CAHIER DE RECHERCHE #1304E
Département de science économique
Faculté des sciences sociales
Université d'Ottawa

WORKING PAPER #1304E
Department of Economics
Faculty of Social Sciences
University of Ottawa

A Theory of Contracts with Limited Enforcement*

David Martimort†, Aggey Semenov‡ and Lars Stole§

October 2013

* Acknowledgements: 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, CREST-ENSAE in Paris, REDECODE in Ottawa. All remaining errors are our own. Special thanks to Sandeep Baliga, Sergei Guriev, Marco Ottavianni, Marco Pagnozzi and Nicola Pavoni for helpful discussions on earlier versions.
† Paris School of Economics-EHESS, Email: david.martimort@parisschoolofeconomics.eu.
‡ Affiliation, Department of Economics, University of Ottawa, 55 Laurier E., Ottawa, Ontario, Canada, K1N 6N5; Email: aggey.semenov@uottawa.ca.
§ University of Chicago Booth School of Business, Email: lars.stole@chicagobooth.edu.

Abstract

We present a Theory of Contracts under costly enforcement in the context of a dynamic relationship between an uninformed buyer and a seller who is privately informed on his persistent cost at the outset. Public enforcement relies on remedies for breach. Private enforcement comes from severing relationships. We first characterize aggregate enforcement constraints ensuring that trading partners do not breach contracts unduly. Whether a long-term contract is enforceable does not depend on the distribution of penalties for breach between the buyer and the seller. While under complete information, the optimal contract would remain stationary, non-stationarity might arise under asymmetric information. Enforcement constraints are time-dependent and easier to satisfy as time passes. Indeed, 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. Yet, such take-the-money-and-run strategy becomes less attractive as time passes and can be prevented with back loaded payments. The optimal contract thus goes through two different phases. First, quantities and prices increase at the inception of the relationship. Later on, the contract looks more stationary. Long-run screening distortions encapsulate the quality of enforcement, offering de facto a link between the quality of the legal system and contractual performances.

Key words: *Asymmetric information, enforcement, breach of contracts, dynamic contracts.*
JEL Classification: D82, D86, K2.

Résumé

Nous présentons une théorie des contrats avec exécution coûteuse dans le contexte d'une relation dynamique entre un acheteur non informé et un vendeur avec information privée quant à son coût persistant au départ. L'exécution publique des contrats s'appuie sur les recours pour violation. L'exécution privée consiste à de rompre les relations. En premier lieu, nous caractérisons les contraintes d'exécution globale. La possible exécution d'un contrat à long terme ne dépend pas de la distribution des pénalités pour rupture entre l'acheteur et le vendeur. Sous information complète, le contrat optimal resterait stationnaire, alors qu'il pourrait être non stationnaire lorsque l'information est asymétrique. Les contraintes de l'application dépendent du temps et sont plus faciles à satisfaire à mesure que le temps passe. En effet, un vendeur avec des coûts élevés peut être tenté d'échanger de grandes quantités à des prix élevés au début de la relation avant de rompre le contrat par la suite. Pourtant, telle stratégie, prendre l'argent et courir, devient moins attrayante à mesure que le temps passe et peut être évitée. Le contrat optimal passe donc par deux phases différentes. Tout d'abord, les quantités et les prix augmentent lors de la création de la relation. Plus tard, on le contrat semble plus stable. Les distorsions dues au dépistage de long terme signalent, ce qui suggère de facto un lien entre la qualité du système légal et les performances contractuelles.

Mots clés : *Asymétrie de l'information, exécution de contrats, violation des contrats, contrats dynamiques.*
Classification JEL : D82, D86, K2.

1. INTRODUCTION

MOTIVATION. The *Theory of Contracts* is a major building block of our current understanding of how markets and organizations perform. From a theoretical viewpoint, its impressive contribution is to allow a full characterization of trading possibilities in environments where privately informed parties might have conflicting interests. Equipped with such a description of the set of *incentive feasible* allocations (Myerson, 1982), modelers might first optimize institutions, organizations and contractual forms and then discuss practical implementation. Although quite successful, this paradigm remains mute over the constraints on enforcing contractual agreements that are observed in practice. At best, economists allude to the mere presence of Courts of Law with crude enforcement capabilities in the background of the contracting environment they entertain.[1] This view of contracts somewhat somewhat contrasts with the perspective taken by Law scholars who actually devote much effort in studying the crucial role played by Courts in enforcing and implementing contracts.[2] Instead, a view broadly shared by economists is that, if any frictions impede contractual performances, those frictions only come from asymmetric information and/or transaction costs, not from the limited ability of Courts in enforcing contracts.

This narrow focus of Contract Theory suffers from at least two oversights. The first weakness of the paradigm challenges its relevance for other fields of social sciences: Contract Theory typically does not draw any precise link between the quality of legal environments and contractual performances. This sharply contrasts with much empirical evidence. In a series of seminal papers, La Porta and al. (1998) and Shleifer and Vishny (1997) demonstrated that the quality of the legal system has a significant impact on transactions in financial, labor and output markets, and is thus a significant determinant of growth and development. Taking an even broader perspective, historians, political scientists (Tilly 2007, Fukuyama 2004) and some economists as well (Acemoglu and Robinson, 2005) emphasized that the enforcement of contracts and regulations is a key ingredient in building state capacities both across space and over history.

The second weakness of received theory is more internal to the field. That Contract Theory does not incorporate the imperfections of the legal system certainly provides a naive and possibly misleading view of optimal contract design. Without an explicit accommodation of legal imperfections, it is unclear *a priori* whether the received lessons of Contract Theory are robust to costly enforcement. Indeed, a richer set of comparative statics and correlations may emerge from a theory that links the quality of enforcement with more standard agency distortions.

To contribute on these fronts, this paper develops a theory of contracting with costly

[1] To illustrate, that some investment/action remains non-verifiable by Courts is a key ingredient not only for the "hold-up model" which is a basic tenet for both Transaction Costs Economics (Williamson, 1985) and the Property Rights literature (Hart, 1995) but also for the moral hazard leg of the Agency literature (Hart and Holmström, 1987).

[2] This view is probably best expressed 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 gap and devise and implement default rules."*

enforcement when parties may renege on their contractual duties and Courts have limited remedies against breaches. Our goal is to link the standard screening distortions which are now well-known in asymmetric information contexts to parameters characterizing the quality of enforcement.

In particular, our approach will emphasize two departures from the literature. As in the relational contracting literature,[3] we assume that the legal enforcement of long-term contracts is limited and therefore dynamic enforcement constraints must be satisfied in each period in order to prevent either the buyer or seller from reneging on their agreement. We will go further, however, and explicitly assume that the parties to the trading relationship can make use of existing (*albeit*, imperfect) legal institutions to obtain a limited amount of enforceability. The precise enforcement limitation that we assume is that courts are unwilling to enforce damages for breach beyond a certain, exogenous upper bound.[4] Within the allowable range of damages, however, the parties are free to design a long-term contract to require payments following any "breach" event. Thus, the buyer and seller will seek to construct their long-term relationship using the constrained remedies for breach and respecting the requirements for dynamic enforceability at each period of time.

Our second departure from much of the relational-contracting research is to assume an inherent asymmetric information contracting problem that would exist even in a world of perfectly-enforceable contracts. In particular, we assume that at the time of contracting, one party (the seller) is privately informed about production costs, and that this production cost is persistent throughout the duration of the relationship. This might seem an odd modeling choice at first glance. If, however, we assumed that the seller's production costs are newly drawn in each period of the relationship and that the parties are able to design their long-term relationship *before* the seller is informed about costs (as in Levin (2003), for example), then the asymmetric-information contracting problem is not a fundamental issue in absence of enforcement constraints. Indeed, with fully-enforceable contracts, the first best would be implemented by the trading parties: ex ante, the buyer would simply offer to purchase output at a price equal to the buyer's surplus, minus a fixed transfer from the seller to the buyer. In this case, the information-incentives problem emerges only when there is also a lack of contract enforcement. It is also in this case that, when contracts are not enforceable, Levin (2003) has established, *inter alia*, that the optimal contract is stationary. In contrast, with a persistent information-incentives problem as in this paper, we will find that the optimal contract will not be stationary and that the dynamics are decidedly more subtle.

Set-up and main results. Consider a highly stylized model of a relationship between an uninformed buyer (the principal, or *she* in the sequel) and a seller (the agent, *he*) who repeatedly trade for the delivery of a good or service. The seller has private information on his cost parameter, and this type is persistent over the whole relationship. We assume

[3]See Section 9 for a review of the relevant literature.

[4]This restriction is known as the *"penalty doctrine"* in Contract Law. Although parties can include in their contract an explicit statement about the damages that should be paid by a breaching party (i.e., "liquidated" or "stipulated" damages), courts have routinely refused to enforce such clauses if the damages seem unreasonably high compared to actual or anticipated losses. In such cases, the courts rule the damages are excessive "penalties" and therefore unenforceable. See Farnsworth (1982) and Posner (2011) for discussions of this doctrine. 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.

that the buyer has all bargaining power in designing a long-term contract, and that the parties can contract on a sequence of trades, specifying prices and quantities over time. Lastly, we assume that any departure from the long-term contract will result in a limited penalty which must be paid by the breaching party.

From the seminal work of Baron and Besanko (1984), the optimal dynamic contract in this context is known to be stationary – a sequence of repetitions in each period of the optimal, one-shot contract. Such a contract trades off the efficiency gains of increased production with the cost of leaving information rent to the informed seller.[5]

Slightly departing from this well-known setting, suppose now that this long-term relationship is subject to bilateral opportunism. In particular, the contract can be breached at any point in time by either party. For instance, the buyer may behave opportunistically and fail to fulfill her payment obligations following delivery while, on his side, the seller may pocket any advance payments and not deliver the requested output. Contracts should now structure the relationship to reduce such bilateral opportunism and ensure that, any point of time, parties prefer to stick to their contractual obligations rather than deviating. The issue at the core of our analysis is thus to understand how to enforce the actual play of a contract when Courts of Law will only impose limited damages or payments following a breach. As noted previously, the 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 compensation. It is hard to articulate an internally consistent theory for this "penalty doctrine," as it is called.[6] Nonetheless, it is a practical constraint that damages for breach cannot be set arbitrarily high. We thus accept this limitation on the *public* enforcement of contracts and ask how parties will optimally design their contract to provide *private* incentives for fulfilling their contractual obligations.

Enforcement constraints. Even when taken in tandem, the private and public sides of enforcement[7] may fail to ensure perfect enforceability. This is so when the perspective of future trading rounds may not suffice to motivate parties to abide to their current obligations or remedies have not been well-designed by Courts. A first step of our analysis thus consists in deriving enforcement constraints ensuring that both the seller and the buyer abide to the contract. Intuitively, the seller is ready to deliver the requested output only if pre-delivery payments are not too large. Simultaneously, the buyer is ready to pay the requested price for service if, on the contrary, post-delivery payments are small enough.

When traders have quasi-linear payoffs, individual enforcement constraints can be pooled into a single forward-looking *aggregate enforcement constraint.* For each realization of the seller's cost, the foregone benefits of future trades plus the overall remedies paid following a breach by either party must be large enough to ensure compliance.[8] The

[5]See Baron and Myerson (1982), and Laffont and Martimort (2002, Chapter 2) for a textbook presentation.

[6]See Posner (2011, Chapter 8) for a discussion. The Courts have also routinely refused to allow for damages when the losses are speculative and look too much to the uncertain future – the "speculative loss" doctrine.

[7]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."*

[8]In the mechanism design literature, it is a standard to trick to pool altogether incentive and partic-

value of this pooling procedure does not depend on whether there is asymmetric information or not even though details differ. It just follows from the possibility to structure within each period of the relationship pre- and post-delivery prices so that enough instruments are available to curb opportunism by both parties, and from the assumption that the court can identify any party failing to perform according to the contract.

Neutrality of the distribution of remedies. It follows from this pooling procedure that aggregate enforcement constraints do not depend on the distribution of punitive remedies between the buyer and the seller but only on their sum. Following a Coasean argument, the exact distribution of breach remedies has no impact. The buyer, who has all bargaining power, can undo any initial distribution and structure payments so as to fully internalize the consequences of a breach even when it might originate from the seller.

Take the money and run. Aggregate enforcement constraints take different forms depending on whether contracting occurs under complete or asymmetric information. Under complete information, those constraints only impose a non-negativity condition on the *enforcement surplus*: The value of future trades must cover the benefits of breaches by either party. This condition on present value is static in nature and does not depend directly on time, except insofar as different trading volumes and prices have been stipulated over the horizon.

Under asymmetric information, a first important novelty is that the value of future trades must be even greater to ensure enforcement. Future surplus must cover the consequences of the new strategic possibilities that are now available to the privately informed seller. The buyer must concede some information rent to the most efficient seller to induce information revelation. With imperfect enforcement, those large payments become attractive even for a high-cost seller who may just mimic efficient trades for a while before reneging on his obligations.

This *take-the-money-and-run* strategy is a fundamental feature of the asymmetric information environment. By construction, it is *time dependent*. As time passes, that strategy becomes less and less attractive for a high-cost seller. At the extreme, always mimicking an efficient type would just contradict incentive compatibility. After enough trading rounds, the enforcement problem then looks like under complete information.[9]

Optimality of non-stationary contracts. Under complete information, time plays no role in the enforcement constraints. A rough intuition might thus suggest that the optimal contract is then stationary. Such result is far from obvious. Relaxing enforcement constraints indeed calls for reducing current prices and volumes so that the current benefits of not delivering the good for the seller or of not paying for the service for the buyer diminish. At the same time, increasing future volumes maximizes future surplus and helps also to

ipation problems through an aggregate feasibility condition that broadly applies in environments with multilateral incentive problems. See, among others, Myerson and Sattherwaite (1983) in bargaining contexts and Mailath and Postelwaite (1999) in public good environments. In the present context, the bilateral incentive problems comes from the ability of either party to breach the contract. In the relational contracting literature which is closer to our framework, pooling incentives on both sides of the relationship is also a common feature as it was shown by Levin (2003) in a different informational context.

[9]Construction contracts offer well-known examples where such *take-the-money-and-run* strategy is pervasive. For anecdotical evidence see for instance http://realestate.msn.com/article.aspx?cp-documentid=20741445.

prevent opportunism by making future trading periods attractive. Because of those conflicting forces, the proof of stationary under complete information is somewhat involved, obliging us to import tools for optimization problems with infinite constraints that were developed in the macro economics literature.[10] Of course, this stationary contract still accounts for the possibility that enforcement constraints might bind.

Under asymmetric information, time now plays a key role due to the existing *take-the-money-and-run* strategy that pertains to such environment. Making such a strategy less attractive requires the parties to backload the low-cost seller's payments. The optimal contract is no longer stationary but goes through two different phases. In the first *growing phase*, trading volumes and payments increase over time as the high-cost seller's incentives to renege progressively decrease. In a more *mature phase*, outputs and prices become stationary. Yet, they differ from those obtained when contracts are *a priori* restricted to be stationary by means of less significant distortions. As time goes on, enforcement constraints have less bite and back-loading payments and volumes becomes less attractive.

Organization of the paper. Section 2 presents the model. Section 3 characterizes the set of allocations that are incentive feasible and enforceable, either under complete or asymmetric information. Section 4 shows that the optimal contract is stationary under complete information. Section 5 characterizes the optimal contract in the case of asymmetric information and provides conditions for their non-stationarity. Section 7 discusses the robustness of our findings under various scenarios. Section 8 draws some implications of our results for organization theory and, more specifically, for the Theory of the Firm. Section 9 concludes with a review of the literature that puts our findings into perspective. Proofs are relegated to an Appendix.

2. THE MODEL

2.1. *Basics*

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

A *trade profile* stipulates 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 per-period average 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 principal's gross surplus function S is increasing and strictly concave ($S' > 0 > S''$) with $S(0) = 0$ and $S'(0) < +\infty$. To ensure positive outputs under all circumstances below (even under asymmetric information and limited enforcement), we also assume that $S'(0)$ is not too small.

INFORMATION. Our goal is to study how screening possibilities interact with limits on enforcement. The seller has thus private information about his cost parameter θ which is

[10] See Rustichini (1998), Marcet and Marimon (2011) and Messner, Pavoni and Sleet (2013) among others.

persistent over the whole relationship but drawn once for all before contracting. The set of possible cost values is $\Theta = \{\underline{\theta}, \overline{\theta}\}$ (with $\Delta\theta > 0$). Let denote by $\nu > 0$ (resp. $1-\nu > 0$) the probability that the seller has a low (resp. high) cost.[11]

2.2. *Mechanisms and Costless Enforcement*

As a benchmark, consider the simple case where enforcement is costless.

Complete information. Under complete information, it is routine to show that the first-best outcome entails a stationary payment-output pair $(t^{fb}(\theta), q^{fb}(\theta))$ (or a stationary rent-output pair $(U^{fb}(\theta), q^{fb}(\theta)))$ defined respectively as:

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

The first condition just states that marginal surplus must be equal to marginal cost while the second shows that the seller's rent is fully extracted.When enforcement is costless, this allocation is repeated in each period of the relationship.

Asymmetric information. Under full commitment, Baron and Besanko (1984) have shown that the *Revelation Principle* applies in such dynamic trading environment. 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 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 the optimal long-term contract is the infinite replica of the static optimal contract. The latter entails first-best production for the low-cost seller and the usual *Baron-Myerson distortion* for the high-cost seller's output $q^{bm}(\overline{\theta})$:

$$S'(q^{bm}(\overline{\theta})) = \overline{\theta} + \frac{\nu}{1-\nu}\Delta\theta.$$ [12]

Under asymmetric information, the buyer's marginal surplus must equal the seller's *virtual cost* $\overline{\theta} + \frac{\nu}{1-\nu}\Delta\theta$ which is above true cost. 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 up 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 breaches by either side. A trader breaches the contract if the current benefit of doing so exceeds the

[11] The case of a continuum of types is analyzed in Section 7.2.

[12] Provided that $S'(0)$ is not too small as assumed, $q^{bm}(\overline{\theta})$ remains positive.

cost. This cost includes the foregone opportunities for future trades (the private side of enforcement) but also the legal remedies that this party may have to pay (the public side of enforcement). A contract stipulates prices and quantities that will be traded in any given period but trade itself is at risk.[13]

To control those opportunistic behaviors, it is important to distinguish, for each date τ, between a *pre-* and a *post-delivery* sub-period. The buyer pays $t_{1,\tau}$ before any production takes place and an additional payment $t_{2,\tau}$ after delivery. Having the buyer pay before production helps controlling her own incentives for breach. Instead post-trading payments help controlling the seller who could be tempted not to deliver the good as requested.[14,15] 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 a type θ is thus $t_\tau(\theta) = t_{1,\tau}(\theta) + t_{2,\tau}(\theta)$ at date τ.

Equipped with those notations, let us see in more details how parties may behave opportunistically and deviate from the expected play of the mechanism.

Seller's breach. After having pocketed the pre-delivery payment $t_{1,\tau}(\theta)$, the seller may not deliver the quantity $q_\tau(\theta)$.[16] This deviation is attractive if $t_{1,\tau}(\theta)$ is large enough.

Buyer's breach. After the delivery of a quantity $q_\tau(\theta)$, the buyer may not pay the post-delivery price $t_{2,\tau}(\theta)$ if this payment is now too large.

Denote by $K \geq 0$ and $L \geq 0$ the remedies that apply respectively to the buyer and the seller in case of a breach.[17] Those penalties are exogenously specified by Courts. Even though our characterization of enforceable contracts will focus on individual breaches,[18] it will nevertheless be useful to define the *aggregate penalty* as $M = K + L$.[19]

Timing. The following timing summarizes how the contracting game unfolds.

[13] By assuming full commitment, we leave aside the issue of renegotiation which is well known from the existing literature (Dewatripont 1989, Hart and Tirole 1988 and Rey and Salanié 1996). In that literature, the only breach of an initial long-term contract comes from the possibility that a party offers a Pareto-improving new contract as information has been revealed by earlier performances. This issue is orthogonal to the enforcement problem that is at the core of our analysis.

[14] Our interest for details of the technology of the trading process within each period and the opportunities that this trading process opens for opportunistic behavior is, at a rough level, similar to Beaudry and Poitevin (1995) and Watson (2007). Those authors argue that technological details shape contracting in environments with limited commitment.

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

[16] An alternative interpretation is that the good might come under two different levels of nonverifiable quality. Producing the low-quality good does not cost anything but is also of no value for the buyer.

[17] Those remedies are effective remedies. They could indeed be considered as expected remedies that incorporate the probability that law is just not enforced by Courts, maybe because of a "speculative loss" doctrine or because of corruption in some weak institutional contexts.

[18] We focus on each party's individual incentives for breach given he expects the other to behave.

[19] When remedies are infinite ($K = L = +\infty$), the optimal contract is of course the infinite replica of the Baron-Myerson allocation. This can be viewed as a stylized modeling of a perfect enforcement system. To focus on less trivial cases, we will thus implicitly assume that penalties are finite ($K, L < +\infty$).

1. At date $\tau = 0^-$, the seller learns his cost parameter θ. The buyer offers the mechanism $\mathcal{C}$. The seller accepts or rejects in which case parties receive their reservation values normalized to zero. If he accepts, the seller reports a type $\hat{\theta}$.
2. Each trading period at date $\tau \geq 0$ then unfolds 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 in which case he must pay 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 by either party, the contract is terminated.[20]

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

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

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

$$\overline{V}_\tau(\theta) = (1-\delta)\sum_{s=0}^{\infty} \delta^s \left(S(q_{\tau+s}(\theta)) - t_{\tau+s}(\theta)\right).$$

We may also define the seller's **backward** output and per-period average rent till date $\tau \geq 0$ respectively as: [21]

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

$$\underline{U}_\tau(\theta) = (1-\delta)\sum_{s=0}^{\tau-1} \delta^s (t_s(\theta) - \theta q_s(\theta)) + \delta^\tau (1-\delta)(t_{1,\tau}(\theta) - L).$$ [22]

Stationary contracts. For future references, it is useful to describe the character of stationary contracts. Those contracts specify payments and outputs that are constant over time. A stationary contract is thus of the form:

$$t_{i,\tau}(\theta) \equiv t_i(\theta) \quad (i = 1, 2), \quad q_\tau(\theta) \equiv q(\theta) \quad \forall \tau, \quad \forall \theta \in \Theta.$$

The seller's per-period rent $U(\theta)$ and the buyer's per period payoff $V(\theta)$ remain constant with such contract and are respectively equal to:

$$U(\theta) \equiv \overline{U}_0(\theta) \text{ and } V(\theta) \equiv \overline{V}_0(\theta).$$

[20] We follow here Abreu (1988) and the literature on relational contracts (Levin 2003, Halac 2012) in specifying that the worst equilibrium is played following a breach.

[21] Of course, the following identities hold: $\underline{U}_\infty(\theta) \equiv \overline{U}_0(\theta)$ and $\underline{q}_\infty(\theta) \equiv \overline{q}_0(\theta) \quad \forall \theta \in \Theta$ where $\overline{U}_0(\theta)$ (resp. $\overline{q}_0(\theta)$) is the seller's intertemporal rent (resp. output) over the whole relationship.

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

From Baron and Besanko (1984), we already know that restricting the analysis to stationary contracts is without loss of generality when enforcement is costless. Putting optimality aside, stationary contracts also remain of particular interest. First, stationary contracts might sometimes be the only feasible agreements when, for instance, contracting periods cannot be specified a priori.[23] Second, stationarity may be required if the buyer starts trading with several potential sellers at different points in time and no third-price discrimination across sellers is possible. Finally, Levin (2003) has also shown that stationary contracts remain optimal when types are independently drawn over time even under limited enforcement. Our analysis below unveils the cost of imposing stationarity on the parties when types are persistent.

3. INCENTIVE-FEASIBLE AND ENFORCEABLE ALLOCATIONS

This section characterizes the set of incentive feasible allocations under costly enforcement. 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 enforcement constraints.

INCENTIVE-FEASIBLE AND ENFORCEABLE ALLOCATIONS. A seller with type θ accepts the long-term contract when the following participation constraint holds:

$$\overline{U}_0(\theta) \geq 0, \quad \forall \theta \in \Theta. \tag{3.1}$$

We do not impose *a priori* a participation constraint from any date $\tau \geq 1$ on. Incentives to renege and leave the relationship are precisely addressed by the enforceability constraints that are more specific to our context.

BUYER'S ENFORCEABILITY. The buyer should pay in each period her due payment to the seller. If she were to deviate and not to pay for delivery, she would incur K and the relationship would end (the private side). These enforcement constraints are described below in compact form.

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

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

The lefthand side represents the buyer's discounted payoff from period $\tau+1$ onwards on the equilibrium path. The righthand side is her deviation payoff for the current period.

SELLER'S ENFORCEABILITY*. We first define the notion of enforceability (marked with a superscript * from now on) under complete information.

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

$$\overline{U}_0(\theta) \geq \underline{U}_\tau(\theta) \quad \forall \theta \in \Theta, \quad \forall \tau. \tag{3.3}$$

[23] For instance, trading opportunities could follow a Poisson process with non-verifiable arrival dates.

The seller should never find it optimal to breach the relationship at any date τ. The benefits of continuing trades must exceed the short-term gains from pocketing a pre-delivery payment at date τ and not delivering the good in the future. In other words, (3.3) also amounts to:

$$\overline{U}_\tau(\theta) \geq (1-\delta)(t_{1,\tau}(\theta) - L). \tag{3.4}$$

Seller's enforceability. Under asymmetric information, the seller's enforcement constraints are more complex because they interact with incentive compatibility requirements. Indeed, when the agent may deviate and not fulfill his obligations, the use of the Revelation Principle must be taken with caution. The possibility for non-compliance affects how incentive constraints should be written.

Definition 3 *The mechanism $\mathcal{C}$ is **seller-enforceable** if and only if:*

$$\overline{U}_0(\theta) \geq \max_{\hat{\theta}\in\Theta} \underline{U}_\tau(\hat{\theta}) + (\hat{\theta} - \theta)\underline{q}_\tau(\hat{\theta}) \quad \forall\theta \in \Theta, \quad \forall\tau. \tag{3.5}$$

The enforcement constraints (3.5) say that a type θ seller prefers to take the corresponding contract rather than adopting a *take-the-money-and-run* strategy.[24] This strategy consists in mimicking a type $\hat{\theta}$ at all dates $0....\ \tau - 1$, delivering the corresponding output, and then breaching at date τ, being only punished from that date onwards.

Playing on the seller's enforcement constraints highlights a few interesting cases. First, taking $\tau = \infty$, the enforceability conditions (3.5) imply standard incentive compatibility:

$$\overline{U}_0(\theta) = \max_{\hat{\theta}\in\Theta} \overline{U}_0(\hat{\theta}) + (\hat{\theta} - \theta)\overline{q}_0(\hat{\theta}), \quad \forall\theta \in \Theta. \tag{3.6}$$

In turn, a well known consequence of incentive compatibility is that only monotonically decreasing output schedules can be implemented:

$$\overline{q}_0(\theta) \text{ weakly decreasing.} \tag{3.7}$$

Second, a breach of contract after some truthful behavior is still an option for the seller. Preventing such behavior again yields constraints (3.4). However, note that the possibility of immediately reneging on the contract before any trade gives also:

$$\overline{U}_0(\theta) \geq (1-\delta)\left(\max_{\hat{\theta}\in\Theta} t_{1,0}(\hat{\theta}) - L\right).^{25} \tag{3.8}$$

Enforceability. Let us conclude this section by stating the definition of an enforceable mechanism, robust to bilateral opportunism.

Definition 4 *The mechanism $\mathcal{C}$ is **enforceable** (resp. **enforceable***) *if and only if it is both buyer- and seller-enforceable (resp. seller-enforceable*).*

[24] Here, we borrow the vocabulary familiar from models with short-term contracting (see for instance, Laffont and Tirole, 1993-Chapter 9, and Rey and Salanié, 1993) although we consider a different sort of limits on commitment.

To get a compact formulation for the aggregate enforcement constraints under various informational scenarios, we now define the *enforcement surplus* as:

$$\Psi(\theta, \overline{\mathbf{q}}_\tau(\theta)) = (1-\delta)\left(\delta \sum_{s=0}^{\infty} \delta^s (S(q_{\tau+s+1}(\theta) - \theta q_{\tau+s+1}(\theta)) - \theta q_\tau(\theta) + M\right). \tag{3.9}$$

The enforcement surplus represents the net gain for the trading partners when they choose to enforce the contract from date τ on. It takes into account future gains from trade but also the foregone penalties when not deviating at date τ.

4. Complete information: stationary contracts are optimal

In this section, we suppose that the cost parameter θ is common knowledge.

Characterization of enforceable* profiles. Next Lemma first explains the role that the enforcement surplus plays in characterizing feasible allocations:

Lemma 1 *A mechanism $\mathcal{C}$ is **enforceable*** if and only if implements a trade profiles with a non-negative enforcement surplus at any date $\tau \geq 0$:*

$$\Psi(\theta, \overline{\mathbf{q}}_\tau(\theta)) \geq 0 \quad \forall \theta \in \Theta. \tag{4.1}$$

The feasibility condition (4.1) is easily obtained by summing the individual enforcement constraints (3.2) and (3.3) under complete information. Under limited enforcement, feasibility requires that the discounted value of future trades covers the sum of the individual costs for breaching the contract.[26] To understand the different terms in the lefthand side of (4.1), it is nevertheless useful to think of the case of a contract stipulating that the buyer should appropriate all gains form trade in every period, i.e.,

$$t_\tau(\theta) = \theta q_\tau(\theta) \quad \forall \theta \in \Theta, \quad \forall \tau.$$

Observe that such payments trivially satisfy (3.3) if the pre-delivery payment is chosen so that $t_{1,\tau}(\theta) = L$ (and thus $t_{2,\tau}(\theta) = \theta q_\tau(\theta) - L$). The aggregate enforcement constraint (4.1) can now be written as:

$$(1-\delta)\delta \sum_{s=0}^{\infty} \delta^s (S(q_{\tau+s+1}(\theta) - \theta q_{\tau+s+1}(\theta)) \geq (1-\delta)(t_\tau(\theta) - M) \quad \forall \theta \in \Theta, \quad \forall \tau. \tag{4.2}$$

The lefthand side of (4.2) is thus the buyer's future gains from trade provided he extracts the seller's surplus in all future periods. The righthand side is simply the net benefit of a deviation, i.e., the benefit from not paying $t_\tau(\theta)$ to the seller minus the overall penalty for breach. With such scheme, everything happens thus as if the buyer was *de facto* residual claimant for all possible breaches that would then take place post-delivery.

Pooling the individual enforcement constraints to get (4.1) only provides a necessary feasibility condition. Reciprocally, suppose that a profile of outputs satisfies (4.1) at all

[26] It is worth noticing at this stage that the same feasibility constraints are obtained whether penalties are liquidated damages paid by the breaching party to its partner, or instead are paid to a third party (or equivalently destroyed).

dates τ; then one can structure pre- and post-delivery prices so that both individual enforcement constraints (3.2) and (3.3) hold. An easy way of doing so consists in taking pre-delivery payments equal to the seller's penalty for breach:

$$t_{1,\tau}(\theta) = L \quad \forall\tau, \quad \forall\theta.$$

At the same time, the post-delivery price should extract the seller's current surplus:

$$t_{2,\tau}(\theta) = -L + \theta q_\tau(\theta) \quad \forall\tau, \quad \forall\theta.$$

For the rest of the analysis, we assume that L is small enough to ensure that this post-delivery price remains positive under all circumstances, i.e.,

$$L \leq \min_{\theta\in\Theta} \theta q(\theta).^{27} \tag{4.3}$$

The payments $(t_{1,\tau}(\theta), t_{2,\tau}(\theta))$ then ensure that (3.3) trivially holds since:

$$\overline{U}_\tau(\theta) = (1-\delta)(t_{1,\tau}(\theta) - L) = 0 \quad \forall\tau.$$

Summarizing our earlier discussion, next proposition (whose proof is obvious) draws some immediate but important implications from (4.1).

Proposition 1 *Everything 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:*

$$M = K + L.$$

The set of enforceable-allocations (and thus the optimal contract under complete information) does not depend on the distribution of remedies for breach but only on M.*

Anticipating our future findings, the same kind of invariance will apply under asymmetric information as long as L is not too large. There also, an enforcement-surplus condition will be obtained by pooling the buyer's and the seller's enforcement constraints. In doing so, we shall nevertheless take care of incentive compatibility.

Enforceable* stationary contracts. Suppose *a priori* that the buyer offers a stationary output profile. The aggregate enforcement constraints (4.1) are all identical at all $\tau \geq 0$ and can be written as:

$$\psi(\theta, q(\theta)) \geq 0 \quad \forall\theta \in \Theta. \tag{4.4}$$

The enforcement surplus becomes:

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

[27] Had (4.3) not been satisfied, the seller could be asked to reimburse the buyer after delivery (i.e., $t_{2,\tau}(\theta) = -L + \theta q_\tau(\theta) < 0$ for some θ) so that the contract still fully extracts his surplus. This makes not paying back an attractive possibility for the seller. When L is too small, this opportunistic behavior could be simply controlled by imposing that post-delivery payments remain non-negative. We leave the corresponding analysis of the optimal contract with this additional constraint for future research. For the time being, we observe that (4.3) will be satisfied by the optimal outputs derived in Theorem 1 below when $L \leq \min_{\theta\in\Theta} \theta q^e(\theta)$ where $q^e(\theta)$ is the highest enforceable output under complete information whose definition is also given below. This gives us an upper bound on L that depends only on exogenous parameters.

Observe that ψ is strictly concave in q and zero at some $q^e(\theta)$. Since the buyer's gross surplus from trade is delayed to the next period in the definition of the enforcement surplus, ψ achieves its maximum at an inefficient low output $q^{db}(\theta) < q^{fb}(\theta)$[28] ($q^{db}(\theta)$ is defined through a first-order condition as $\delta S'(q^{db}(\theta)) = \theta$). It also immediately follows from ψ concavity in q that $q^e(\theta) > q^{db}(\theta)$.

THE OPTIMAL CONTRACT. Under complete information, the buyer's objective is to maximize the discounted net surplus that she withdraws from trade, subject to the seller's individual rationality, and the aggregate enforcement constraint that altogether characterize feasible trades. Formally, this problem can be stated as:

$$(\mathcal{P}^*): \qquad \max_{(\mathbf{q}(\theta),\overline{U}_0(\theta))} (1-\delta)\sum_{\tau=0}^{+\infty}\delta^\tau(S(q_\tau(\theta))-\theta q_\tau(\theta))-\overline{U}_0(\theta) \text{ subject to (3.1) and (4.1).}$$

The solution to this problem depends on whether enforcement constraints are binding or not.

THEOREM 1 *Under complete information, the optimal contract is stationary but not always efficient:*

$$q_\tau^{sb}(\theta) = \begin{cases} q^{fb}(\theta) & \text{if } \psi(\theta, q^{fb}(\theta)) \geq 0 \ , \\ q^e(\theta) < q^{fb}(\theta) & \text{otherwise.} \end{cases}$$

This Theorem deserves some comments. Some are directly related to the economic intuition behind the existing distortions if any. Others are more technical but still deeply connected to the economics of the problem.

Turning first to intuition, observe that when the aggregate enforcement constraint (4.1) always holds, trade is efficient in every period and the seller's surplus is fully extracted. Of course, this outcome is stationary. This efficient scenario arises when it is easy to prevent breaches. This can be because the cost of producing is small and the seller is not tempted by not delivering. This might also be because the buyer's benefit is large and pursuing trade in the future is attractive. To illustrate, observe that, for a fixed level of public enforcement M, the inequality $q^e(\theta) \geq q^{fb}(\theta)$ always holds when δ is close enough to 1. Enforcement is not an issue if parties value sufficiently the future even with finite remedies. The private side of enforcement suffices.

The interesting case precisely arises for low discount factors and public remedies are also needed. Then, efficient trades are no longer enforceable. Enforcement constraints are binding at all dates. Yet, the fact that those constraints are *not directly time-dependent* and similar from one period to the other is enough to ensure stationarity of the optimal contract. However, production must be reduced below the first best so that payments are also lowered and incentives for breaches less attractive on both sides of the transaction.

With inefficient trades, the stationarity of the optimal contract is much less obvious. Indeed, relaxing the aggregate enforcement constraint (4.1) calls for reducing current output so as the seller does not breach while at the same time increasing future trades to also control the buyer. There is *a priori* no reason for those two forces to be always of

[28] The superscript is meant for *delayed benefits*.

equal magnitude. For instance, the buyer could choose to implement a trade profile with outputs which are continuously increasing over time simply by promising always even more surplus in future rounds of trading. Avoiding such *roll-balling* strategy requires that the multiplier of the enforcement constraint (4.1) at any date τ does not grow too fast. Indeed, much of the technical works that is developed in the Appendix precisely consists in proving that such exploding dynamics do not arise.[29]

In a related context, Levin (2003) has ingeniously showed that there is no loss of generality in looking for an optimal contract within the class of stationary ones provided that an optimal contract exists in the first place. This qualifier is an important one. The Lagragean techniques that we use to prove Theorem 1 allow us to prove existence altogether with properties of the optimal contract. Moreover, this more direct approach turns out to be applicable also under asymmetric information in contexts where stationarity might be lost and no obvious class of contracts might appear attractive in the first place.

5. ASYMMETRIC INFORMATION: CHARACTERIZING ENFORCEABLE PROFILES

The seller's enforceability conditions (3.5) have now more stringent implications beyond the usual incentive compatibility constraints (3.6). Indeed, aggregating (3.2) and (3.5) gives us a new set of feasibility conditions that are described in next Lemma.

LEMMA 2 *An incentive compatible mechanism $\mathcal{C}$ is* ***enforceable*** *if and only if*

$$\Psi(\theta, \overline{\mathbf{q}}_\tau(\theta)) \geq \delta^{-\tau} \max_{\hat{\theta}\in\Theta} \left\{ \overline{U}_0(\hat{\theta}) - \overline{U}_0(\theta) + (\hat{\theta} - \theta)\underline{q}_\tau(\hat{\theta}) \right\}, \quad \forall\theta \in \Theta, \quad \forall\tau \geq 0. \tag{5.1}$$

To understand how those feasibility conditions are constructed, let us first come back on the seller's enforcement constraints. In the very same spirit as under complete information, having pre-delivery payments equal to the seller's remedy, i.e., $t_{1,\tau}(\theta) \equiv L$, ensures that the buyer becomes residual claimant for all breaches, at least as long as the seller behaves truthfully. Yet, the seller may behave so only if constraints (3.5) hold.

In this respect, it should be intuitive that the most salient enforcement constraints are those of a high-cost seller. 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 those rents over the whole relationship.

Instead, a *take-the-money-and-run* strategy is particularly attractive for a high-cost seller. That type may find it worth to behave as being more efficient and pocket the large payment that such low-cost type receives for the first τ periods before breaching.

To illustrate, suppose that the payment for the high-cost seller fully extracts his surplus in every period, i.e., $t_\tau(\overline{\theta}) = \overline{\theta} q_\tau(\overline{\theta})$. That simple scheme satisfies (3.5) if and only if

$$\overline{U}_0(\overline{\theta}) \geq \max\{\overline{U}_0(\underline{\theta}) - \Delta\theta \underline{q}_\tau(\underline{\theta}) - \delta^\tau(\overline{U}_\tau(\underline{\theta}) - (1-\delta)(t_{1,\tau}(\underline{\theta}) - L)); 0\}.$$

The first term on the righthand side is what can be grasped by a high-cost seller when behaving as a low-cost one for $\tau - 1$ periods. This strategy entails getting a high price

[29] From a technical perspective, it is interesting to observe that the constraints in $(\mathcal{P}^*)$ are forward looking; a feature often found in the macroeconomic literature. See Marcet and Marimon (2011) for references and Messner, Pavoni and Sleet (2013) for recent advances on that front.

but also producing a large quantity in any of those earlier periods. The second term is simply the zero surplus option that this high-cost seller secures from reporting truthfully.

Taking into account that $t_{1,\tau}(\theta) \equiv L$ and $\overline{U}_0(\underline{\theta}) = \Delta\theta\overline{q}_0(\overline{\theta}) \geq \overline{U}_0(\overline{\theta}) = 0$, we get:

$$\overline{U}_\tau(\underline{\theta}) \geq \delta^{-\tau} \max\{\Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_\tau(\underline{\theta})); 0\} \quad \forall \tau \geq 0. \tag{5.2}$$

To prevent the high-cost seller's *take-the-money-and-run* strategy, the buyer must push the low-cost seller's payments towards the future. Payments are back-loaded to facilitate enforcement. Of course, this *take-the-money-and-run* strategy only matters earlier on in the relationship. Had the high-cost seller misbehaved over the whole relationship, he would just produce too much, a behavior that would violate incentive compatibility. The beginning and the tail of the relationship do not look the same for the high-cost seller. This important feature explains the non-stationarity of the optimal contract.

Pooling the buyer's enforcement constraint (3.2) when trading with a low-cost seller with the high-cost seller's enforcement constraint (5.2) yields the following important aggregate enforcement constraint:

$$\Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta})) \geq \delta^{-\tau} \max\left\{\Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_\tau(\underline{\theta})), 0\right\}.^{30} \tag{5.3}$$

Equipped with the constraint on forward rents (5.2), we may reinterpret (5.3) as saying that the enforcement surplus must always cover the forward rent left to a low-cost seller. There are two specific cases that deserve more attention. First, observe that, since the enforcement constraints hold at all τ, we may as well take their limits as τ goes to infinity. Doing so shows that those conditions look the same as under complete information in the limit if the monotonic conditions (3.7) is strict because the *max* on the right-hand side is then zero. In other words, in the long run the consequences for contract design of asymmetric information and limited enforcement can be somewhat disentangled. Limited enforcement might thus only impact for at most a finite number of early periods.

Second, turning now to the case $\tau = 0$, constraint (5.3) becomes:

$$\Psi(\underline{\theta}, \overline{\mathbf{q}}_0(\underline{\theta})) \geq \Delta\theta\overline{q}_0(\overline{\theta}). \tag{5.4}$$

This condition is similar to (4.2) with the only difference being now that the first-period payment made to the low-cost seller has also to pay for the extra information rent that this type can get, namely $\Delta\theta\overline{q}_0(\overline{\theta})$. Asymmetric information hardens the aggregate enforceability constraint at date $\tau = 0$.

6. THE OPTIMALITY OF NON-STATIONARY CONTRACTS UNDER ASYMMETRIC INFORMATION

We now suppose that $q^{fb}(\theta) \leq q^e(\theta)$ for all θ. This assumption ensures that, under complete information, the enforcement surplus is always non-negative. From Theorem 1, trade is thus always efficient in that case. Focusing on such scenario isolates the most interesting contracting patterns that may now arise under asymmetric information.

[30] It is immediate to see that, whenever (5.3) is binding, the enforcement constraints (3.2) and (3.5) for the inefficient seller are both also binding.

THE OPTIMAL CONTRACT. The buyer's objective is now to maximize the discounted net surplus that she withdraws from trade, subject to the seller's individual rationality, incentive compatibility and the new aggregate enforcement constraints. Formally, we get:

$$(\mathcal{P}): \max_{(\mathbf{q}(\theta),\overline{U}_0(\theta))} E_\theta \left((1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \left(S(q_\tau(\theta)) - \theta q_\tau(\theta) \right) - \overline{U}_0(\theta) \right)$$

subject to (3.1), (3.6), and (5.1).

6.1. *Implementing the Baron-Myerson Outcome*

We first present conditions under which enforceability is not an issue. As a by-product, the optimal contract then turns out to be stationary.

THEOREM 2 *Suppose that*

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) \geq \Delta\theta q^{bm}(\overline{\theta}) \text{ and } \psi(\overline{\theta}, q^{bm}(\overline{\theta})) \geq 0. \tag{6.1}$$

An optimal contract is stationary and corresponds to the infinite replica of the Baron-Myerson outcome.

When there are enough gains from trade with a low-cost seller to prevent breaches by both parties, the Baron-Myerson allocation is enforceable at no extra cost. This is so even if the price received by this low-cost seller is greater than under complete information as a result of an extra information rent.

To give more intuition, observe that (6.1) implies that (5.3) holds at date 0 when all payments to the low-cost agent are paid upfront. The most attractive deviation for the buyer would then be not to pay that amount at date 0 and breach the contract right away but (6.1) ensures that such deviation is not valuable.

In the scenario depicted by Theorem 2, restricting the analysis to stationary contracts has no bite and the Baron-Myerson allocation is optimal within the broadest possible class of incentive mechanisms that can be envisioned.

6.2. *Optimal Stationary Contracts*

When (6.1) does not hold, the stationary Baron-Myerson outcome can no longer be implemented. Let us now *a contrario* assume that the information rent left to a low-cost seller in that Baron-Myerson scenario is too large to allow enforcement:

$$0 \leq \psi(\underline{\theta}, q^{fb}(\underline{\theta})) < \Delta\theta q^{bm}(\overline{\theta}) \text{ and } \psi(\overline{\theta}, q^{bm}(\overline{\theta})) \geq 0. \tag{6.2}$$

A first possibility of interest to satisfy the enforcement constraints is to choose a stationary contract. In that case, (5.3) amounts to:

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

For any output profile satisfying (3.7) this enforcement constraint holds at $\tau \geq 1$ if it already holds at $\tau = 0$. Intuitively, with a stationary contract, if any breach happens, it is not delayed and takes place right away. This leads to the simpler requirement:

$$\psi(\underline{\theta}, q(\underline{\theta})) \geq \Delta\theta q(\overline{\theta}). \tag{6.3}$$

Equipped with this feasibility condition, we can now obtain the following characterization.

PROPOSITION 2 *When condition (6.2) holds, the optimal stationary contract entails:*

1. *Downward output distortions below the Baron-Myerson outcome*

$$S'(q^{st}(\underline{\theta})) = \underline{\theta} + \frac{(1-\delta)\Lambda}{\nu + \Lambda\delta}\Delta\theta \tag{6.4}$$

$$S'(q^{st}(\overline{\theta})) = \overline{\theta} + \frac{\nu + \Lambda}{1-\nu}\Delta\theta \tag{6.5}$$

where $\Lambda > 0$ is the Lagrange multiplier of the binding enforcement constraint (6.3);

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

$$\overline{U}_0^{st}(\underline{\theta}) = \Delta\theta q^{st}(\overline{\theta}) > 0 = \overline{U}_0^{st}(\overline{\theta}).$$

The Lagrange multiplier Λ provides an indirect measure of the quality of the judicial system, i.e., the fact that the aggregate remedies M is not enough to ensure enforceability. Because that multiplier impacts on screening distortions, the quality of the judicial system and contract performances are now linked altogether.

With stationary contracts, an imperfect enforcement impacts on the seller's output for any realization of his cost. Both outputs are reduced as enforcement problems become more acute. This result will change when non-stationary contracts are considered below.

The explanation behind those output distortions is simple. To relax the binding enforceability constraint (6.3), the buyer would like to reduce the price paid to a low-cost seller so as to make breaches less attractive. To do so, two instruments are used in tandem. First, the buyer may commit to buy even less from a high-cost seller than in the Baron-Myerson scenario. This distortion reduces the information rent of the low-cost seller. Second, the buyer can also ask for less output from a low-cost seller.

With stationary contracts, reducing the low-cost seller's output has nevertheless two conflicting effects. First, it reduces the benefits of a current breach which definitively relaxes the enforceability constraint (6.3). Second, it reduces also the surplus in future trading rounds which is, on the contrary, detrimental for enforceability. Those two conflicting roles can be disentangled with non-stationary contracts as we will see below.

6.3. *Non-Stationary Optimal Contract*

To prepare for a characterization of the dynamics of optimal contracts, we need to introduce some notations and some technical conditions. In this respect, we first suppose that the following inequality holds:

$$0 \leq \psi(\underline{\theta}, q^{fb}(\underline{\theta})) < \Delta\theta q^{bm}(\overline{\theta}) \leq \psi(\overline{\theta}, q^{bm}(\overline{\theta})).^{31} \tag{6.6}$$

This assumption strengthens condition (6.2). It ensures that only the enforcement constraint (5.3) related to the enforcement surplus with a low-cost type might be binding at the optimum.[32]

Condition (6.6) implies that $q^{bm}(\overline{\theta}) < q^e(\overline{\theta})$ and $q^e(\overline{\theta}) < q^{fb}(\underline{\theta})$. Let us thus start from any arbitrary point $q_0(\underline{\theta}) \in \left[q^e(\overline{\theta}), q^{fb}(\underline{\theta})\right]$ to construct recursively the sequence $\mathbf{q}(\underline{\theta})$ as:

$$\begin{cases} q_0(\underline{\theta}) \in \left[q^e(\overline{\theta}), q^{fb}(\underline{\theta})\right] \\ q_{\tau+1}(\underline{\theta}) = \Phi(q_\tau(\underline{\theta})) \end{cases} \tag{6.7}$$

[32] This condition is more likely to be satisfied when $q^{bm}(\overline{\theta})$ is itself small, which arises for example when the probability of having a high-cost seller is low enough.

where the function $\Phi(q) = S^{-1}\left(\frac{1}{\delta}\left(\overline{\theta}q - (1-\delta)M\right)\right)$ is defined over the interval $\left[\frac{(1-\delta)M}{\overline{\theta}}, +\infty\right)$. Observe that Φ is increasing and convex. Because $\Phi\left(\frac{(1-\delta)M}{\overline{\theta}}\right) = 0$ and $\Phi(q^{bm}(\overline{\theta})) > q^{bm}(\overline{\theta})$ when $q^{bm}(\overline{\theta}) < q^{e}(\overline{\theta})$, there exists a unique fixed point to Φ within the open interval $\left(0, q^{bm}(\overline{\theta})\right)$. This fixed point is $q^{e}(\overline{\theta})$. 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.

Theorem 3 below shows that the dynamic pattern of trading with a low-cost seller entails two distinct phases. In the first early periods, his output continuously increases while remaining inefficiently low. The fact that output is distorted with both types over this transitory phase bears some similarity with the contracting pattern obtained when the analysis is *a priori* restricted to stationary contracts. In the mature phase of contracting that follows, trade with a low-cost seller becomes efficient and the sole distortion concerns the high-cost seller's output. Yet this distortion remains constant over time. The optimal contract in the long run exhibits features that come close 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 3 *Suppose that condition (6.6) holds. There almost always*[33] *exists* $\tau^* \geq 1$ *such that the optimal contract passes through two different phases.*

1. EARLY PHASE OF CONTRACTING WITH GROWING OUTPUT. *For* $\tau \leq \tau^*$, *the optimal output* $q^{sb}_{\tau}(\underline{\theta})$ *of the low-cost seller strictly increases over time but remains inefficient:*

   $$q^{e}(\overline{\theta}) < q^{sb}_{\tau}(\underline{\theta}) < q^{fb}(\underline{\theta}). \tag{6.8}$$

   The enforcement constraint (5.3) is binding at all dates $\tau \leq \tau^*$ *and the sequence* $q^{sb}_{\tau}(\underline{\theta})$ *obeys the recursive condition (6.7).*

2. MATURE PHASE OF CONTRACTING WITH STATIONARY OUTPUT. *For* $\tau > \tau^*$, *the optimal output* $q^{sb}_{\tau}(\underline{\theta})$ *of a low-cost seller is equal to its first-best level:*

   $$q^{sb}_{\tau}(\underline{\theta}) = q^{fb}(\underline{\theta}). \tag{6.9}$$

   The enforcement constraint (5.3) is slack at all those dates.

3. *At all dates* τ, *the high-cost seller produces the same quantity* $q^{sb}(\overline{\theta})$ *which remains below the Baron-Myerson level:*

   $$S'(q^{sb}(\overline{\theta})) = \overline{\theta} + \frac{\nu + \Lambda_{\infty}}{1-\nu}\Delta\theta \tag{6.10}$$

   for some $\Lambda_{\infty} > 0$.

[33] The qualifier *"almost always"* in the statement of Theorem 3 comes from the fact that the buyer may sometimes be indifferent between keeping the enforcement constraint binding till date $\tau^* - 1$ and till date τ^*.

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

$$\overline{U}_0^{sb}(\underline{\theta}) = \Delta\theta q^{sb}(\overline{\theta}) > 0 = \overline{U}_0^{sb}(\overline{\theta}).$$

The buyer is somewhat torn between two objectives: first, making contracts easier to enforce and second implementing an allocation close to the Baron-Myerson outcome which would maximize gains from trade if enforcement was costless.

There are two ways to relax enforceability constraints. The first one consists in increasing future surpluses. This is obtained by increasing the low-cost seller's output towards its first-best level.

With non-stationary contracts, the buyer can also backload the payments needed to induce information revelation from a low-cost seller while, at the same time, ensuring that the *take-the-money-and-run* strategy of high-cost seller is not a threat. Pushing payments towards the future again relaxes enforceability constraints till output distortions with a low-cost seller are no longer needed. In this mature phase of contracting, the contract becomes stationary. However, and 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 as time passes contrary to what arises with stationary contracts.[34]

The parameter Λ_∞ is now a cumulative of all Lagrange multipliers for all the enforcement constraints that bind in earlier periods. It keeps track of all distortions due to enforcement over the whole contract duration. Yet, this parameter again links altogether the nature of the screening distortions and the quality of the legal system. Remarkably, the output distortion for a high-cost seller obeys a *modified Baron-Myerson* formula (6.10) that shows how virtual costs must now be modified with limited enforcement. Lower remedies means that payments and outputs should be reduced to facilitate enforcement. A greater value of Λ_∞ comes thus with more output distortions. Therefore, we expect greater output distortions and back-loaded payments when enforcement is more difficult.

Implementation. The fact that the aggregate enforcement constraint at a given date binds puts quite a bit of structure on the intertemporal profile of payments. This stands in sharp contrast with the case of costless enforcement (Baron and Besanko, 1984) since then only the values of the overall intertemporal payments to both types are known from the binding incentive and participation constraints.[35] Next Proposition exhibits payments that implement the optimal allocation described in Theorem 3.

Proposition 3 *Suppose that condition (6.6) holds.*

1. *Pre-delivery payments always cover the seller's penalty for breach:*

$$t_{1,\tau}^{sb}(\theta) = L \quad \forall\theta \in \Theta, \quad \forall\tau \geq 0. \tag{6.11}$$

[34]The fact that the optimal contract is not stationary implies that, at the beginning of the relationship, the buyer would like to commit to a trades pattern which is no longer optimal as times passes. Yet, in this model with persistent types, leaving the possibility of renegotiating the continuation of the contract and re-optimizing trading profiles from a given date on would modify the seller's incentives to reveal his type earlier on, introducing all well-known difficulties from the renegotiation literature.

[35]Those payments can be chosen to be stationary as shown above but there is much leeway beyond.

2. *the high-cost seller's payment can be taken to be stationary:*

$$t_\tau^{sb}(\overline{\theta}) = \overline{\theta} q^{sb}(\overline{\theta}) \quad \forall \tau \geq 0. \tag{6.12}$$

3. *the low-cost seller's payment is increasing over the early phase of contracting*

$$t_\tau^{sb}(\underline{\theta}) = \overline{\theta} q_\tau^{sb}(\underline{\theta}) \quad \forall \tau \leq \tau^* \tag{6.13}$$

and can be taken to be stationary over the mature phase

$$t_\tau^{sb}(\underline{\theta}) = \underline{\theta} q^{fb}(\underline{\theta}) + \delta^{-\tau^*} \Delta\theta (q^{sb}(\overline{\theta}) - \underline{q}_{\tau^*}^{sb}(\underline{\theta})) \quad \forall \tau > \tau^*. \tag{6.14}$$

The pre-delivery payments in (6.11) just make the buyer residual claimant for all penalties for breach. The overall payments (6.12) fully extract the high-cost seller's rent in each period. Finally, Theorem 3 has shown that the enforcement constraints (5.3) are binding at all dates $\tau \leq \tau^*$. An immediate corollary is that the inefficient seller's enforcement constraints (3.5) are then also binding. The high-cost seller's *take-the-money-and-run* strategy is thus prevented by keeping this type just indifferent between pretending being more efficient and getting no profit in each period.[36]

The final step in constructing payments consists in ensuring that the stationary payment in (6.14) offered over the mature phase together with the prices over the growing phase described (6.13) just ensure that the low-cost seller receives enough rent to truthfully reveal his type at date 0.

Finally, the specification of those payments allows us to find an upper bound on L so that the post-delivery payment $t_{2\tau}(\theta)$ remains non-negative as posited, namely:

$$L \leq \min\{\overline{\theta} q^{sb}(\overline{\theta}), \underline{\theta} q^{fb}(\underline{\theta}) + \delta^{-\tau^*} \Delta\theta (q^{sb}(\overline{\theta}) - \underline{q}_{\tau^*}^{sb}(\underline{\theta}))\}.$$

Length of the growing phase. The length of the first phase trades off two competing effects. First, increasing the number of periods where the enforcement constraints (5.3) are binding smoothes the cost of the *take-the-money-and-run* strategy. However doing so might unnecessarily delay efficient trade with a low-cost seller. We now provide a simple upper bound on the length of this earlier phase, which is computed from fundamentals only.

Proposition 4 *Suppose that condition (6.6) holds. An upper bound on the number of periods in the growing phase of contracting is:*

$$int\left(-\frac{ln\left(\frac{\psi(\overline{\theta}, q^{fb}(\underline{\theta}))}{\Delta\theta(q^{bm}(\overline{\theta}) - q^e(\overline{\theta}))}\right)}{ln(\delta)}\right) + 1. \tag{6.15}$$

[36] This feature is reminiscent of the countervailing incentives literature (Lewis and Sappington, 1989). In our context, countervailing incentives apply to different kinds of constraints (incentive and enforcement constraints) and full separation remains possible.

When δ goes to 0, only the public side of enforcement matters; future surplus is of no help to prevent opportunism. In that limiting case, the numerator of (6.15) remains bounded while the denominator goes to infinity so that the upper bound is equal to one. Since condition (6.6) also implies that the enforcement constraint is necessarily binding, there is precisely a single period, date $\tau = 0$, where the enforcement constraint (5.3) is binding when δ is small enough. The enforcement constraint (5.3) at date 0 then almost boils down to a simple constraint on output and forward rent at this date, namely:

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

This (binding) enforcement constraint is actually the same as (6.3) (taken for $\delta = 0$) which arises when the buyer is forced to use a stationary contract over infinitely many periods. Output distortions are identical. Intuitively, when the future no longer matters, the relationship boils down to a static relationship with the sole possibilities for enforcement coming from public remedies.[37]

Substitutability between public and private enforcements. Getting general comparative statics on the length of the growing phase described in Theorem 3 is not easy. Nevertheless, to get simple insights on how the length of this phase varies, we now ask which variations of fundamentals keep it constant.

To simplify, consider the case where the parameters M and δ are such that there is only one round in the growing phase (i.e., the enforcement constraint (5.3) is only binding at date 0). Starting from such constellation of parameters, suppose that the continuation of the relationship matters marginally more for the partners (i.e., δ increases). Certainly, a first impact of relying more on the private side of enforcement is to increase the enforcement surplus with an efficient seller $\psi(\underline{\theta}, q^{fb}(\underline{\theta})) = \delta S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta}) + (1-\delta)M$ (provided that $S(q^{fb}(\underline{\theta})) > M$). This of course relaxes the binding enforcement constraint at date 0 and calls for lower distortions in the low-cost seller's production. On the other hand, increasing δ makes it also more attractive to adopt a *take-the-money-and-run* strategy and breach contract at date 1 for the high-cost seller. This hardens the enforcement constraint (5.3) at date 1.

Next proposition elaborates on a joint change of M and δ that offers unambiguous conclusions on the length of the growing phase.

Proposition 5 *Suppose that condition (6.6) holds and that the earlier stage with growing output lasts only one period (i.e., $q_0(\underline{\theta}) < q_\tau(\underline{\theta}) = q^{fb}(\underline{\theta})$ for all $\tau \geq 1$). Increasing δ while keeping $\psi(\underline{\theta}, q^{fb}(\underline{\theta}))$ fixed maintains the length of the growing phase of contracting at only one round.*

To keep $\psi(\underline{\theta}, q^{fb}(\underline{\theta}))$ fixed, an increase in δ must be compensated by a simultaneous decrease in M, i.e., by reducing the importance public enforcement. Proposition 5 highlights some form of substitutability between the two kinds of enforcement.

7. Robustness

This section discusses various extensions of our basic framework.

[37] When instead δ goes to 1 so as the future matters a lot in the relationship and the private side of enforcement more significant, the upper bound in (6.15) gives us much less information. As already noticed, this is precisely under such scenario that enforcement is no longer an issue.

7.1. *More General Results*

The strength of condition (6.6) was to ensure that only the enforcement constraint in state $\underline{\theta}$ matters for contract dynamics. Yet, even when the weaker condition (6.2) only holds and the enforcement constraint might bind also in state $\overline{\theta}$, much can still be said on trading patterns, at least in their long run limits. Indeed, Theorem 4 below demonstrates that efficiency and extra distortions beyond the Baron-Myerson level remain important features of contract dynamics.

THEOREM 4 *Suppose that condition (6.2) holds. The optimal contract has always the following properties in the long run.*

1. *Trade with a low-cost seller is always efficient:*

$$\lim_{\tau\to+\infty} S'(q_\tau(\underline{\theta})) = \underline{\theta}. \tag{7.1}$$

2. *Trade with a high-cost seller is lower than in the Baron-Myerson outcome:*

$$\lim_{\tau\to+\infty} q_\tau(\overline{\theta}) = q^{sb}(\overline{\theta}) < q^{bm}(\overline{\theta}) \tag{7.2}$$

7.2. *Continuum of Types*

In our analysis so far, types were discrete while the optimal output could take values in a continuum. This flexibility allowed us to fully describe the pattern of increasing trades over the growing phase of contracting with much details. In this section, we entertain the possibility that types are continuous and investigate the robustness of our previous findings. To simplify the analysis and avoid unnecessary technicalities, let us assume that θ is uniformly distributed on $\Theta = \left[\underline{\theta}, \overline{\theta}\right]$.

For future references, let us also define an output profile $q^\Psi(\theta)$ and a cut-off θ^Ψ such that

$$\begin{cases} S'(q^\Psi(\theta)) = 2\theta - \underline{\theta} + \Psi & \text{if } \theta \le \theta^\Psi \\ q^\Psi(\theta) = 0 & \text{otherwise} \end{cases} \tag{7.3}$$

where θ^Ψ solves $S'(0) = 2\theta^\Psi - \underline{\theta} + \Psi$ if such solution exists in Θ and $\theta^\Psi = \overline{\theta}$ otherwise.

The optimal static contract is characterized by the usual Baron-Myerson distortion $q^{bm}(\theta) \equiv q^0(\theta)$ and a positive rent for the seller (at least for those types $\theta \le \theta^{bm} \equiv \theta^0$) given by $U^{bm}(\theta) = \int_\theta^{\theta^{bm}} q^{bm}(x)dx = t^{bm}(\theta) - \theta q^{bm}(\theta)$. Inefficient types $\theta \ge \theta^{bm}$ do not produce and get thus no rent. Again, in this framework, the buyer's marginal surplus must equal the seller's *virtual cost* which exceeds true cost under asymmetric information since indeed $2\theta - \underline{\theta} \ge \theta$. The corresponding downward output distortion helps reducing the costly information rent left to the seller.

CHARACTERIZING ENFORCEABLE ALLOCATIONS. Working with a continuum of types allows us to get an elegant characterization of aggregate enforcement constraints.

LEMMA 3 *An incentive compatible mechanism $\mathcal{C}$ is enforceable if and only if*

$$\Psi(\theta, \overline{\mathbf{q}}_\tau(\theta)) \ge \delta^{-\tau} \max_{\hat{\theta}\in\Theta} \left\{ \int_\theta^{\hat{\theta}} \overline{q}_0(x)dx + (\theta - \hat{\theta})\underline{q}_\tau(\theta) \right\}, \quad \forall\theta \in \Theta, \quad \tau \ge 0. \tag{7.4}$$

Clearly, the righthand side above is always strictly positive, at least at any $\theta < \overline{\theta}$. The enforcement constraints are thus always harder to satisfy under asymmetric information.

Implementability of the Baron-Myerson allocation. Mimicking our analysis of the discrete case, we first highlight conditions under which enforceability is not an issue. The restriction to stationary contracts has again no bite in such contexts and the Baron-Myerson allocation is thus optimal within the broadest possible class of mechanisms.

Theorem 5 *Suppose that*

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) \geq \int_{\underline{\theta}}^{\overline{\theta}} q^{bm}(x)dx \text{ and } q^{bm}(\theta) \geq q^{db}(\theta) \quad \forall \theta \in \Theta. \tag{7.5}$$ [38]

The optimal contract is stationary and implements in each period the Baron-Myerson allocation $(U^{bm}(\theta), q^{bm}(\theta))$.

The conditions (7.5) ensure that, if there is enough enforcement surplus with the most efficient type to cover his information rent then, enforcement constraints are automatically satisfied for less efficient types. Even though the enforcement surplus is lower with those types, it decreases less quickly than the corresponding information rent.

Optimal Stationary Contract. Lemma 3 admits a simpler expression when the analysis is restricted to stationary contracts.

Lemma 4 *A stationary incentive compatible mechanism* $\mathcal{C}$ *is enforceable if and only if:*

$$\psi(\theta, q(\theta)) \geq U(\theta) = \int_{\theta}^{\overline{\theta}} q(x)dx \quad \forall \theta \in \Theta. \tag{7.6}$$

From now on, we shall thus assume that the Baron-Myerson allocation is not always enforceable. Under that scenario, distortions arise to relax the enforcement constraints.

Proposition 6 *Suppose that*

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) < \int_{\underline{\theta}}^{\overline{\theta}} q^{bm}(x)dx. \tag{7.7}$$

The optimal continuous stationary contract entails a non-increasing output schedule[39] $q^{st}(\theta)$ *with a non-trivial interval of bunching on* $\left[\underline{\theta}, \hat{\theta}\right]$ *(with* $\hat{\theta} > \underline{\theta}$*) over which the enforcement constraint (7.6) is binding:*

$$q^{st}(\theta) = \begin{cases} q^{\Psi}(\hat{\theta}) & \text{if } \theta \leq \hat{\theta}, \\ q^{\Psi}(\theta) & \text{if } \theta \geq \hat{\theta} \end{cases} \tag{7.8}$$

[38] Observe that $q^{db}(\theta)$ is small enough when δ is also small so that the second part of (7.5) trivially holds in that case.

[39] It can be shown that the optimal stationary contract is indeed continuous.

for some $\Psi > 0$ *provided that* $q^{\Psi}(\hat{\theta}) \geq q^{db}(\theta)$ *for all* $\theta \in \Theta$.[40]

To understand the downward distortion, observe that the enforcement constraint (7.6) is relaxed by moving the output closer downward to the value $q^{db}(\theta)$ that maximizes the enforcement surplus and thus the lefthand side of this constraint. At the same time, such downward reduction of output reduces the seller's information rent, decreases his payments and eases enforcement as well; a second effect that impacts on the righthand side of (7.6).

Those two effects are familiar from our analysis of the discrete case. However, with a continuum of types, a new phenomenon arises: bunching on the lower tail of the types distribution. To understand its nature, it is easier to think of the case where δ is close to zero and only public remedies can be used. Then, the enforcement constraint (7.6) with a stationary contract amounts to a simple upper bound on payments:

$$M \geq t(\theta) = U(\theta) + \theta q(\theta), \quad \forall \theta \in \Theta. \tag{7.9}$$

Incentive compatibility ensures that payments are increasing so that, whenever (7.9) starts being binding it is also so for all most efficient types.[41,42] The only possibility is then to offer a constant price and output for those most efficient types. By incentive compatibility, compressing payments for those most efficient types also requires further downward output distortions for least efficient ones. The output requested from those types is everywhere distorted downwards below the Baron-Myerson solution.

Non-Stationary Contracts Are Again Optimal. Finding the optimal mechanism with a continuum of non-stationary enforcement constraints (7.4) that apply at all dates is a daunting task. Yet, the characterization of the optimal stationary contract is simple enough so that we can easily exhibit a valuable non-stationary deviation.

Consider indeed a contract that differs from the optimal stationary contract characterized in Proposition 6 only at date 0 where the output is:

$$q_0(\theta) = q^{st}(\theta) - \varepsilon \tag{7.10}$$

[40]Take M small enough so that the Baron-Myerson allocation is just implementable, i.e.,

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) = \int_{\underline{\theta}}^{\overline{\theta}} q^{bm}(x)dx \text{ and } q^{bm}(\theta) \geq q^{db}(\theta) \quad \forall \theta \in \Theta.$$

Starting from this benchmark and slightly reducing M, the enforcement constraint (7.6) is only binding on a small interval around $\underline{\theta}$ and Ψ is small enough. Then the condition $q^{\Psi}(\hat{\theta}) \geq q^{db}(\theta)$ also holds.

[41] Thomas (2002) and Gautier and Mitra (2006) have also studied screening problems with exogenous bounds on payments. The constraint in payments coming here from imperfect enforcement constraint is by construction endogenous. The optimal contract with an exogenous constraint nevertheless shares some features with the optimal contract here; namely bunching at the top and a modified Baron-Myerson downward distortion of production for least efficient types. Levin (2003) also stresses the existence of bunching on the lower tail of the cost distributions for the optimal relational contract. Although, an optimal stationary contract in our framework also exhibits this feature, technical details differ. First, when focusing on i.i.d. types as in Levin (2003), the future surplus that facilitates enforcement is independent on current types so that distortions on current outputs have no impact on future surplus. Second, there is no asymmetric information at the inception of the relationship and rent extraction is not an issue in Levin (2003)'s framework.

[42]The non-responsiveness of the contract (Guesnerie and Laffont, 1984) on the lower tail of the cost distribution can be interpreted as an *endogenous contract incompleteness*. Not all contingencies can be contracted upon as a result of the enforcement problem.

for some $\varepsilon \geq 0$ small enough.

With stationarity, the output $q^{st}(\theta)$ plays two competing roles as we already stressed. On the one hand, increasing this output towards the first best increases future gains from trade which eases enforcement. On the other hand, decreasing this output also reduces the seller's current rent and payment which also facilitates enforcement. The non-stationary deviation (7.10) allows at the same time to reduce date 0-output, decrease date 0-payment and facilitate enforcement at that date without jeopardizing future surplus.

THEOREM 6 *Assume that $q^{st}(\theta)$ is everywhere non-negative, the optimal contract is non-stationary.*[43]

7.3. *One-Sided Breach*

In our main analysis, bilateral opportunism was a key ingredient allowing us to pool enforcement constraints on each side of the relationship into a single condition. When opportunism is only on one side, pre- and post-delivery payments can be found which implement the Baron-Myerson at no extra cost.

To illustrate this point, we first consider the case where only the buyer may behave opportunistically. In each period, pre- and post-delivery payments can be constructed so as to induce the buyer to behave. More precisely, define

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

Post-delivery payments are such that the buyer is indifferent at each period between breaching, 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) hold at any date. Pre-delivery payments are such that the seller receives the corresponding Baron-Myerson payments.

Suppose now that only the seller might behave opportunistically. Suppose also, to harden the case, that he is privately informed. We want to construct a mechanism that implements the Baron-Myerson allocation and is seller-enforceable. To do so, it must be that (3.5) always holds at any date and for any type. By an argument which is by now familiar, the low-cost seller's enforcement constraint is not an issue if the stationary payment extracts the high-cost seller's surplus in each period, namely:

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

Enforcement constraints with a high-cost seller are thus satisfied with the following pre- and post-delivery payments:

$$t_{\tau}(\overline{\theta}) = L, \text{ and } t_{2,\tau}(\overline{\theta}) = \overline{\theta} q^{bm}(\overline{\theta}) - L \quad \forall \tau \geq 0.$$

Observe that post-delivery payments are here also non-negative if L is not too large.

More interesting maybe are transfers that prevent the *take-the-money-and-run* strategy of a high-cost seller. From (3.5), backward payoffs $\underline{U}_{\tau}(\underline{\theta})$ must then satisfy:

$$\overline{U}_0(\overline{\theta}) = 0 \geq \underline{U}_{\tau}(\underline{\theta}) - (1-\delta^{\tau})\Delta\theta q^{fb}(\underline{\theta}), \quad \forall \tau \geq 0.$$

[43] Building on this insight might suggest a recursive construction of the optimal non-stationary contract in the case of a continuum of types. The new contract, so constructed, could be pushed forwards from date 1 on and we could look for a date 0-optimal contract with this contract as a continuation; we could then iterate the process. The characterization of such optimal non-stationary contracts is very complex and left for future research.

Finding such payoffs is now easy. Define 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, we now take pre-delivery payments that make the high-cost seller just indifferent between breaching or not in each period:

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

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

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

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

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

Observe again that post-delivery payments are non-negative when L is again small enough. We can thus conclude:

THEOREM 7 *One-sided opportunism is costless.*

8. APPLICATIONS

Our framework is useful to address or revisit a number of important questions in organization theory.

Asset specificity and contract enforcement. The contractual setting entertained in this paper can be viewed as a stylized modeling of an ongoing relationship between a contractor and his long-term supplier for an essential input. The Transaction Costs Economics literature 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 on contract duration.[44] Yet, we are not aware of any empirical work that would have investigated how trading volumes evolve over time when parties are stuck in long-term bilateral contracts because of their earlier specific investments. Similarly, how differences in legal environments may impact on contract duration remains, to the best of our knowledge, an unexplored issue.

To shed some light on contract duration through the lens of our model, observe that the optimal contractual dynamics found above can also be viewed as a metaphor for how transactions go from spot contracts at their inception towards more mature relationships. Over the growing phase where enforcements constraints are binding, the relationship must build safeguards against bilateral opportunism. Instead, in the long run, mature relationships are no longer subject to such threat.

Even though specific investments are not present in our baseline model, our framework can be readily extended to link asset specificity and enforcement issues. 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 the trading surplus in each period τ by an amount $B(i)$ (with $B' > 0$ and $B'' < 0$):

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

[44] See Joskow (1987, 1988) and Crocker and Masten (1988, 1996) among many others.

Either under complete information, or even under asymmetric information but in the absence of any enforcement problem, the efficient level of investment i^{fb} thus satisfies:

$$B'(i^{fb}) = 1.$$ [45]

Now denote by $C(i)$ the buyer's opportunity costs for the foregone use of dedicated assets if the contract is breached (with $C' > 0$ and $C'' > 0$[46]). Thus, private costs also increase with asset specificity. Everything happens as if this extra cost was now counted as an implicit penalty for breach. Keeping the same expression for the enforcement surplus $\Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta}))$ as above, the enforcement constraint at date τ now is:

$$\Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta})) + \delta B(i) + (1-\delta)C(i) \geq \delta^{-\tau} \max\left\{\Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_\tau(\underline{\theta})), 0\right\}. \tag{8.1}$$

The optimal investment level i^e must now take into account its impact on contracts enforceability. There is of course a first direct effect in relaxing enforcement constraints and second, an indirect effect, that comes from changing optimal trade profiles as a result. By the Envelope Theorem, this indirect effect vanishes and i^e solves

$$(1 + \Lambda_\infty \delta)B'(i^e) + \Lambda_\infty(1-\delta)C'(i^e) = 1,$$

and, as a consequence, we 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 the cumulative Lagrange multiplier Λ_∞ is bigger, i.e., when enforcement is more difficult. Since we also know that output distortions are greater in that case, the model predicts that specific investments and outputs may covary negatively.

Relational contracting and firm's boundaries. Gibbons (2005a) argues 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 paper contributes to this important debate. Suppose that the informed seller may perform some specific investment i_s prior to contracting. That investment now improves the value of trade by increasing the probability $\nu(i_s)$ of being efficient. Following Riordan (1990), we may assume that ownership gives access to information. Consider thus a vertically integrated firm owned by the buyer. The seller has thus no information rent and no incentives to make any ex ante investment; $i_s = 0$. The seller becomes an employee of the firm and, at any point in time, this employee has the right to leave the firm.[47]

Under vertical separation, the seller owns the assets, retains private information and thus some expected informational rent worth $\nu(i_s)q^{sb}(\bar{\theta})$. This expected rent acts as an engine for investment. Of course, this rent also hardens the enforcement problem. Our model first predicts that *take-the-money-and-run* strategy will only arise in market relationships between vertically separate units. Second, market relationships come with greater volumes over time whereas within firms exchanges may exhibit more stable patterns through stationarity. Finally, output distortions being greater when enforcement problems are more acute, vertical integration becomes more attractive in weak enforcement contexts.

[45] Assuming, for simplicity, an interior solution.

[46] We assume that C'' is small enough so that s $\delta B(i) + (1-\delta)C(i)$ (on the righthand side of (8.1)) remains a concave function of i.

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

Another related interpretation of our findings is that reducing the cost of enforcing contracts requires the ability to backload payments. Such back-loading might be more easily achieved for transactions in thin markets than when firms can more easily found alternative partners. This case indeed amounts to an implicit reduction of the discount rate and thus makes private enforcement of contracts more difficult. In that sense also, our model predicts more vertical integration and long-term contracts in environments where enforcement is a key concern.

9. LITERATURE REVIEW

We end this paper by situating our findings with respect to the existing literature. Indeed, our paper merges issues found in various trends of the contracting literature.

Relational contracting. Basics. That a continuing relationship improves contracting even when contractual possibilities are limited is the basic tenet of an important literature on relational contracting (Bull 1987, MacLeod and Malcomson 1989 and 1998, Baker and al. 1994, among others).[48] There, the seller's effort (or output) is non-verifiable but remains observable. Short-term bonuses suffice to provide enough incentives provided the relationship is sufficiently repeated. This framework has been extended to study how relational contracts are modified when some extra variables relevant for trade might remain contractible.[49] In our context, prices and traded quantities are always contractible and parties can commit to a long-term contract. The only non-verifiable variables at any point in time in the game are the parties' decisions on whether to fulfill contractual obligations or not.[50] The contracting environment we consider is thus somewhere *in between* relational contracting and the full commitment/full enforcement scenario analyzed when complete contracts are feasible. As a result, while finding relational contracts is mostly a pure exercise in repeated game theory, our analysis still preserves the flavor and techniques of the mechanism design literature in dynamic environments with full commitment.[51] Yet, it modifies this framework by appending enforcement constraints.

Following Klein and Leffler (1981) and Shapiro (1983), a sizable share of the relational contracting literature analyzes durable goods. A recent advance on this front is Wolitzky (2010) who studies the dynamic incentives of a seller to cut prices and quantities. Keeping prices high acts as a commitment to deliver a promised quantity. Although there are many differences between his setting and ours (we allow public enforcement, non-atomistic sellers, and focus on the opposite allocation of bargaining power between buyers and sellers), we share a common interest for the design of dynamic trading procedures under limited enforcement.

[48] Malcomson (2012b) offers an exhaustive survey of this literature. Gibbons (2005b) draws the implication of this paradigm for organization theory.

[49] Bernheim and Whinston (1998), Schmidt and Schnitzer (1995), Iossa and Spagnolo (2009) and Ishihara (2013) among others contributed to that line of research.

[50] This quote from Malcomson (2012b) nicely summarizes why the literature on relational contracts makes no account of Courts: *"Relational contracts are concerned with agreements that can be enforced without resort to courts. The spirit of much of the theory discussed here is that, although an effective legal system exists, important elements of the relationship cannot be enforced legally because courts do not have the information to do that."* In this paper, we instead keep an active but imperfect legal system and assume that trade can be verified when it takes place. Yet, the issue remains of ensuring that it indeed takes place.

[51] See the seminal paper by Baron and Besanko (1984) and Battaglini (2005), Pavan, Segal and Toikka (2011) for most recent vintages.

Relational contracting. Stationarity or growing stakes. In an important paper, Levin (2003) characterizes optimal relational contracts when types are independently distributed over time, there is no asymmetric information at the time of contracting, and incentives to renege are evaluated before the realization of the productivity shock in any period. A major result of his analysis is that optimal relational contracts are stationary when payoff functions are quasi-linear. This stationarity does not survive with persistent types and asymmetric information at the time of contracting as in our framework. The enforcement constraints have now to be written for each cost realization and the possibility of adopting a *take-the-money-and-run* strategy makes the optimal contract non-stationary. This feature of contracting would remain even with any tiny serial correlation of types.

Other papers addressing how types persistence affects relational contracts include Horner (2002), Fong and Li (2010), Halac (2012), and Malcomson (2012a). Those authors are interested in how private information of the supplier diffuses over time in environments with no commitment. The focus is on the amount of pooling that arises endogenously in earlier periods of the relationship and how it determines non-stationary stakes.[52] Assuming instead full commitment, our mechanism design perspective allows us to address how types persistence is already enough to generate non-stationarity.

Finally, optimal non-stationary contracts in our framework and, especially their shape over the growing phase, also share features found in the reputation literature (Sobel 1985, Ghosh and Ray 1996, Watson 1999 and 2002). Relationships might start "small" there to ease reputations building when there is uncertainty on traders' opportunism.

Other modelings of enforcement. Beyond the literature on relational contracts, enforcement has also received some attention in models that stress the limited ability of Courts to enforce obligations. In that vein, Schwartz and Watson (2004), Doornik (2010) and Kvaløy and Olsen (2009) study the costs of writing contracts and its consequences on contract design. Those costs are nevertheless different from the enforcement costs highlighted above. Krasa and Villamil (2000) study the enforcement of financial contracts and show the optimality of debt contracts when Courts can enforce repayments. As pointed out by La Porta et al. (1998), ineffective enforcement of financial contracts may follow from corruption of the legal system or from existing laws that offer little protection to outside creditors attempting to enforce contracts. We differ from these papers both in terms of the definition of enforcement we are using and in terms of the role played by Courts.

Finally, Laffont and Martimort (2002, Chapter 9) and Guasch, Laffont and Straub (2003) in the more specific context of developing countries present static models with moral hazard ingredients that also discuss how incentive and enforcement constraints might interact.

REFERENCES

Abreu, D. (1988), "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica,* 56: 383-396.

Acemoglu, D. and J.A Robinson (2005), "Economic Origins of Dictatorship and Democracy," Cambridge University Press, Cambridge.

[52]The non-stationary of relational contracts may also come from learning persistent types as in a model of the labor market proposed by Yang (2012). Hemsley (2013) extends Levin (2003)'s model to the case of risk aversion and show that the optimal contract is non-stationary.

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

Baker, G., R. Gibbons and K. Murphy (1994), "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics,* 109: 1125-56.

Baron, D. and D. Besanko (1984), "Regulation and Information in a Continuing Relationship," *Information Economics and Policy,* 1: 267-302.

Baron, D. and R. Myerson (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.

Beaudry, P. and M. Poitevin (1994), "The Commitment Value of Contracts under Dynamic Renegotiation," *The RAND Journal of Economics,* 25: 501-517.

Bernheim, B. and M. Whinston (1998), "Incomplete Contracts and Strategic Ambiguity," *American Economic Review,* 88: 902-932.

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

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

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

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.

Farnsworth, E. A., (1982), *Contracts*, Little Brown Publishers.

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

Fukuyama, F. (2004), *State Building. Governance and World Order in the Twenty-First Century,* Cornell University Press, Cornell.

Gautier, A. and M. Mitra (2006), "Regulating a Monopolist with Limited Funds," *Economic Theory*, 27: 705-718.

Guasch, J. L., J.J. Laffont and S. Straub (2003), "Renegotiation of Concession Contracts in Latin America," (Vol. 3011), World Bank Publications.

Guesnerie, R. and J.J. Laffont (1984), "A Complete Solution to a Class of Principal-Agent Problems With an Application to the Control of Self-Managed Firm," *Journal of Public Economics*, 25: 329-369.

Gibbons, R. (2005a), "Incentives Between Firms (and Within)," *Management Science,* 51: 2-17.

Gibbons, R. (2005b), "Four Formal(izable) Theories of the Firm," *Journal of Economic Behavior and Organization,* 58: 200-245.

Ghosh, P., and D. Ray (1996), "Cooperation in Community Interaction without Information Flows," *Review of Economic Studies,* 63: 491-519.

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

Hart, O. (1995), *Firms, Contracts, and Financial Structure,* Oxford University Press, Oxford.

Hart, O. and B. Holmström (1987), "The Theory of Contracts," in *Advances in Economic*

Theory: Fifth World Congress, 75-106 (T. Bewley ed.).

Hart, O. and J. Tirole (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.

Horner, J. (2002), "Reputation and Competition," *American Economic Review*, 92: 644-663.

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

Ishihara, A. (2013), "Role of Informational Rents in Relational Contracts," Mimeo Kyoto University.

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.

Klein, B. and K. Leffler (1981), "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89: 615?641.

Krasa, S. and A. Villamil (2000), "Optimal Contracts When Enforcement is a Decision Variable," *Econometrica,* 68: 119-134.

Kvaløy, O. and T. Olsen (2009), "Endogenous Verifiability and Relational Contracting," *American Economic Review*, 99: 2193-2208.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R. Vishny (1998), "Law and Finance," *Journal of Political Economy,* 106: 1113-1155.

Laffont, J.J. and D. Martimort (2002), *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press.

Levin, J. (2003), "Relational Incentive Contracts," *American Economic Review,* 93: 835-857.

Lewis, T., and D. Sappington (1989), "Countervailing Incentives in Agency Problems," *Journal of Economic Theory*, 49: 294-313.

MacLeod, B. and J. Malcomson (1989), "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica,* 57: 447-480.

MacLeod, B. and J. Malcomson (1998), "Motivation and Markets," *American Economic Review,* 88: 388-411.

Mailath, G. and A. Postlewaite (1990), "Asymmetric Information Bargaining Problems with Many Agents," *Review of Economic Studies*, 57: 351-368.

Malcomson, J. (2012a), "Relational Incentive Contracts with Persistent Private Information," Mimeo University of Oxford, Department of Economics.

Malcomson, J. (2012b), "Relational Incentive Contracts," in *Handbook of Organizational Economics*, R. Gibbons and J. Roberts eds. Princeton University Press.

Marcet, A. and R. Marimon (2011), "Recursive Contracts," Mimeo University Pompeu Fabra.

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

Messner, M., N. Pavoni and C. Sleet (2013), "The Dual Approach to Recursive Optimization: Theory and Examples, " Mimeo Bocconi University.

Myerson, R. (1982), "Optimal Coordination Mechanisms in Generalized Principal-Agent Models,"*Journal of Mathematical Economics,* 10: 67-81.

Myerson, R. and M. Satterthwaite (1983), "Efficient Mechanisms for Bilateral Trading,"

Journal of Economic Theory, 29: 265-281.

Pavan, A., Segal, I., and J. Toikka (2011), "Dynamic Mechanism Design: Incentive Compatibility, Profit Maximization and Information Disclosure," mimeo, Stanford University.

Posner, E. A. (2011). *Contract Law and Theory.* Aspen Student Treatise Series. Wolters Kluwer Legal Publications.

Rey, P. and B. Salanié (1996), "On the Value of Commitment with Asymmetric Information," *Econometrica*, 64: 1395-1414.

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

Rustichini, A. (1998), "Lagrange Multipliers in Incentive-Constrained Problems," *Journal of Mathematical Economics,* 29: 365-380.

Schmidt, K. and M. Schnitzer (1995), "The Interaction of Explicit and Implicit Contracts," *Economics Letters*, 48: 193-199.

Shleifer, A. and R. Vishny (1997), "A Survey of Corporate Governance," *Journal of Finance,* 52: 737-753.

Seierstad, A. and K. Sydsaeter (1987), *Optimal Control Theory with Economic Applications*, North Holland, Amsterdam.

Schwartz, A. and J. Watson (2004), "The Law and Economics of Costly Contracting," *Journal of Law, Economics, and Organization*, 20: 2-31.

Shapiro, C. (1983), "Premiums for High Quality Products as Returns to Reputations," *Quarterly Journal of Economics,* 98: 659-679.

Sobel, J. (1985), "A Theory of Credibility," *Review of Economic Studies,* 52: 557-573.

Tilly, C. (2007), *Democracy,* Cambridge University Press, Cambridge.

Thomas, L. (2002), "Non-Linear Pricing with Budget Constraint," *Economics Letters*, 75: 257-263.

Watson, J. (2007), "Contract, Mechanism Design, and Technological Detail," *Econometrica*, 75: 55-81.

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.

Williamson, O. (1985), *The Economic Institutions of Capitalism*, Free Press, New-York.

Wolitzky, A. (2010), "Dynamic Monopoly with Relational Incentives," *Theoretical Economics,* 5: 479-518.

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

APPENDIX

Proof of Lemma 1: Necessity. We can rewrite (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:

(A.1)

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

We sum up (A.1) with (3.4) to get (4.1).

Sufficiency. Note first that $\overline{U}_0(\theta)=0$. If $\overline{U}_0(\theta)>0$ we can decrease $t_{1,0}(\theta)$ by $\varepsilon>0$ thereby improving the buyer's payoff without modifying other constraints. If the solution satisfies the aggregate constraint (4.1) we can define $\overline{U}_\tau(\theta)=0$ and $t_{1,\tau}(\theta)=L$ for all $\tau\geq 0$ then the seller's enforceability constraints (3.4) hold with equality and we obtain (A.1) from (4.1). *Q.E.D.*

Proof of Theorem 1: Theorem A.1 below provides a Lagrangian formulation for optimization problems with infinitely many constraints, an ingredient that requires careful use of duality arguments. We first prepare for the use of this theorem before formulating it and then applying it to our specific contexts.

Preliminaries. We allow for the possibility that the set of feasible outputs is not all $\mathbb{R}_+$, but an interval of the form $\mathcal{Q}=[0,\overline{\mathcal{Q}}]$. Denote then by $\mathbf{Q}_\infty$ the set of non-negative output sequences $\mathbf{q}=\{q_\tau\}_{\tau=0}^{\infty}$ such that $(1-\delta)\sum_{\tau=0}^{\infty}\delta^\tau|q_\tau|<\infty$. This is a closed convex set of the Banach space $l_1(\delta)$ of all sequences $\mathbf{x}$ such that $\|\mathbf{x}\|_\delta\equiv(1-\delta)\sum_{\tau=0}^{\infty}\delta^\tau|x_\tau|<\infty$. The topological dual of this space is l_∞, i.e., the space of bounded sequences $\mathbf{y}$ such that $\|\mathbf{y}\|_\infty\equiv\sup|y_\tau|<\infty$. Observe also that, for any $\mathbf{q}$ such that all components are uniformly bounded away from zero, the conical hull $C(\mathbf{q})=\{\lambda(\mathbf{Q}_\infty-\mathbf{q})\geq 0\}$ is equal to l_∞ whose closure is $l_1(\delta)$.

Properties of the maximand and constraints. The participation constraint (3.1) is binding at the optimum of the principal's problem, $\overline{U}_0(\theta)=0$. Taking this into account, we simplify the objective function accordingly. The maximand in the problem $(\mathcal{P}^*)$ so obtained is indeed a strictly concave and Fréchet differentiable function defined as:

$$f^*(\mathbf{q})=(1-\delta)\delta^\tau\left(\sum_{\tau=0}^{\infty}\delta^s(S(q_\tau)-\theta q_\tau)\right).$$

The Fréchet derivative at any $\mathbf{q}\in\mathbf{Q}^\infty$ is a linear and continuous operator $f^{*'}(\mathbf{q})$ such that, for any $\mathbf{h}$ such that $\mathbf{q}+\mathbf{h}\in\mathbf{Q}^\infty$:

$$f^{*'}(\mathbf{q})\mathbf{h}=(1-\delta)\delta^\tau\left(\sum_{\tau=0}^{\infty}\delta^s(S'(q_\tau)-\theta)h_\tau\right).$$

Turning now to the constrained set, the aggregate enforcement constraint (4.1) at date τ also defines a convex set since all functions

$$g_\tau^*(\mathbf{q})=(1-\delta)\delta^\tau\left(\delta\sum_{s=0}^{\infty}\delta^s(S(q_{\tau+s+1})-\theta q_{\tau+s+1})-\theta q_\tau+M\right)$$

are concave in $\mathbf{q}$ for all $\tau\geq 0$. Let denote by $g^*(\mathbf{q})=\{g_\tau^*(\mathbf{q})\}_{\tau\geq 0}$ the corresponding mapping from $l_1(\delta)$ onto itself. This mapping is Fréchet differentiable with a Fréchet derivative at any $\mathbf{q}\in\mathbf{Q}^\infty$ which is a linear operator $g^{*'}(\mathbf{q})$. Indeed consider such $\mathbf{q}$ and let $\mathbf{h}$ such that $\mathbf{q}+\mathbf{h}\in\mathbf{Q}_\infty$, we can easily compute:

$$g_\tau^{*'}(\mathbf{q})\mathbf{h}=(1-\delta)\delta^\tau\left(\delta\sum_{s=0}^{\infty}\delta^s(S'(q_{\tau+s+1})-\theta)h_{\tau+s+1}-\theta h_\tau\right).$$

Observe that $g_\tau^{*'}(\mathbf{q})$ is a bounded continuous operator onto $l_1(\delta)$ (and so is $g^{*'}(\mathbf{q})$). Indeed, we have:

$$\|g_\tau^{*'}(\mathbf{q})\mathbf{h}\|_\delta \leq (1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s \sup_q |S'(q)-\theta||h_{\tau+s+1}| + \overline{\theta}|h_\tau|\right)$$

$$\leq \max\{\sup_q |S'(q)-\theta|, |\overline{\theta}|\}(1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s |h_{\tau+s+1}| + |h_\tau|\right)$$

$$\leq \max\{\sup_q |S'(q)-\theta|, |\overline{\theta}|\} \left((1-\delta)\sum_{s=0}^{\infty} \delta^{\tau+s+1}|h_{\tau+s+1}| + (1-\delta)\delta^\tau |h_\tau|\right)$$

which finally gives us

$$\|g_\tau^{*'}(\mathbf{q})\mathbf{h}\|_\delta \leq N^* \|\mathbf{h}\|_\delta$$

where $N^* = 2\max\left\{\sup_q |S'(q)-\theta|, |\overline{\theta}|\right\} < +\infty$ from the fact that S' is bounded above over $\mathcal{Q}$ since S is concave and $S'(0) < +\infty$.

Formulation. With help of those notations, we can rewrite the maximization problem in a more compact form as:

$$(\mathcal{P}^*): \quad \max_{\mathbf{q}\in\mathbf{Q}_\infty} f^*(\mathbf{q}) \text{ subject to } g^*(\mathbf{q}) \geq 0.$$ [53]

We now define the Lagrangean for $(\mathcal{P}^*)$ as:

$$\mathcal{L}^*(\mathbf{q},\lambda) = f^*(\mathbf{q}) + \lambda g^*(\mathbf{q}) = (1-\delta)\sum_{\tau=0}^{\infty} \delta^\tau \left(S(q_\tau) - \theta q_\tau\right)$$

$$+ \sum_{\tau=0}^{\infty} \lambda_\tau \left((1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s (S(q_{\tau+s+1}) - \theta q_{\tau+s+1}) - \theta q_\tau + M\right)\right).$$

Next Theorem reminds an important result due to Rustichini (1998) that ensures the existence of a sequence of non-negative Lagrange multipliers $\lambda = \{\lambda_\tau\}_{\tau=0}^{\infty}$ that lies in $l_1(\delta)$ for the problem $(\mathcal{P}^*)$ so that an optimal solution $\mathbf{q}^*$ to $(\mathcal{P}^*)$ satisfies:

$$\mathcal{L}^*(\mathbf{q}^*,\lambda) \geq \mathcal{L}^*(\mathbf{q},\lambda) \quad \forall \mathbf{q} \in \mathbf{Q}^\infty.$$

THEOREM A.1 *(Rustichini, 1998, Theorem 6.6.) Take a sequence $\mathbf{q}^*$ in the interior of $\mathbf{Q}^\infty$ such that $g(\mathbf{q}^*) \geq 0$ which is an optimal solution to $(\mathcal{P}^*)$ with f and g Fréchet differentiable and concave, and suppose that the continuous operator $g'(\mathbf{q}^*)$ maps $l_1(\delta)$ onto itself. Then, there exists a non-negative Lagrange multiplier λ which is an element of $l_1(\delta)$ such that:*

(A.2) $\quad \lambda g^*(\mathbf{q}^*) = 0,$ [54]

(A.3) $\quad f^{*\prime}(\mathbf{q}^*) + \lambda g^{*'}(\mathbf{q}^*) = 0.$

[53] We adopt the standard convention that vector inequalities should be understood as being taken coordinate wise.

[54] 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 $\tau \geq 0$.

The fact that $g^{*\prime}(\mathbf{q}^*)$ maps $l_1(\delta)$ onto as shown above allows us to apply this theorem and conclude that Lagrange multipliers exist in $l_1(\delta)$

Optimization. The following optimality condition w.r.t. q_τ for $\tau \geq 0$ can be written as:

$$S'(q_\tau(\theta)) - \theta - \lambda_\tau \theta + \left(\sum_{s=0}^{\tau-1} \lambda_s\right)(S'(q_\tau(\theta)) - \theta) = 0. \tag{A.4}$$

This condition can be further simplified as:

$$S'(q_\tau(\theta)) = \theta + \frac{\lambda_\tau \theta}{1 + \sum_{s=0}^{\tau-1} \lambda_s}. \tag{A.5}$$

Case 1. $q^{fb}(\theta) \leq q^e(\theta)$. Then setting $\lambda_\tau = 0$ for all $\tau \geq 0$ yields $q_\tau(\theta) = q^{fb}(\theta)$ at all dates τ. Finally, (4.1) writes as

$$\Gamma(q^{fb}(\theta), \theta) \geq 0$$

which holds since $q^{fb}(\theta) \leq q^e(\underline{\theta})$. Of course, $\lambda = (0, 0, ..., 0, ...) \in l_1(\delta)$.

Case 2. $q^{fb}(\theta) > q^e(\theta)$. We look for a sequence of non-negative multipliers $\{\lambda_\tau\}_{\tau=0}^{\infty}$ in $l_1(\delta)$ which altogether with the stationary output $q_\tau(\theta) = q^e(\theta)$ at all dates τ implements the optimal solution. Because $q^{fb}(\theta) > q^e(\theta) > q^{db}(\theta)$, there exists $\lambda_0 > 0$ and $\lambda_0 < \frac{1}{\delta} - 1$ such that we can write:

$$S'(q^e(\theta)) = \theta(1 + \lambda_0). \tag{A.6}$$

More generally, consider thus the sequence $\lambda = \{\lambda_\tau\}_{\tau=0}^{\infty}$ defined recursively through the following condition inherited from (A.5);

$$\lambda_\tau = \left(\frac{S'(q^e(\theta))}{\theta} - 1\right)\left(1 + \sum_{s=0}^{\tau-1} \lambda_s\right). \tag{A.7}$$

From there, let us now define the sequence $\Lambda = \{\Lambda_\tau\}_{\tau=0}^{\infty}$ of cumulative multipliers as:

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

From (A.6), we have $\Lambda_1 = \lambda_0 > 0$. Moreover, (A.7) can be written as:

$$\Lambda_{\tau+1} + 1 = \frac{S'(q^e(\theta))}{\theta}(\Lambda_\tau + 1).$$

or equivalently:

$$\Lambda_\tau + 1 = \left(\frac{S'(q^e(\theta))}{\theta}\right)^{\tau-1}(\Lambda_1 + 1) = \left(\frac{S'(q^e(\theta))}{\theta}\right)^{\tau}.$$

It follows from this latter equation and the fact that $S'(q^e(\theta)) > \theta$ that Λ_τ is strictly increasing. This implies that the multiplier λ_τ is positive at all dates τ. Indeed, we have:

$$\lambda_\tau = \left(\frac{S'(q^e(\theta))}{\theta}\right)^{\tau}\left(\frac{S'(q^e(\theta))}{\theta} - 1\right).$$

Finally, $\lambda \in l_1(\delta)$ when $\frac{\delta S'(q^e(\theta))}{\theta} < 1$; a condition that holds since $q^e(\theta)) > q^{db}(\theta)$. *Q.E.D.*

Proof of Lemma 2: Observe that

$$\overline{U}_0(\hat{\theta}) = \underline{U}_\tau(\hat{\theta}) - \delta^\tau(1-\tau)(t_{1,\tau}(\hat{\theta}) - L) + \delta^\tau \overline{U}_\tau(\hat{\theta}) \quad \forall \hat{\theta} \in \Theta.$$

Using this condition, we rewrite (3.5) as:

$$\overline{U}_0(\theta) \geq \overline{U}_0(\hat{\theta}) + \delta^\tau(1-\tau)(t_{1,\tau}(\hat{\theta}) - L) - \delta^\tau \overline{U}_\tau(\hat{\theta}) + (\hat{\theta} - \theta)\underline{q}_\tau(\hat{\theta}) \quad \forall (\theta, \hat{\theta})^2 \in \Theta^2, \quad \forall \tau \geq 0.$$

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

$$\text{(A.8)} \quad \delta^\tau \left(\overline{U}_\tau(\theta) - (1-\tau)(t_{1,\tau}(\theta) - L)\right) \geq \overline{U}_0(\hat{\theta}) - \overline{U}_0(\theta) + (\theta - \hat{\theta})\underline{q}_\tau(\theta) \quad \forall (\theta, \hat{\theta})^2 \in \Theta^2, \quad \forall \tau \geq 0.$$

Let now multiply (A.1) by δ^τ to get:

(A.9)

$$(1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s (S(q_{\tau+s+1}(\theta)) - \theta q_{\tau+s+1}(\theta)) - \theta q_\tau(\theta) + M\right) \geq \delta^\tau \left(\overline{U}_\tau(\theta) - (1-\delta)(t_{1,\tau}(\theta) - L)\right).$$

Taken together (A.8) and (A.9) are compatible if and only if:

$$(1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s (S(q_{\tau+s+1}(\theta)) - \theta q_{\tau+s+1}(\theta)) - \theta q_\tau(\theta) + M\right) \geq \max_{\hat{\theta} \in \Theta} \{\overline{U}_0(\hat{\theta}) - \overline{U}_0(\theta) + (\theta - \hat{\theta})\underline{q}_\tau(\hat{\theta})\}$$

(where (3.6) holds) which can be rewritten as (5.1).

We rewrite Lemma 2 by developing those constraints as:

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

$$\text{(A.10)} \quad \Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta})) \geq \max\left\{\delta^{-\tau}\Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_\tau(\underline{\theta})), 0\right\},$$

$$\text{(A.11)} \quad \Psi(\overline{\theta}, \overline{\mathbf{q}}_\tau(\overline{\theta})) \geq \max\left\{\delta^{-\tau}\Delta\theta(\underline{q}_\tau(\overline{\theta}) - \overline{q}_0(\underline{\theta})), 0\right\},$$

(A.12)

$$\min\left\{\delta^\tau \Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta})) + \Delta\theta \underline{q}_\tau(\underline{\theta}); \Delta\theta \overline{q}_0(\underline{\theta})\right\} \geq \max\left\{-\delta^\tau \Psi(\overline{\theta}, \overline{\mathbf{q}}_\tau(\overline{\theta})) + \Delta\theta \underline{q}_\tau(\overline{\theta}); \Delta\theta \overline{q}_0(\overline{\theta})\right\}.$$

Proof. The incentive compatibility conditions (3.6) imply:

$$\text{(A.13)} \quad \Delta\theta \overline{q}_0(\underline{\theta}) \geq \overline{U}_0(\underline{\theta}) - \overline{U}_0(\overline{\theta}) \geq \Delta\theta \overline{q}_0(\overline{\theta}).$$

Inserting the second (resp. first) of these inequalities into (5.1) taken for $\theta = \underline{\theta}$ (resp. taken for $\theta = \overline{\theta}$) yields (A.10) (resp. (A.11)).

Finally, there exist values of $\overline{U}_0(\underline{\theta}) - \overline{U}_0(\overline{\theta})$ that satisfy (A.13) and (5.1) if and only if the following condition holds:

(A.14)

$$\min\left\{(1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s (S(q_{\tau+s+1}(\underline{\theta})) - \underline{\theta} q_{\tau+s+1}(\underline{\theta})) - \underline{\theta} q_\tau(\underline{\theta}) + M\right) + \Delta\theta \underline{q}_\tau(\underline{\theta}); \Delta\theta \overline{q}_0(\underline{\theta})\right\} \geq$$

$$\max\left\{-(1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s (S(q_{\tau+s+1}(\overline{\theta})) - \overline{\theta} q_{\tau+s+1}(\overline{\theta})) - \overline{\theta} q_\tau(\overline{\theta}) + M\right) + \Delta\theta \underline{q}_\tau(\overline{\theta}); \Delta\theta \overline{q}_0(\overline{\theta})\right\}.$$

This latter condition can be written in a more compact form as (A.12). ■

Q.E.D.

Proof of Theorem 2: Consider problem $(\mathcal{P})$. We first neglect the enforcement constraints (A.10), (A.11) and (A.12) and only consider, as usual in the screening literature, the participation constraint (3.1) for type $\overline{\theta}$ and the incentive constraint (3.6) for type $\underline{\theta}$. Of course, those two constraints are 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 (5.3), it becomes:

$$\delta^{\tau}\psi(\underline{\theta}, q^{fb}(\underline{\theta})) \geq \max\left\{\Delta\theta q^{bm}(\overline{\theta}) - (1-\delta^{\tau})\Delta\theta q^{fb}(\underline{\theta}); 0\right\}. \tag{A.15}$$

Developing, we obtain the following pair of inequalities:

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) \geq 0, \tag{A.16}$$

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

Because $q^{bm}(\overline{\theta}) < q^{fb}(\underline{\theta})$, the right-hand side of (A.17) is maximum at $\tau = 0$. Manipulating leads to the first inequality in (6.1) which in turn implies (A.16).

Observe that (A.11) now becomes:

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

Developing, we obtain the following pair of inequalities:

$$\psi(\overline{\theta}, q^{bm}(\overline{\theta})) \geq 0, \tag{A.18}$$

$$\psi(\underline{\theta}, q^{bm}(\overline{\theta})) \geq \delta^{-\tau}\Delta\theta((1-\delta^{\tau})\, q^{bm}(\overline{\theta}) - q^{fb}(\underline{\theta})). \tag{A.19}$$

Because $q^{bm}(\overline{\theta}) < q^{fb}(\underline{\theta})$, The righthand side of (A.19) is maximum at $\tau = 0$ and thus

$$\psi(\overline{\theta}, q^{bm}(\overline{\theta})) \geq -\Delta\theta q^{fb}(\underline{\theta}). \tag{A.20}$$

which in turn is implied by (A.18).

Turning now to (A.12). This condition becomes:

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

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

When (6.1) holds, the min on the lefthand side and the the max on the righthand side are achieved for $\tau = +\infty$ so that the condition becomes $q^{fb}(\underline{\theta}) \geq q^{bm}(\overline{\theta})$ which holds. *Q.E.D.*

Proof of Proposition 2: When restricted to stationary contracts, the buyer's problem expressed in per-period payoff and output becomes:

$$(\mathcal{P}): \quad \max_{\{(q(\theta),U(\theta))\}_{\theta\in\Theta}} E_{\theta}\left(S(q(\theta)) - \theta q(\theta) - U(\theta)\right) \text{ subject to (6.3) and}$$

$$U(\underline{\theta}) \geq U(\overline{\theta}) + \Delta\theta q(\overline{\theta}), \tag{A.21}$$

(A.22) $U(\overline{\theta}) \geq 0$

where we restrict the constrained set to the relevant binding incentive and participation constraints (checking ex post that omitted constraints are satisfied). Of course, (A.21) and (A.22) are both binding at the optimum. Inserting those expressions of the rent into the maximand and denoting by Λ the non-negative multiplier of the enforcement constraint (6.3), we may rewrite the Lagrangean of the corresponding problem as:

$$\mathcal{L}(\mathbf{q}, \Lambda) = E_\theta \left(S(q(\theta)) - m(\theta)q(\theta)\right) + \Lambda \left(\psi(\underline{\theta}, q(\underline{\theta})) - \Delta\theta q(\overline{\theta})\right)$$

where $\mathbf{q} = (\mathbf{q}(\underline{\theta}), \mathbf{q}(\overline{\theta}))$ and where the virtual cost parameter $m(\theta)$ is such that $m(\theta) = \begin{cases} \underline{\theta} & \text{if } \theta = \underline{\theta}, \\ \overline{\theta} + \frac{\nu}{1-\nu}\Delta\theta & \text{if } \theta = \overline{\theta}. \end{cases}$ The objective and constraint are both concave with the constrained set obviously satisfying the Slater condition. The Karush-Khün-Tucker conditions for optimality are then necessary and sufficient to characterize the optimal outputs given by (6.4) and (6.5). Of course, $\Lambda > 0$ because (6.6) holds. Otherwise, we would get a contradiction. *Q.E.D.*

Proof of Theorem 3: *Preliminaries.* We first consider a relaxed problem $(\mathcal{P})$ with only (3.6) taken for $\theta = \underline{\theta}$ and (3.1) taken for $\theta = \overline{\theta}$. We are thus neglecting (3.6) taken for $\theta = \overline{\theta}$ and (3.1) taken for $\theta = \underline{\theta}$ that will both be checked ex post. Second, we neglect (A.11) and (A.12) which are checked ex post.

Of course, the participation constraint (3.1) taken for $\theta = \overline{\theta}$ and the incentive constraint (3.6) taken for $\theta = \underline{\theta}$ must both be binding at the optimum of the so relaxed $(\mathcal{P})$. Inserting the expression $\overline{U}_0(\underline{\theta}) = \Delta\theta\overline{q}_0(\overline{\theta})$ into the maximand allows us to simplify the objective function accordingly. The maximand in the problem $(\mathcal{P})$ so obtained is indeed a strictly concave and Fréchet differentiable function defined as:

$$f(\mathbf{q}) = E_\theta \left((1-\delta)\delta^\tau \left(\sum_{\tau=0}^{\infty} \delta^s (S(q_\tau(\theta)) - m(\theta)q_\tau(\theta)) \right) \right).$$

Properties of the maximand and constraints. The sequence $\mathbf{q} = (\mathbf{q}(\underline{\theta}), \mathbf{q}(\overline{\theta}))$ lies in the Banach space $l_1(\delta) \times l_1(\delta)$ which is the cross product of the spaces of sequences considered in the proof of Theorem 1. Let us thus endow the Banach space $l_1(\delta) \times l_1(\delta)$ with the obvious product topology and the norm $\|\mathbf{q}(\theta)\| = \sum_{\theta\in\Theta} \|\mathbf{q}(\theta)\|_\delta$.

The Fréchet derivative of f at any $\mathbf{q} = (\mathbf{q}(\underline{\theta}), \mathbf{q}(\overline{\theta})) \in \mathbf{Q}^\infty \times \mathbf{Q}^\infty$ is a linear and continuous operator $f'(\mathbf{q})$ such that, for any $\mathbf{h} = (\mathbf{h}(\underline{\theta}), \mathbf{h}(\overline{\theta}))$ with $\mathbf{q} + \mathbf{h} \in \mathbf{Q}^\infty \times \mathbf{Q}^\infty$:

$$f'(\mathbf{q})\mathbf{h} = E_\theta \left((1-\delta)\delta^\tau \left(\sum_{\tau=0}^{\infty} \delta^s (S'(q_\tau(\theta)) - m(\theta))h_\tau(\theta) \right) \right).$$

We now assume that the righthand side of (5.3) is positive. We will check ex post that this is indeed the case at the optimum. Thus, (5.3) can be rewritten as:

(A.23) $g_\tau(\mathbf{q}) = \delta^\tau \Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta})) - \Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_\tau(\underline{\theta})) \geq 0.$

Taken together, the aggregate enforcement constraints (A.23) for all dates τ also define a convex set since all functions $g_\tau(\mathbf{q})$ are concave in $\mathbf{q}$ for all $\tau \geq 0$. Let denote by $g(\mathbf{q}) = \{g_\tau(\mathbf{q})\}_{\tau\geq 0}$ the corresponding mapping from $l_1(\delta) \times l_1(\delta)$ onto $l_1(\delta)$.

The mapping $g_\tau(\mathbf{q})$ is obviously Fréchet differentiable with a Fréchet derivative at any $\mathbf{q} \in \mathbf{Q}^\infty$ which is a linear operator $g^{*'}(\mathbf{q})$. Indeed consider such $\mathbf{q}$ and let $\mathbf{h}$ such that $\mathbf{q}+\mathbf{h} \in \mathbf{Q}_\infty$, we can easily compute:

$$g'_\tau(\mathbf{q})\mathbf{h} = (1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s (S'(q_{\tau+s+1}(\overline{\theta})) - \underline{\theta}) h_{\tau+s+1}(\overline{\theta}) - \underline{\theta} h_\tau(\overline{\theta}) \right)$$

$$-\Delta\theta(1-\delta) \left(\sum_{s=0}^{\infty} \delta^s h_s(\overline{\theta}) - \sum_{s=0}^{\tau-1} \delta^s h_s(\underline{\theta} \right).$$

Observe that $g'_\tau(\mathbf{q})$ is a bounded continuous operator. Indeed, following steps similar to those in the proof of Theorem 1, we have:

$$\|g'_\tau(\mathbf{q})\mathbf{h}\| \leq \Delta\theta \left(\|\mathbf{h}(\underline{\theta})\|_\delta + \|\mathbf{h}(\overline{\theta})\|_\delta \right) + (1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s \sup_q |S'(q) - \underline{\theta}||h_{\tau+s+1}(\overline{\theta})| + \overline{\theta}|h_\tau(\overline{\theta})| \right)$$

$$\leq N\|\mathbf{h}\|$$

with $N = \Delta\theta + \max\left\{\sup_q |S'(q) - \theta|, |\overline{\theta}|\right\} < +\infty$ from the fact that S' is bounded above over $\mathcal{Q}$ since S is concave and $S'(0) < +\infty$.

Formulation. With those compact notations, we rewrite the maximization problem as:

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

The corresponding Lagrangean becomes

$$\mathcal{L}(\mathbf{q}, \lambda) = f(\mathbf{q}) + \lambda g(\mathbf{q}) = E_\theta \left((1-\delta) \left(\sum_{\tau=0}^{\infty} \delta^\tau (S(q_\tau(\theta)) - \tilde{\theta}(\theta) q_\tau(\theta)) \right) \right)$$

$$+ \sum_{\tau=0}^{\infty} \lambda_\tau \left(\delta^\tau \Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta})) - \Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_\tau(\underline{\theta})) \right).$$

Optimization. First, observe that, because $\lambda \in l_1(\delta)$, we have also $\sum_{s=0}^{\infty} \lambda_s < +\infty$. We can now explore the implications of (A.3). The following first-order conditions hold:

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

$$S'(q_\tau(\underline{\theta})) - \underline{\theta} = \frac{\lambda_\tau \underline{\theta} - \left(\sum_{s=\tau+1}^{\infty} \lambda_s\right) \Delta\theta}{\nu + \sum_{s=0}^{\tau-1} \lambda_s}. \tag{A.24}$$

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

$$\begin{cases} S'(q_\tau(\overline{\theta})) - \overline{\theta} = \frac{\nu + \sum_{s=0}^{\infty} \lambda_s}{1-\nu} \Delta\theta & \text{if } S'(0) \geq \overline{\theta} + \frac{\nu + \sum_{s=0}^{\infty} \lambda_s}{1-\nu} \Delta\theta, \\ q_\tau(\overline{\theta}) = 0 & \text{otherwise.} \end{cases} \tag{A.25}$$

Observation 1. We necessarily have $q_\tau(\overline{\theta}) \leq q^{bm}(\overline{\theta})$ and the inequality is strict provided one multiplier at least is positive, something we know is true when condition (6.6) holds.

Observation 2. Because $\lambda_\tau \geq 0$ and $\sum_{s=0}^{\infty} \lambda_s < +\infty$, we necessarily have $\lim_{\tau \to +\infty} \lambda_\tau = 0$. Inserting into (A.24), it follows that:

$$\lim_{\tau \to +\infty} S'(q_\tau(\underline{\theta})) = \underline{\theta}. \tag{A.26}$$

The first-best output for a low-cost seller is implemented in the limit.

Non-stationarity. We prove that the contract cannot be stationary. Suppose the contrary. Then, (A.26) implies that necessarily, $q_\tau(\underline{\theta})) = q^{fb}(\underline{\theta})$ for all τ. Turning to (A.24), it implies that $\lambda_\tau = 0$ for all τ which from (A.25) also implies $q_\tau(\overline{\theta}) = q^{bm}(\overline{\theta})$ for all τ.

Inserting these expressions of the outputs into (5.3), we obtain that (A.15) should hold at all τ. But this condition does not hold at date 0 from (6.6).

Binding enforcement constraints. We now investigate the structure of the binding enforcement constraints.

Case 1. Suppose that all constraints (A.23) are binding. We rewrite (A.23) taken at dates τ and $\tau + 1$ respectively as:

$$\Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta})) = \delta^{-\tau} \Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_\tau(\underline{\theta})), \tag{A.27}$$

$$\delta\Psi(\underline{\theta}, \overline{\mathbf{q}}_{\tau+1}(\underline{\theta})) = \delta^{-\tau} \Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_{\tau+1}(\underline{\theta})). \tag{A.28}$$

By sustracting (A.28) from (A.27), 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 yields that the sequence $q_\tau(\underline{\theta})$ must satisfy the recursive condition (6.7) for all pairs τ and $\tau + 1$ for which (A.23) is binding.

Case 1.1. Fix $q_0(\underline{\theta}) \leq q^e(\overline{\theta})$. Because Φ is increasing and $\Phi(q^e(\overline{\theta})) = q^e(\overline{\theta})$, the whole sequence $\mathbf{q}(\underline{\theta})$ defined by (6.7) is (weakly) increasing and such that:

$$q_\tau(\underline{\theta}) \leq q^e(\overline{\theta}) \quad \forall \tau \geq 0. \tag{A.29}$$

Observe that condition (6.6) implies

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) < \Delta\theta q^{bm}(\overline{\theta}) < \Delta\theta q^{fb}(\underline{\theta})$$

which amounts to

$$q^e(\overline{\theta}) < q^{fb}(\underline{\theta}).$$

This inequality taken together with (A.29) yields a contradiction with (A.26).

Case 1.2. $q_0(\underline{\theta}) > q^e(\overline{\theta})$. Using the (strict) concavity of S, (6.7) implies:

$$\delta S'(q^e(\overline{\theta}))(q_{\tau+1}(\underline{\theta}) - q^e(\overline{\theta})) > \overline{\theta}(q_\tau(\underline{\theta}) - q^e(\overline{\theta}))$$

and thus

$$q_\tau(\underline{\theta}) - q^e(\overline{\theta})) > \left(\frac{\overline{\theta}}{\delta S'(q^e(\overline{\theta}))}\right)^\tau (q_0(\underline{\theta}) - q^e(\overline{\theta}))$$

Since $q^e(\overline{\theta}) > q^{db}(\overline{\theta})$, we have $\frac{\overline{\theta}}{\delta S'(q^e(\overline{\theta}))} > 1$, the sequence $\mathbf{q}(\underline{\theta})$ defined by (6.7) then diverges towards $+\infty$. This is again a contradiction with (A.26).

Case 2. Not all constraints (A.23) are binding. Since *Case 1* above cannot arise, we conjecture a solution where (A.23) is binding for all dates $\tau \leq \tau^*$ for some $\tau^* \geq 0$ and slack for $\tau > \tau^*$, i.e., $\lambda_\tau \geq 0$ for $\tau \leq \tau^*$ and $\lambda_\tau = 0$ for $\tau \geq \tau^* + 1$. We are now going to construct the corresponding non-negative multipliers and outputs accordingly.

Observe first that (A.24) and the fact that $\lambda_\tau = 0$ for $\tau \geq \tau^* + 1$ altogether imply:

$$S'(q_\tau(\underline{\theta})) = \underline{\theta} \quad \forall \tau \geq \tau^* + 1. \tag{A.30}$$

The optimal output of the low-cost seller is first-best from date $\tau^* + 1$ on.

Let us now define the sequence $\mathbf{\Lambda}$ of cumulative multipliers as:

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

with our previous convention that $\Lambda_0 = 0$. Because all multipliers λ_s are non-negative, this sequence is by definition non-decreasing and non-negative with terminal value $\Lambda_{\tau^*+1} = \Lambda_\infty$. From the optimality condition (A.24), the sequence $\mathbf{\Lambda}$ 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:

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

We are looking for a sequence such that $\lambda_\tau = \Lambda_{\tau+1} - \Lambda_\tau > 0$ for all $\tau \leq \tau^*$ and $\lambda_\tau = \Lambda_{\tau+1} - \Lambda_\tau = 0$ for $\tau \geq \tau^* + 1$. This implies also that the sequence $\mathbf{q}(\underline{\theta})$ is such that the recursive condition (6.7) holds for all dates $\tau \leq \tau^*$.

Observe also that (A.24) implies

$$S'(q_{\tau^*}(\underline{\theta})) = \underline{\theta} + \frac{\underline{\theta}\lambda_{\tau^*}}{\nu + \Lambda_\infty - \lambda_{\tau^*}} \geq \underline{\theta}$$

and thus

$$q_{\tau^*}(\underline{\theta}) \leq q^{fb}(\underline{\theta}).$$

Consider $q_{\tau^*}(\underline{\theta}) \in \left[\Gamma(q^{fb}(\underline{\theta})), q^{fb}(\underline{\theta})\right)$. Starting from such $q_{\tau^*}(\underline{\theta})$, we may then construct the following 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, this sequence $q_\tau(\underline{\theta})$ for all $\tau \leq \tau^*$ is thus increasing in τ. Moreover, that $q^e(\overline{\theta}) \leq \Gamma(q^{fb}(\underline{\theta})) < q^{fb}(\underline{\theta})$ with $q^e(\overline{\theta}) = \Gamma(q^e(\overline{\theta}))$ implies also that:

$$q^e(\overline{\theta}) \leq q_\tau(\underline{\theta}) \quad \forall \tau \leq \tau^* - 1. \tag{A.32}$$

We can rewrite (A.31) as

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

with

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

Observe that $q_\tau(\underline{\theta}) > q^e(\overline{\theta}) > q^{fb}(\overline{\theta})$ by assumption which implies that $\alpha_\tau < 1$ for all τ. Also, $q_{\tau+1}(\underline{\theta}) < q^{fb}(\underline{\theta})$ implies that $\beta_\tau > 0$ for all τ.

As long as the sequence Λ is monotonically increasing (so that the multipliers λ_τ remain non-negative) for a given sequence of outputs satisfying (6.7) for $\tau \leq \tau^*$ and $q_{\tau^*}(\underline{\theta}) \in \left[\Gamma(q^{fb}(\underline{\theta})), q^{fb}(\underline{\theta})\right)$, we have the solution. Slightly abusing notations, we will thus consider an arbitrary sequence satisfying (A.33) and study its monotonicity properties. In this respect, define now two sequences C and $\tilde{\Lambda}$ such that:

$$\Lambda_\tau = C_\tau \tilde{\Lambda}_\tau \text{ and } \tilde{\Lambda}_{\tau+1} = \alpha_\tau \tilde{\Lambda}_\tau.$$

Observe that it is just a normalization to set $\tilde{\Lambda}_0 = 1$ provided that C_0 is conveniently adjusted. From this, we immediately obtain:

$$\tilde{\Lambda}_\tau = \Pi_{s=0}^{\tau-1} \alpha_s.$$

Inserting into (A.33) then yields:

$$C_{\tau+1} - C_\tau = \gamma_\tau \text{ where } \gamma_\tau = \frac{\beta_\tau}{\Pi_{s=0}^{\tau} \alpha_s}.$$

From which we get:

$$C_\tau = C_0 + \sum_{s=0}^{\tau-1} \gamma_s$$

where C_0 is a constant to be found. Finally, we obtain:

$$\text{(A.35)} \quad \Lambda_\tau = \left(C_0 + \sum_{s=0}^{\tau-1} \gamma_s \right) \Pi_{s=0}^{\tau-1} \alpha_s.$$

This expression leads us immediately to:

$$\frac{\Lambda_{\tau+1}}{\Lambda_\tau} = \alpha_\tau \frac{C_0 + \sum_{s=0}^{\tau} \gamma_s}{C_0 + \sum_{s=0}^{\tau-1} \gamma_s}.$$

Because $\beta_\tau > 0$, $\gamma_\tau > 0$ and thus $\frac{C_0 + \sum_{s=0}^{\tau} \gamma_s}{C_0 + \sum_{s=0}^{\tau-1} \gamma_s} > 1$. Hence, the ratio $\frac{\Lambda_{\tau+1}}{\Lambda_\tau}$ is the product of two terms, one α_τ being less than one while the second is greater than one. Observe that $\frac{\Lambda_{\tau+1}}{\Lambda_\tau} \geq 1$ when

$$C_0 \leq v_\tau = \frac{\gamma_\tau \alpha_\tau}{1 - \alpha_\tau} - \sum_{s=0}^{\tau-1} \gamma_s.$$

The sequence v_τ is monotonically decreasing whenever:

$$\text{(A.36)} \quad \frac{\gamma_\tau}{1 - \alpha_\tau} \geq \frac{\gamma_{\tau+1} \alpha_{\tau+1}}{1 - \alpha_{\tau+1}}.$$

Notice that

$$\gamma_{\tau+1} = \frac{\gamma_\tau \beta_{\tau+1}}{\alpha_{\tau+1} \beta_\tau}.$$

Inserting into (A.36), we see that $v_{\tau+1} \leq v_\tau$ if and only if

$$\text{(A.37)} \quad \frac{\beta_\tau}{1-\alpha_\tau} \geq \frac{\beta_{\tau+1}}{1-\alpha_{\tau+1}}.$$

We have:

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

Because, starting from any $q_0(\underline{\theta}) > q^e(\overline{\theta})$, the sequence $q_\tau(\underline{\theta})$ defined by the recursive equation (6.7) is monotonically increasing in τ and $q_\tau(\underline{\theta}) > q^e(\overline{\theta}) > q^{fb}(\overline{\theta})$ for all τ, the sequence $\frac{\beta_\tau}{1-\alpha_\tau}$ is monotonically decreasing in τ. Thus, the sequence v_τ is also monotonically decreasing. It implies that, for any $\tau^* \geq 1$ there always exists a non-empty interval of values for C_0 such that:

$$\text{(A.38)} \quad v_{\tau^*} \leq C_0 \leq v_{\tau^*-1}.$$

For such C_0, the monotonicity of v_τ implies that

$$C_0 < v_{\tau'} \quad \forall \tau' \leq \tau^* - 1 \text{ and } C_0 > v_{\tau'} \quad \forall \tau' \geq \tau^*.$$

For such values of C_0,

$$\frac{\Lambda_{\tau'+1}}{\Lambda_{\tau'}} \geq 1 \quad \forall \tau'^* - 1 \text{ and } \frac{\Lambda_{\tau'+1}}{\Lambda_{\tau'}} < 1 \quad \forall \tau' \geq \tau^*.$$

Hence, by choosing C_0 such that (A.38) holds, the sequence Λ_τ achieves its maximum at τ^*. We thus have:

$$\Lambda_{\tau^*} = \Lambda_\infty.$$

Taking $C_0 = v_{\tau^*}$ and using the definition of Λ_{τ^*} coming from (A.35), we obtain the following condition linking Λ_∞ and τ^*

$$\text{(A.39)} \quad \Lambda_\infty = \frac{\Pi_{s=0}^{\tau^*}\alpha_s}{1-\alpha_{\tau^*}}\gamma_{\tau^*}(\Lambda_\infty) = \frac{\beta_{\tau^*}(\Lambda_\infty)}{1-\alpha_{\tau^*}}$$

where we make explicit the dependence of γ_{τ^*}, and β_{τ^*} on Λ_∞ from (A.34) explicit.

Denote $\eta(\tau^*, \Lambda_\infty)$ the righthand side of (A.39). Observe that $\eta(\tau^*, 0) > 0$ and that (A.39) has a unique solution $\Lambda_\infty^{\tau^*} > 0$. Moreover, (A.37) implies that $\Lambda_\infty^{\tau^*}$ is an increasing sequence in τ^*. From this, we may as well define the step function $\tau^*(\Lambda_\infty)$ with the convention that $\tau^*(\Lambda_\infty^{\tau^*}) = \tau^*$.

We conclude that, whenever $\Lambda_\infty \in \left[\Lambda_\infty^{\tau^*}, \Lambda_\infty^{\tau^*+1}\right)$, the sequence $\mathbf{\Lambda}$ satisfying (A.33) and such that $\Lambda_{\tau^*} = \Lambda_\infty$ is monotonically increasing over the first τ^* periods.

To find the value of Λ_∞ at the optimum, remember that we are conjecturing an optimal contract with all enforcement constraints (A.23) being binding for $\tau \leq \tau^*(\Lambda_\infty)$ and slack for $\tau > \tau^*(\Lambda_\infty)$. The fact that those constraints are all binding for $\tau \leq \tau^*(\Lambda_\infty)$ implies that the low-cost seller's outputs follow the recursive sequence $\Gamma(q_\tau(\underline{\theta})) = q_{\tau-1}(\underline{\theta})$ with $q_{\tau^*(\Lambda_\infty)}(\underline{\theta})$ (or alternatively $q_0(\underline{\theta})$ such that $q_0(\underline{\theta}) = \Gamma^{\tau^*(\Lambda_\infty)}(q_{\tau^*(\Lambda_\infty)}(\underline{\theta}))$) to be found. the fact that (A.23) is binding at all date $\tau \leq \tau^*(\Lambda_\infty)$ and slack thereafter, so that $q_\tau(\underline{\theta}) = q^{fb}(\underline{\theta})$ for $\tau > \tau^*(\Lambda_\infty)$, we can rewrite the binding constraint (A.23) at date $\tau^*(\Lambda_\infty)$ to get the following condition:

$$\text{(A.40)} \quad \delta(S(q^{fb}(\underline{\theta})) - \underline{\theta}q^{fb}(\underline{\theta})) - (1-\delta)\underline{\theta}q_{\tau^*(\Lambda_\infty)}(\underline{\theta}) + (1-\delta)M = \vartheta(\Lambda_\infty)$$

where

$$\vartheta(\Lambda_\infty) = \delta^{-\tau^*(\Lambda_\infty)}\Delta\theta\left(q(\overline{\theta},\Lambda_\infty) - (1-\delta)\sum_{s=0}^{\tau^*(\Lambda_\infty)-1}\delta^s q_s(\underline{\theta})\right)$$

and where from (A.25), we make explicit the dependence of $q(\overline{\theta},\Lambda_\infty)$ on Λ_∞ as:

$$S'(q(\overline{\theta},\Lambda_\infty)) = \overline{\theta} + \frac{\nu+\Lambda_\infty}{1-\nu}\Delta\theta \tag{A.41}$$

where we assume that $S'(0)$ is large enough so that $q(\overline{\theta},\Lambda_\infty)$ remains nonnegative. Observe that $q(\overline{\theta},\Lambda_\infty)$ is decreasing with Λ_∞. Observe also that ϑ is strictly decreasing on each interval $\left[\Lambda_\infty^k,\Lambda_\infty^{k+1}\right)$ where $\tau^*(\Lambda_\infty)=k$ remains constant. At a discontinuity point of ϑ, say $\Lambda_\infty^{\tau+1}$, where $\tau^*(\Lambda_\infty)$ jumps from τ to $\tau+1$; we have $\vartheta(\Lambda_\infty^{\tau+1-}) \geq \vartheta(\Lambda_\infty^{\tau+1})$ if and only if:

$$\delta^{-\tau}\Delta\theta\left(q(\overline{\theta},\Lambda_\infty^{\tau+1}) - (1-\delta)\sum_{s=0}^{\tau-1}\delta^s q_s(\underline{\theta})\right) \geq \delta^{-\tau-1}\Delta\theta\left(q(\overline{\theta},\Lambda_\infty^{\tau+1}) - (1-\delta)\sum_{s=0}^{\tau}\delta^s q_s(\underline{\theta})\right),$$

or equivalently

$$q(\overline{\theta},\Lambda_\infty^{\tau+1}) \leq \delta^\tau q_\tau(\underline{\theta}) + (1-\delta)\sum_{s=0}^{\tau-1}\delta^s q_s(\underline{\theta}). \tag{A.42}$$

Now observe that $q_\tau(\underline{\theta}) \geq q^e(\overline{\theta})$ for all τ. Hence, the righthand side of (A.42) is bounded below by $q^e(\overline{\theta})$. Finally, the second condition in condition 6.6 implies that $q^e(\overline{\theta}) \geq q^{bm}(\overline{\theta})$. Since $q^{bm}(\overline{\theta}) \geq q(\overline{\theta},\Lambda_\infty^{\tau+1})$, (A.42) immediately follows. Hence, ϑ has a jump downwards at the discontinuity point $\Lambda_\infty^{\tau+1}$. There are of course only a countable number of such discontinuities. We conclude that ϑ is everywhere decreasing. Moreover, we have $\tau^*(0)=0$ and thus $\vartheta(0) = \Delta\theta q^{bm}(\overline{\theta}) > 0$ while, for Λ_∞ large enough, $q(\overline{\theta},\Lambda_\infty) < q^e(\overline{\theta}) < (1-\delta)\sum_{s=0}^{\tau^*(\Lambda_\infty)-1}\delta^s q_s(\underline{\theta})$ so that $\vartheta(\Lambda_\infty) < 0$ for such Λ_∞. Hence, there exists Λ_∞^{sup} such that $\vartheta(\Lambda_\infty) \geq 0 \quad \forall \Lambda_\infty \leq \Lambda_\infty^{sup}$. Let denote the possible discontinuity (open) interval as: $\mathcal{J}_\tau = \left(\vartheta(\Lambda_\infty^{\tau+1}), \vartheta(\Lambda_\infty^{\tau+1-})\right)$.

Two cases (mutually exclusive) have now to be considered.

Case 1. Suppose that there exists Λ_∞^* such that $A)$ Λ_τ is an increasing sequence with $\max\Lambda_\tau = \Lambda_{\tau^*} = \Lambda_\infty^*$ and $B)$ $q_{\tau^*(\Lambda_\infty)}(\underline{\theta}) \in \left[\Gamma(q^{fb}(\underline{\theta})), q^{fb}(\underline{\theta})\right)$ is such that (A.40) holds. From condition 6.6 , the righthand side above is positive and less than $\Delta\theta q^{bm}(\overline{\theta})$. Hence, Λ_∞^* is positive, less than Λ_∞^{sup} and $0 < \tau^*(\Lambda_\infty^*) < +\infty$. Starting from the values $q_{\tau^*(\Lambda_\infty)}(\underline{\theta})$ and Λ_∞^*, we can reconstruct the sequences of outputs $q_\tau(\underline{\theta})$ for $\tau \leq \tau^*(\Lambda_\infty)-1$ and the corresponding cumulative multipliers Λ_τ which in turn fully determines the positive multipliers λ_τ for $\tau \leq \tau^*(\Lambda_\infty^*)$.

Case 2. Suppose that there exists τ^* such that

$$\psi(\underline{\theta}, q^{fb}(\underline{\theta})) \in \mathcal{J}_{\tau^*}. \tag{A.43}$$

Because there is a countable set of discontinuities for ϑ, this case almost never arises. When it does, one must adjust the value of $q_{\tau^*}(\underline{\theta})$ so that (A.40) holds for τ^* and $\Lambda_\infty^* = \Lambda_\infty^{\tau*}$.

For further references, we note $q^{sb}(\overline{\theta}) = q(\overline{\theta},\Lambda_\infty^*)$ as defined in (A.41) and $q_\tau^{sb}(\underline{\theta}) = q_\tau(\underline{\theta})$ where the latter sequence is defined recursively for $\tau \leq \tau^*(\Lambda_\infty^*)$ through (6.7) with $q_{\tau^*(\Lambda_\infty^*)}(\underline{\theta}) = q^{fb}(\underline{\theta})$.

Checking the omitted constraints. It is routine to check that (3.6) for $\overline{\theta}$ and (3.1) for $\underline{\theta}$ are both satisfied.

From the fact that $\psi(\underline{\theta}, q^{fb}(\underline{\theta})) > 0$ (i.e., $q^{fb}(\underline{\theta}) < q^e(\underline{\theta})$), the righthand side of (5.3) is strictly positive for all $\tau \leq \tau^*(\Lambda_\infty)$ so g_τ can be written as (A.23) as conjectured.

Now we check the remaining enforcement constraints. First, (A.11) amounts to

$$\delta^\tau \psi(\overline{\theta}, q^{sb}(\overline{\theta})) \geq \max\left\{\Delta\theta((1-\delta^\tau)q^{sb}(\overline{\theta}) - \overline{q}_0^{sb}(\underline{\theta})), 0\right\}.$$

Observe that $q^{sb}(\overline{\theta}) < q^{bm}(\overline{\theta}) < q^e(\overline{\theta})$. Thus,

$$\psi(\overline{\theta}, q^{sb}(\overline{\theta})) > \psi(\overline{\theta}, q^{bm}(\overline{\theta})) > 0 \tag{A.44}$$

where the last inequality follows from the assumption $q^{fb}(\overline{\theta}) \leq q^e(\overline{\theta})$. Observe also that

$$\delta^\tau \psi(\overline{\theta}, q^{sb}(\overline{\theta})) \geq \Delta\theta((1-\delta^\tau)q^{sb}(\overline{\theta}) - \overline{q}_0^{sb}(\underline{\theta})) \tag{A.45}$$

since

$$\delta^\tau \psi(\underline{\theta}, q^{sb}(\overline{\theta})) \geq 0 \geq \Delta\theta(q^{sb}(\overline{\theta}) - \overline{q}_0^{sb}(\underline{\theta}))$$

where the last inequality follows from the fact that $q^{sb}(\overline{\theta}) < q^{bm}(\overline{\theta}) < q_\tau^e(\underline{\theta})$ for all τ and the first inequality follows from $q^{sb}(\overline{\theta}) < q^{bm}(\overline{\theta}) < q^e(\overline{\theta}) < q^e(\underline{\theta})$. Finally (A.44) and (A.45) taken altogether ensures that (A.11) holds.

Observe that (A.12) now amounts to

$$\min\left\{\delta^\tau \psi(\overline{\theta}, q^{sb}(\overline{\theta})) + \Delta\theta \underline{q}_\tau^{sb}(\underline{\theta}); \Delta\theta \overline{q}_0^{sb}(\underline{\theta})\right\} \geq \max\left\{-\delta^\tau \Psi(\underline{\theta}, \overline{q}_\tau^{sb}(\underline{\theta})) + \Delta\theta(1-\delta^\tau)q^{sb}(\overline{\theta}); \Delta\theta q^{sb}(\overline{\theta})\right\}.$$

Several cases should be considered.

1. Observe that $q^{sb}(\overline{\theta}) < q^{bm}(\overline{\theta}) < q^e(\overline{\theta})$ implies

$$\Delta\theta \overline{q}_0^{sb}(\underline{\theta}) \geq \Delta\theta q^e(\overline{\theta}) \geq \Delta\theta q^{sb}(\overline{\theta}).$$

2. Note that $\Psi(\underline{\theta}, \overline{q}_\tau^{sb}(\underline{\theta})) \geq 0$. Thus, we have

$$\Delta\theta \overline{q}_0^{sb}(\underline{\theta}) \geq -\delta^\tau \Psi(\underline{\theta}, \overline{q}_\tau^{sb}(\underline{\theta})) + \Delta\theta q^{sb}(\overline{\theta}) > -\delta^\tau \Psi(\underline{\theta}, \overline{q}_\tau^{sb}(\underline{\theta})) + \Delta\theta(1-\delta^\tau)q^{sb}(\overline{\theta}).$$

3. We have $\psi(\overline{\theta}, q^{sb}(\overline{\theta})) \geq 0$. Thus, we get:

$$\delta^\tau \psi(\overline{\theta}, q^{sb}(\overline{\theta})) + \Delta\theta \underline{q}_\tau^{sb}(\underline{\theta}) \geq -\delta^\tau \Psi(\underline{\theta}, \overline{q}_\tau^{sb}(\underline{\theta})) + \Delta\theta(1-\delta^\tau)q^{sb}(\overline{\theta}).$$

4. Observe that the following string of inequalities holds:

$$\begin{aligned}
&\delta^\tau \psi(\overline{\theta}, q^{sb}(\overline{\theta})) + \Delta\theta \underline{q}_\tau^{sb}(\underline{\theta}) \geq \delta^\tau \psi(\overline{\theta}, q^{sb}(\overline{\theta})) + \Delta\theta(1-\delta^\tau)q_\tau^{sb}(\underline{\theta}) \\
&\geq \delta^\tau \psi(\overline{\theta}, q^{sb}(\overline{\theta})) + \Delta\theta(1-\delta^\tau)q^e(\overline{\theta}) \geq \delta^\tau \psi(\overline{\theta}, q^{sb}(\overline{\theta})) + \Delta\theta(1-\delta^\tau)q^{bm}(\overline{\theta}) \\
&\geq \delta^\tau \psi(\overline{\theta}, q^{bm}(\overline{\theta})) + \Delta\theta(1-\delta^\tau)q^{bm}(\overline{\theta}) = \Delta\theta q^{bm}(\overline{\theta}) + \delta^\tau\left(\psi(\overline{\theta}, q^{bm}(\overline{\theta})) - \Delta\theta q^{bm}(\overline{\theta})\right) \\
&\geq \Delta\theta q^{sb}(\overline{\theta})
\end{aligned}$$

where the last inequality follows from $\psi(\overline{\theta}, q^{bm}(\overline{\theta})) \geq \Delta\theta q^{bm}(\overline{\theta})$ in (6.6).

Finally, note also that (A.11) follows from (A.12). Indeed (A.12) implies

$$\delta^\tau \Psi(\overline{\theta}, \overline{\mathbf{q}}_\tau(\overline{\theta})) + \Delta\theta \underline{q}_\tau(\underline{\theta}) \geq \Delta\theta \overline{q}_0(\overline{\theta}).$$

Because $\overline{q}_0(\underline{\theta}) \geq \underline{q}_\tau(\underline{\theta})$ and $\overline{q}_0(\overline{\theta}) \geq \underline{q}_\tau(\overline{\theta})$ we have

$$\delta^\tau \Psi(\overline{\theta}, \overline{\mathbf{q}}_\tau(\overline{\theta})) + \Delta\theta \overline{q}_0(\underline{\theta}) \geq \delta^\tau \Psi(\overline{\theta}, \overline{\mathbf{q}}_\tau(\overline{\theta})) + \Delta\theta \underline{q}_\tau(\underline{\theta}) \geq \Delta\theta \overline{q}_0(\overline{\theta}) \geq \Delta\theta \underline{q}_\tau(\overline{\theta}).$$

Q.E.D.

Proof of Proposition 3: From the Proof of Lemma A.1, the fact that the enforcement constraint (5.3) is binding at a given date $\tau \leq \tau^*$ implies that all constraints below are also binding at that date:

$$\overline{U}_0(\underline{\theta}) \geq \overline{U}_0(\overline{\theta}) + \Delta\theta \underline{q}_\tau(\overline{\theta}) - \delta^\tau(\overline{U}_\tau(\overline{\theta}) - (1-\delta)(t_{1,\tau}(\overline{\theta}) - L)), \tag{A.46}$$

$$\overline{U}_0(\overline{\theta}) = \overline{U}_0(\underline{\theta}) - \Delta\theta \underline{q}_\tau(\underline{\theta}) - \delta^\tau(\overline{U}_\tau(\underline{\theta}) - (1-\delta)(t_{1,\tau}(\underline{\theta}) - L)), \tag{A.47}$$

$$(1-\delta)\delta^\tau \left(\delta \sum_{s=0}^{\infty} \delta^s (S(q_{\tau+s+1}(\underline{\theta})) - \underline{\theta} q_{\tau+s+1}(\underline{\theta})) - \underline{\theta} q_\tau(\underline{\theta}) + M \right) \geq \delta^\tau \left(\overline{U}_\tau(\underline{\theta}) - (1-\delta)(t_{1,\tau}(\underline{\theta}) - L) \right). \tag{A.48}$$

Pre-delivery payments. The Proof of Lemma A.1 taken in tandem with the Proof of Theorem 3 also show in passing that there is no loss of generality in choosing pre-delivery payments that satisfy (6.11), provided that $\overline{U}_\tau^{sb}(\theta) - (1-\delta)(t_{1,\tau}^{sb}(\theta) - L) \geq 0$, i.e., $\overline{U}_\tau^{sb}(\theta)$ remains non-negative.

Payments to the high-cost seller. Remember that (3.1) and (3.6) are respectively binding and slack for $\theta = \overline{\theta}$ and the output of such high-cost type is stationary. Without loss of generality, we take a stationary payment satisfying (6.12). Then, we get

$$\overline{U}_\tau^{sb}(\overline{\theta}) = 0 \quad \forall \tau. \tag{A.49}$$

Observe that this stationary payment satisfies the enforcement constraints (A.46) and

$$\overline{U}_\tau(\underline{\theta}) \geq (1-\delta)(t_{1,\tau}(\underline{\theta}) - L) \tag{A.50}$$

which, taking into account that (A.49) and (6.11), rewrite respectively as

$$\overline{U}_0^{sb}(\underline{\theta}) = \Delta q^{sb}(\overline{\theta}) \geq \Delta\theta \underline{q}_\tau^{sb}(\overline{\theta}) \quad \forall \tau, \tag{A.51}$$

and

$$\overline{U}_\tau^{sb}(\underline{\theta}) \geq 0 \quad \forall \tau. \tag{A.52}$$

We already know from the Proof of Theorem 3 that $\overline{U}_\tau^{sb}(\underline{\theta}) = \delta^{-\tau}\Delta\theta(q^{sb}(\overline{\theta}) - \underline{q}_\tau^{sb}(\underline{\theta})) \geq 0$ for $\tau \leq \tau^*$. Hence, (A.51) holds for all $\tau \leq \tau^*$ and *a fortiori* for $\tau > \tau^*$. Condition (A.52) is trivially satisfied.

Payments to the low-cost seller. From the fact that (A.47) is binding at the optimal contract and the fact that (3.1) and (3.6) are respectively binding for $\theta = \overline{\theta}$ and $\theta = \underline{\theta}$, we get:

$$\text{(A.53)} \quad \overline{U}_\tau^{sb}(\underline{\theta}) = \delta^{-\tau}\Delta\theta(q^{sb}(\overline{\theta}) - \underline{q}_\tau^{sb}(\underline{\theta})) \quad \forall \tau \leq \tau^*.$$

From the Proof of Theorem 3, we also know that $\overline{U}_\tau^{sb}(\underline{\theta}) = \delta^{-\tau}\Delta\theta(q^{sb}(\overline{\theta}) - \underline{q}_\tau^{sb}(\underline{\theta})) \geq 0$ for $\tau \leq \tau^*$. From the definition of $\overline{U}_\tau^{sb}(\underline{\theta})$, we have:

$$\overline{U}_\tau^{sb}(\underline{\theta}) = (1-\delta)(t_\tau^{sb}(\underline{\theta}) - \underline{\theta} q_\tau^{sb}(\underline{\theta})) + \delta \overline{U}_{\tau+1}^{sb}(\underline{\theta}).$$

Inserting the expression of $\overline{U}_\tau^{sb}(\underline{\theta})$ and $\overline{U}_{\tau+1}^{sb}(\underline{\theta})$ obtained from (A.53) as long as $\tau \leq \tau^* - 1$, we get (6.13).

Let us look for a stationary payment $t_\tau^{sb}(\underline{\theta}) \equiv t^{sb}(\underline{\theta})$ for $\tau \geq \tau^*$. We can rewrite the binding incentive constraint (3.6) for $\theta = \underline{\theta}$ as:

$$\overline{U}_0^{sb}(\underline{\theta}) = (1-\delta)\sum_{s=0}^{\tau^*-1} \delta^s (t_s^{sb}(\underline{\theta}) - \underline{\theta} q_s^{sb}(\underline{\theta})) + (1-\delta)\sum_{\tau^*}^{+\infty} \delta^s (t_\tau^{sb}(\underline{\theta}) - \underline{\theta} q_s^{sb}(\underline{\theta})).$$

Taking into account that $\overline{U}_0^{sb}(\underline{\theta}) = \Delta\theta(q^{sb}(\overline{\theta})$, we finally obtain (6.14). *Q.E.D.*

Proof of Proposition 4: From (A.40) and the fact that $q_{\tau^*}(\Lambda_\infty)(\underline{\theta}) < q^{fb}(\underline{\theta})$, we get:

$$\delta(S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta})) - (1-\delta) q_{\tau^*(\Lambda_\infty)}(\underline{\theta}) + (1-\delta)M < \delta(S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta})) - (1-\delta)\underline{\theta} q^{fb}(\underline{\theta})) + (1-\delta)M$$

$$= \delta^{-\tau^*(\Lambda_\infty)}\Delta\theta \left(q(\overline{\theta}, \Lambda_\infty) - (1-\delta)\sum_{s=0}^{\tau^*(\Lambda_\infty)-1} \delta^s q_s(\underline{\theta}) \right) \leq \delta^{-\tau^*}\Delta\theta \left(q^{bm}(\overline{\theta}) - (1-\delta^{\tau^*}) q^e(\overline{\theta}) \right)$$

where the last inequality follows from the fact that $q^{bm}(\overline{\theta}) \geq q(\overline{\theta}, \Lambda_\infty)$ and $q_s(\underline{\theta}) \geq q^e(\overline{\theta})$ for all $s \leq \tau^*$. From this, it follows that:

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

Manipulating this inequality and taking the closest higher integer gives the righthand side of (6.15). *Q.E.D.*

Proof of Proposition 5: There is exactly one round in the growing phase when the conditions of optimality with respect to outputs and the binding constraint (A.40) for $\tau^* = 0$ defines a unique positive λ_0. Coming back on the optimality conditions w.r.t. $q_0(\underline{\theta})$ and $q_0(\overline{\theta})$ respectively and making the dependence of those outputs on λ_0 explicit, we rewrite:

$$\text{(A.54)} \quad S'(q_0(\underline{\theta}, \lambda_0)) - \underline{\theta} = \frac{\lambda_0}{\nu}\underline{\theta},$$

$$\text{(A.55)} \quad S'(q_0(\overline{\theta}, \lambda_0)) - \overline{\theta} = \frac{\nu + \lambda_0}{1-\nu}\Delta\theta.$$

Making the dependence of those outputs on λ_0 explicit, the fact that (A.40) is binding for $\tau^* = 0$ can now be rewritten as:

$$\text{(A.56)} \quad \delta\left(S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta})\right) + (1-\delta)M = \Delta\theta q_0(\overline{\theta}, \lambda_0) + (1-\delta)\underline{\theta} q_0(\underline{\theta}, \lambda_0).$$

Solving the system (A.54)-(A.55)-(A.56) yields $\lambda_0 > 0$ when condition (6.6) holds.

On the other hand, (A.40) is slack for $\tau \geq 1$ when it is already slack at $\tau = 1$, i.e., when:

$$\text{(A.57)} \quad \delta S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta}) + (1-\delta)M \geq \delta^{-1}\left(\Delta\theta q_0(\overline{\theta}, \lambda_0) - (1-\delta)\Delta\theta q_0(\underline{\theta}, \lambda_0)\right).$$

Taken together, (A.56) and (A.57) imply:

$$\text{(A.58)} \quad q_0(\underline{\theta}, \lambda_0) \leq \Gamma(q^{fb}(\underline{\theta})).$$

This condition ensures that there is a unique round on the growing phase:

Denote $x = \delta(S(q^{fb}(\underline{\theta})) - \underline{\theta} q^{fb}(\underline{\theta})) + (1-\delta)M = \psi(\underline{\theta}, q^{fb}(\underline{\theta})) + (1-\delta)\underline{\theta} q^{fb}(\underline{\theta})$ and observe that λ_0 is actually a decreasing function of x only. The condition (A.58) can thus be rewritten as:

$$\text{(A.59)} \quad q_0(\underline{\theta}, \lambda_0(\psi(\underline{\theta}, q^{fb}(\underline{\theta}))) + (1-\delta)\underline{\theta} q^{fb}(\underline{\theta})) \leq \Gamma(q^{fb}(\underline{\theta})) = \frac{1}{\overline{\theta}}(\psi(\underline{\theta}, q^{fb}(\underline{\theta})) + \underline{\theta} q^{fb}(\underline{\theta})).$$

Let us fix $\psi(\underline{\theta}, q^{fb}(\underline{\theta}))$ and δ such that (A.59) holds. Increasing δ and decreasing accordingly M ensures that (A.59) still holds so that the growing phase has still one period only.

Q.E.D.

Proof of Theorem 4: Consider the case where only condition (6.2) holds. Our optimization cannot rule out the possibility that both (5.3) and (A.12) are binding.[55] In fact, it is enough to include in our optimization a simplified version of (A.12) namely:

$$\text{(A.60)} \quad \delta^\tau \Psi(\overline{\theta}, \overline{\mathbf{q}}_\tau(\overline{\theta})) + \Delta\theta \underline{q}_\tau(\underline{\theta}) \geq \Delta\theta \overline{q}_0(\overline{\theta}).$$

Denoting by μ the sequence of multipliers for constraints (A.60), the Lagrangean becomes:

$$\mathcal{L}(\mathbf{q}, \lambda, \mu) = E_\theta\left((1-\delta)\left(\sum_{\tau=0}^{\infty} \delta^\tau (S(q_\tau(\theta)) - \tilde{\theta}(\theta) q_\tau(\theta))\right)\right)$$

$$+\sum_{\tau=0}^{\infty} \lambda_\tau \left(\delta^\tau \Psi(\underline{\theta}, \overline{\mathbf{q}}_\tau(\underline{\theta})) - \left(\Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_\tau(\underline{\theta}))\right)\right) + \sum_{\tau=0}^{\infty} \mu_\tau \left(\delta^\tau \Psi(\overline{\theta}, \overline{\mathbf{q}}_\tau(\overline{\theta})) - \left(\Delta\theta(\overline{q}_0(\overline{\theta}) - \underline{q}_\tau(\underline{\theta}))\right)\right).$$

Optimization. We proceed as in the Proof of Theorem 3 and thus much details are omitted. The following first-order conditions hold:

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

$$S'(q_\tau(\underline{\theta})) - \underline{\theta} = \frac{\lambda_\tau \underline{\theta} - \left(\sum_{s=\tau+1}^{\infty} (\lambda_s + \mu_s)\right)\Delta\theta}{\nu + \sum_{s=0}^{\tau-1} \lambda_s};$$

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

$$S'(q_\tau(\overline{\theta})) - \overline{\theta} = \frac{(\nu + \sum_{s=0}^{\infty} \lambda_s)\Delta\theta + \mu_\tau \overline{\theta}}{1 - \nu + \sum_{s=0}^{\tau-1} \mu_s}$$

[55] From Item 5. in the Proof of Theorem 3, we already know that (A.11) is implied by (A.12).

Observation 1. Because $\lambda_\tau, \mu_\tau \geq 0$ and $\sum_{s=0}^{\infty} \lambda_s < +\infty$ and $\sum_{s=0}^{\infty} (\lambda_s + \mu_s) < +\infty$ we necessarily have $\lim_{\tau \to +\infty} \lambda_\tau = 0$. Inserting into the first-order condition, we get (7.1).

Observation 2. Note that now the output of the inefficient seller is not necessarily stationary but, yet, (7.2) holds.

Q.E.D.

Proof of Lemma 3: The following Lemma fully characterizes incentive compatible and individually rational allocations $\left(\overline{U}_0(\theta), \overline{q}_0(\theta)\right)$.

Lemma A.2 *$\overline{U}_0(\theta)$ is convex, absolutely continuous and satisfies:*

$$\overline{U}_0(\theta) - \overline{U}_0(\overline{\theta}) = \int_{\theta}^{\overline{\theta}} \overline{q}_0(x)dx \tag{A.61}$$

with $\overline{q}_0(\theta)$ satisfiying (3.7) and

$$\overline{U}_0(\overline{\theta}) \geq 0. \tag{A.62}$$

Proof. Observe that $\overline{U}_0(\theta)$ is the maximum of linear functions and as such is convex. Convexity implies that $\overline{q}_0(\theta)$ should be weakly decreasing. Absolute continuity follows from an argument in Milgrom and Segal (2002). This property immediately implies (A.61). Last, the participation constraint (3.1) can be reduced to satisfying (A.62) only because $\overline{U}_0(\theta)$ is weakly decreasing.■

Condition (7.4) is then obtained by inserting (A.61) into (5.1).

Observe that the maximand on the righthand side of (7.4) is a concave function of $\hat{\theta}$ since $\overline{q}_0(\theta)$ is weakly decreasing by (3.7). It achieves its maximum at $\hat{\theta}^*_\tau(\theta)$ such that

$$\overline{q}_0(\hat{\theta}^*_\tau(\theta)) = \underline{q}_\tau(\theta).$$

Because $\overline{q}_0(\theta) > \underline{q}_\tau(\theta)$, we necessarily have $\hat{\theta}^*_\tau(\theta) > \theta$ at all $\theta < \overline{\theta}$.

Q.E.D.

Proof of Theorem 5: We just want to check that (7.4) holds for the Baron-Myerson allocation $(U^{bm}(\theta), q^{bm}(\theta))$. To this end, observe that the righthand side can be rewritten as:

$$\varphi(\tau, \theta) = \delta^{-\tau} \max_{\hat{\theta} \in \Theta} \left\{ \int_{\theta}^{\hat{\theta}} q^{bm}(x)dx + (\theta - \hat{\theta})(1 - \delta^\tau) q^{bm}(\theta) \right\}.$$

The above maximand is a concave function of $\hat{\theta}$ since $q^{bm}(\theta)$ is weakly decreasing with a uniform distribution. It achieves its maximum at $\hat{\theta}^*_\tau(\theta)$ such that

$$q^{bm}(\hat{\theta}^*_\tau(\theta)) = (1 - \delta^\tau) q^{bm}(\theta)$$

so that $\hat{\theta}^*_\tau(\theta) > \theta$ immediately follows. Treating (somewhat abusively) τ as a continuous variable, we get:

$$\frac{\partial \varphi}{\partial \tau}(\tau, \theta) = -ln\delta \left(\varphi(\tau, \theta) + (\theta - \hat{\theta}^*_\tau(\theta)) q^{bm}(\theta) \right). \tag{A.63}$$

Because $q^{bm}(\theta)$ is weakly decreasing, we have:

$$\int_{\theta}^{\hat{\theta}^*_\tau(\theta)} q^{bm}(x)dx + (\theta - \hat{\theta}^*_\tau(\theta))(1-\delta^\tau)q^{bm}(\theta) \leq (\hat{\theta}^*_\tau(\theta) - \theta)\delta^\tau q^{bm}(\theta).$$

Therefore,

$$\varphi(\tau,\theta) + (\theta - \hat{\theta}^*_\tau(\theta))q^{bm}(\theta) \leq 0.$$

Inserting into (A.63) yields:

$$\frac{\partial \varphi}{\partial \tau}(\tau,\theta) \leq 0.$$

Hence, (7.4) holds for the Baron-Myerson allocation at any $\tau \geq 0$ when it holds at $\tau = 0$, which means:

$$\text{(A.64)} \quad \psi(\theta, q^{bm}(\theta)) \geq \max_{\hat{\theta}\in\Theta}\left\{\int_{\theta}^{\hat{\theta}} q^{bm}(x)dx\right\} \equiv \int_{\theta}^{\overline{\theta}} q^{bm}(x)dx, \quad \forall\theta\in\Theta.$$

Now, let us define:

$$Z(\theta) = \psi(\theta, q^{bm}(\theta)) - \int_{\theta}^{\overline{\theta}} q^{bm}(x)dx.$$

We get:

$$\dot{Z}(\theta) = (\delta S'(q^{bm}(\theta)) - \theta)\dot{q}^{bm}(\theta) = (\delta(\theta - \underline{\theta}) - (1-\delta)\theta)\,\dot{q}^{bm}(\theta) \geq 0$$

where the last inequality follows from $\delta S'(q^{bm}(\theta)) - \theta \leq 0$ when $q^{bm}(\theta) \geq q^{db}(\theta)$ for all θ (second part of condition (7.5)), and the fact that $q^{bm}(\theta)$ is weakly decreasing. Hence, (A.64) holds if it holds at $\underline{\theta}$ as requested by (7.5). *Q.E.D.*

Proof of Lemma 4: Observe that, for a stationary contract, the maximand on the righthand side of (7.4) can be rewritten as:

$$\varphi(\tau,\theta) = \delta^{-\tau}\max_{\hat{\theta}\in\Theta}\left\{\int_{\theta}^{\hat{\theta}} q(x)dx + (\theta - \hat{\theta})(1-\delta^\tau)q(\theta)\right\}$$

where $q(\theta)$ is weakly decreasing by (3.7). One can again show that the maximand above is achieved at $\hat{\theta}^*_\tau(\theta)$ such that

$$q(\hat{\theta}^*_\tau(\theta)) = (1-\delta^\tau)q(\theta)$$

and that we necessarily have $\hat{\theta}^*_\tau(\theta) > \theta$ at all $\theta < \overline{\theta}$.

Proceeding as in the Proof of Theorem 5, one can show that (7.4) holds for a stationary allocation at any $\tau \geq 0$ when it holds at $\tau = 0$, which can be rewritten as (7.6). *Q.E.D.*

Proof of Proposition 6: The principal's objective is again to maximize the expected discounted net surplus that she withdraws from trade,

$$(\mathcal{P}): \quad \max_{\mathcal{C}}(1-\delta)\sum_{\tau=0}^{\infty}\delta^{\tau}\int_{\underline{\theta}}^{\overline{\theta}}\left(S(q_{\tau}(\theta))-t_{1,\tau}(\theta)-t_{2,\tau}(\theta)\right)dF(\theta)$$

subject to individual rationality, incentive compatibility and enforcement constraints. When restricted to stationary contracts, this problem can be rewritten in terms of per-period payoffs as follows:

$$(\mathcal{P}^{st}): \quad \max_{(U(\theta),q(\theta))\in\mathcal{A}}\int_{\underline{\theta}}^{\overline{\theta}}\left(S(q(\theta))-\theta q(\theta)-U(\theta)\right)\frac{d\theta}{\Delta\theta} \text{ subject to (7.6)}$$

where $\mathcal{A}$ is the set of admissible stationary allocations defined as follows:

Definition A.1 *A stationary allocation $(U(\theta),q(\theta))$ is **admissible** if and only if $q(\theta)$ is piecewise continuous and it satisfies (3.7), (A.62) and (A.61) or equivalently*

$$\dot{U}(\theta)=-q(\theta). \tag{A.65}$$

$(\mathcal{P}^{st})$ is a control problem with a *mixed constraint* (7.6) which links both the state variable U and the control q.[56] The Hamiltonian can be written as:

$$H(\theta,U,q,\lambda)=(S(q)-\theta q-U)\frac{1}{\Delta\theta}-\lambda q$$

where λ is the co-state variable associated with (A.65), while the Lagrangian writes as

$$L(\theta,U,q,\lambda,\mu)=H(\theta,U,q,\lambda)+\mu(\psi(\theta,q)-U)$$

where μ is the multiplier of the mixed constraint (7.6).

First, observe that the Hamiltonian and the expression $\psi(\theta,q)-U$ are both concave in (U,q) for all θ. Hence, the necessary conditions for optimality of a program with mixed constraint (Seierstad and Sydsaeter, 1987, Theorem 1, p. 276) are also sufficient conditions of the Mangasarian type (Seierstad and Sydsaeter, 1987, Theorem 5, p. 287). The admissible pair $(U(\theta),q(\theta))$ is thus optimal if and only if there exist a continuous and piecewise continuously differentiable function $\lambda(\theta)$ and a piecewise continuous function $\mu(\theta)$ such that the following conditions hold.

1. *Optimality w.r.t. output:*

$$(S'(q(\theta))-\theta)\frac{1}{\Delta\theta}+\mu(\theta)(\delta S'(q(\theta)))-\theta)=\lambda(\theta) \quad \forall\theta\in\Theta. \tag{A.66}$$

2. *Condition on the co-state variable:*

$$\dot{\lambda}(\theta)=\frac{1}{\Delta\theta}+\mu(\theta), \quad \text{for a.e. } \theta\in\Theta. \tag{A.67}$$

3. *Boundary and transversality conditions:*

$$\lambda(\underline{\theta})=0 \text{ and } U(\bar{\theta})=0. \tag{A.68}$$

[56] Filippov-Cesari Existence Theorem can be used to prove existence of a solution to this optimal control problem with mixed constraint (Seierstad and Sydsaeter, 1987, Theorem 2, p. 285). Proof available upon request.

4. *Complementarity slackness:*

$$\mu(\theta) \geq 0, \psi(\theta, q(\theta)) - U(\theta) \geq 0 \text{ and } \mu(\theta)(\psi(\theta, q(\theta)) - U(\theta)) = 0 \quad \forall \theta \in \Theta. \tag{A.69}$$

Some important facts follow from the analysis of our problem.

1. Using (A.67) and (A.68) immediately yields the expression of the adjoint variable:

$$\lambda(\theta) = \frac{\theta - \underline{\theta}}{\Delta\theta} + \Psi(\theta) \quad \forall \theta \in \Theta, \text{ where } \Psi(\theta) = \int_{\underline{\theta}}^{\theta} \mu(\xi)\, d\xi. \tag{A.70}$$

2. It is also convenient to rewrite condition (A.66) as:

$$S'(q(\theta))\,(1 + \delta\mu(\theta)\Delta\theta) = 2\theta - \underline{\theta} + \theta\mu(\theta)\Delta\theta + \Psi(\theta)\Delta\theta \quad \forall \theta \in \Theta. \tag{A.71}$$

Whenever (7.6) is slack and $\mu(\theta) = 0$, we get:

$$S'(q(\theta)) = 2\theta - \underline{\theta} + \Psi(\theta)\Delta\theta. \tag{A.72}$$

In particular, continuity of the optimal output at $\hat{\theta}$ implies:

$$S'(\hat{q}) = 2\hat{\theta} - \underline{\theta} + \Psi(\hat{\theta})\Delta\theta. \tag{A.73}$$

3. On any interval with non-empty interior where (7.6) is binding, differentiating w.r.t. θ yields:

$$\dot{q}(\theta)(\delta S'(q(\theta)) - \theta) = 0. \tag{A.74}$$

We now conjecture the form of a solution with $q(\theta)$ being continuous and we then check the optimality conditions above.

- On the interval $\Theta_1 = [\underline{\theta}, \hat{\theta}]$ over which the allocation entails some bunching $q(\theta) = \hat{q}$ (so that (A.74) holds), (7.6) is binding, and $\mu(\theta) \geq 0$. In particular, we have $\mu(\underline{\theta}) \geq 0$. Equations (A.66) and (A.68) then yield:

$$\left(S'(\hat{q}) - \underline{\theta}\right)\frac{1}{\Delta\theta} + \mu(\underline{\theta})\left(\delta S'(\hat{q}) - \underline{\theta}\right) = \lambda(\underline{\theta}) = 0$$

and thus

$$\delta S'(\hat{q}) - \theta \leq \delta S'(\hat{q}) - \underline{\theta} \leq 0 \leq S'(\hat{q}) - \underline{\theta} \quad \forall \theta \in \Theta. \tag{A.75}$$

Finally, (7.6) being binding at $\hat{\theta}$, we get:

$$\psi(\hat{\theta}, \hat{q}) = \int_{\hat{\theta}}^{\overline{\theta}} q(x)dx. \tag{A.76}$$

- On the interval $\Theta_2 = (\hat{\theta}, \overline{\theta}]$, (7.6) is slack and $\mu(\theta) = 0$. From (A.72), we get:

$$S'(q(\theta)) = 2\theta - \underline{\theta} + \Psi(\hat{\theta})\Delta\theta \quad \forall \theta \in \Theta_2 \tag{A.77}$$

so that $q(\theta)$ is weakly decreasing on Θ_2. Second, together with the continuity of $q(\theta)$ at $\hat{\theta}$, it also ensures that

$$S'(\hat{q}) = 2\hat{\theta} - \underline{\theta} + \Psi(\hat{\theta})\Delta\theta \geq 2\theta - \underline{\theta} + \Psi(\hat{\theta})\Delta\theta \quad \forall \theta \leq \hat{\theta}. \tag{A.78}$$

Inserting also $\mu(\theta) = \dot{\Psi}(\theta)$ into (A.71) gives us:

$$(\theta - \delta S'(\hat{q}))\dot{\Psi}(\theta) + \Psi(\theta) = \frac{1}{\Delta\theta}\left(S'(\hat{q}) - \underline{\theta}\right) \quad \forall\theta \in \Theta_1. \tag{A.79}$$

Taking into account that $\Psi(\underline{\theta}) = 0$ yields the solution to this differential equation:

$$\Psi(\theta) = \frac{(S'(\hat{q}) - \underline{\theta})\,(\theta - \underline{\theta})}{(\theta - \delta S'(\hat{q}))\Delta\theta}. \tag{A.80}$$

Differentiating this function, and taking into account the conditions (A.75) yields the expression of the non-negative Lagrange multiplier:

$$\mu(\theta) = \frac{(S'(\hat{q}) - \underline{\theta})\,(\underline{\theta} - \delta S'(\hat{q}))}{(\delta S'(\hat{q}) - \theta)^2\,\Delta\theta} \geq 0 \tag{A.81}$$

Finally, to ensure that the output so constructed is optimal, we need to show that (7.6) also holds for all $\theta \in \Theta_2$. To this end, note that $\psi(\hat{\theta}, q(\hat{\theta})) - U(\hat{\theta}) = 0$ and that the derivative of $\psi(\theta, q(\theta)) - U(\theta)$ w.r.t. θ is equal to $\dot{q}(\theta)\,(\delta S'(q(\theta)) - \theta)$. Because $\dot{q}(\theta) < 0$, this derivative is positive when $q(\theta) \geq q^{db}(\theta)$ for all θ. This gives the result.

We are now ready to describe the solution to $(\mathcal{P}^{st})$.

1. *The enforcement constraint (7.6) is always slack.* Theorem 5 then shows that the Baron-Myerson allocation is optimal.

2. *The enforcement constraint (7.6) is binding on the lower tail.* When (7.7) holds, the Baron-Myerson output no longer satisfies (7.6). There must be some bunching for the efficient types and the optimal solution is of the form (7.8) with $\hat{\theta} > \underline{\theta}$ and $\hat{q} < q^{fb}(\underline{\theta})$. The parameters $\hat{\theta}$, $\Psi(\hat{\theta})$ and $\hat{q}$ are solutions to the system (A.73), (A.76) and (A.80).

Q.E.D.

Proof of Theorem 6 : Starting from any stationary contract with a per-period positive output profile $q_1(\theta)$, let construct a new contract where only date 0-output is modified (referred to as an *almost stationary contract* in the sequel) as:

$$q_0(\theta) = q_1(\theta) - \varepsilon$$

for some $\varepsilon \geq 0$ small enough. Lemma A.2 together with a binding participation constraint (A.62) at $\overline{\theta}$ then fully determines the whole rent profile.

First, observe that the enforcement constraint (7.4) for date 0 can be rewritten as:

$$\psi(\theta, q_1(\theta)) + (1-\delta)\theta\varepsilon \geq \max_{\hat{\theta}\in\Theta}\left\{\int_{\theta}^{\hat{\theta}} q_1(x)dx - (1-\delta)(\hat{\theta} - \theta)\varepsilon\right\}, \quad \forall\theta \in \Theta.$$

The maximum on the above righthand side is still achieved at $\overline{\theta}$ provided $q_1(\theta) + \varepsilon$ remains positive which is true when ε is small enough. Thus (7.4) for date 0 becomes:

$$\psi(\theta, q_1(\theta)) \geq \int_{\theta}^{\overline{\theta}} q_1(x)dx - (1-\delta)\overline{\theta}\varepsilon, \quad \forall\theta \in \Theta. \tag{A.82}$$

Second, (7.4) for all $\tau \geq 1$ is kept unchanged:

$$\psi(\theta, q_1(\theta)) \geq \delta^{-\tau} \max_{\hat{\theta}\in\Theta} \left\{ \int_{\theta}^{\hat{\theta}} q_1(x)dx - (1-\delta)\varepsilon(\hat{\theta}-\theta) + (1-\delta)(\theta-\hat{\theta})q_1(\theta) - (1-\delta)\varepsilon(\hat{\theta}-\theta) \right\}$$

$$= \delta^{-\tau} \max_{\hat{\theta}\in\Theta} \left\{ \int_{\theta}^{\hat{\theta}} q_1(x)dx + (1-\delta)(\theta-\hat{\theta})q_1(\theta) \right\} \quad \forall\theta \in \Theta.$$

Using an argument similar to one made in the proof of Theorem 5, the above righthand side is maximum for $\tau = 1$ so that (7.4) for date 1 on can be rewritten as:

$$\text{(A.83)} \quad \psi(\theta, q_1(\theta)) \geq \varphi(1,\theta) = \delta^{-1} \max_{\hat{\theta}\in\Theta} \left\{ \int_{\theta}^{\hat{\theta}} q_1(x)dx + (1-\delta)(\theta-\hat{\theta})q_1(\theta) \right\} \quad \forall\theta \in \Theta.$$

Still from the proof of Theorem 5, we know that

$$\int_{\theta}^{\overline{\theta}} q^{st}(x)dx \geq \delta^{-1} \max_{\hat{\theta}\in\Theta} \left\{ \int_{\theta}^{\hat{\theta}} q^{st}(x)dx + (1-\delta)(\theta-\hat{\theta})q^{st}(\theta) \right\} \quad \forall\theta \in \Theta$$

with a strict inequality on an interval $[\underline{\theta}, \hat{\theta}+\eta]$ for η small enough. Hence the most stringent enforcement constraint between (A.82) and (A.83) is still (A.82) when ε is small enough. Taking stock of this observation, expressing the intertemporal rent as $\overline{U}_0(\theta) = \int_{\theta}^{\overline{\theta}}((1-\delta)q_0(x) + \delta q_1(x))dx = \int_{\theta}^{\overline{\theta}} q_1(x)dx - (1-\delta)\varepsilon(\overline{\theta}-\theta)$ and integrating by parts, we can rewrite the principal's problem as when restricted to almost stationary contracts as:

$$\max_{(\varepsilon, q_1(\theta))} (1-\delta) \int_{\underline{\theta}}^{\overline{\theta}} (S(q_1(\theta)-\epsilon) - S(q_1(\theta)) - (2\theta-\underline{\theta})\varepsilon) \frac{d\theta}{\Delta\theta} + \int_{\underline{\theta}}^{\overline{\theta}} (S(q_1(\theta)) - (2\theta-\underline{\theta})q_1(\theta))\, d\theta$$

subject to (A.82).

If the optimal mechanism is stationary the solution to this problem is $(\varepsilon = 0, q_1(\theta) = q^{st}(\theta))$. Let us first fix ε and denote by $\mathcal{W}(\varepsilon)$ the objective function so obtained after optimization w.r.t. $q_1(\theta)$. We evaluate the derivative $\mathcal{W}'(\varepsilon)$ of this objective w.r.t. ε at $\varepsilon = 0$ and thus $q_1(\theta) = q^{st}(\theta)$. A simple application of the Envelope Theorem yields:

$$\mathcal{W}'(0) = -(1-\delta) \left(\int_{\underline{\theta}}^{\overline{\theta}} \left(S'(q^{st}(\theta)) - (2\theta-\underline{\theta})\right) \frac{d\theta}{\Delta\theta} - \Psi(\hat{\theta})\overline{\theta} \right).$$

Taking into account the expression of $q^{st}(\theta)$ yields:

$$\text{(A.84)} \quad \mathcal{W}'(0) = -(1-\delta) \left(\int_{\underline{\theta}}^{\hat{\theta}} 2\left(\hat{\theta}-\theta\right) \frac{d\theta}{\Delta\theta} - \Psi(\hat{\theta})\underline{\theta} \right)$$

From (A.80) and (A.78), we get:

$$\Psi(\hat{\theta}) \left(\underline{\theta} - \delta(2\hat{\theta} - \underline{\theta} + \Psi(\hat{\theta})\Delta\theta) \right) = 2\frac{(\hat{\theta}-\underline{\theta})^2}{\Delta\theta}.$$

Inserting into (A.84) yields:

$$\mathcal{W}'(0) = \frac{(1-\delta)}{2}\Psi(\hat{\theta}) \left(\underline{\theta} + \delta(2\hat{\theta} - \underline{\theta} + \Psi(\hat{\theta})\Delta\theta) \right) > 0.$$

Hence, offering a quasi-stationary contract with $q_0(\theta) < q^{st}(\theta)$ improves the buyer's intertemporal welfare.

Q.E.D.