamic Mechanism Design: A Myersonian Approach", *Econometrica*, **82**, 601–653.

PEARCE, D. and STACCHETTI, E. (1998), "The Interaction of Implicit and Explicit Contracts in Repeated Agency," *Games and Economic Behavior*, **23**, 75–96.

POSNER, E. A. (2011), *Contract Law and Theory.* Aspen Student Treatise Series. (Alphen aan den Rijn, Netherlands: Wolters Kluwer Legal Publications).

RAHMAN, D. (2012), "But Who Will Monitor the Monitor?" *American Economic Review*, **102**, 2767–2797.

RAMEY, G. and WATSON, J. (2001), "Bilateral Trade and Opportunism in a Matching Market", *The B.E. Journal of Theoretical Economics*, **1**, 1–35.

REINGANUM, J. and WILDE, L. (1986), "Settlement, Litigation, and the Allocation of Litigation Costs", *The RAND Journal of Economics*, **17**, 557–566.

REY, P. and SALANIÉ, B. (1996), "On the Value of Commitment with Asymmetric Information," *Econometrica*, **64**, 1395–1414.

RIORDAN, M. (1990), "What Is Vertical Integration," in Aoki, M., Gustaffson, B. and Williamson, O. (eds), *The Firm as a Nexus of Treaties* (Thousand Oaks: Sage Publications).

SCHMIDT, K. and SCHNITZER, M. (1995), "The Interaction of Explicit and Implicit Contracts", *Economics Letters*, **48**, 193–199.

SCHWARTZ, A. and WATSON, J. (2004), "The Law and Economics of Costly Contracting", *Journal of Law, Economics, and Organization*, **20**, 2–31.

SHAVELL, S. (1980), "Damage Measures for Breach of Contract", *Bell Journal of Economics*, **11**, 466–499.

SHAVELL, S. (2004), *Foundations of Economic Analysis of Law* (Cambridge: Harvard University Press).

SOBEL, J. (1985), "A Theory of Credibility", *Review of Economic Studies*, **52**, 557–573.

VINTER, G. and PRICE, G. (2006), *Project Finance: A Legal Guide* (London: Sweet Maxwell).

WATSON, J. (1999), "Starting Small and Renegotiation", *Journal of Economic Theory*, **85**, 52–90.

WATSON, J. (2002), "Starting Small and Commitment", *Games and Economic Behavior*, **38**, 176–199.

YANG, H. (2012), "Nonstationary Relational Contracts with Adverse Selection", *International Economic Review*, **54**, 525–547.