Economics Letters 159 (2017) 18–22

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Optimal stationary contract with two-sided imperfect enforcement and persistent adverse selection

David Martimort [a], Aggey Semenov [b,*], Lars Stole [c]

[a] Paris School of Economics-EHESS, France
[b] Economics Department, University of Ottawa, 120 University private, Ottawa, ON, K1N 6N5 Canada
[c] University of Chicago Booth School of Business, United States

HIGHLIGHTS

- The optimal stationary contract in an infinitely repeated relationship is proposed.
- The contract is made of two distinct pieces.
- For the most efficient types of the agent, the contract entails bunching.
- For less efficient types, the contract exhibits downward output distortions.
- Distortions are set below the Baron–Myerson level.

ARTICLE INFO

Article history:
Received 17 March 2017
Received in revised form 28 June 2017
Accepted 6 July 2017
Available online 14 July 2017

JEL classification:
D82
D86

Keywords:
Contract enforcement
Optimal control
Adverse selection
Stationary contract

ABSTRACT

We consider an infinitely-repeated principal–agent relationship run with stationary contracts. The agent has private information on his persistent cost parameter and, under limited enforcement, both parties can breach the contract. The optimal stationary contract with limited enforcement is made of two distinct pieces. For the most efficient types of the agent, the contract entails bunching with a fixed payment and a fixed output. For less efficient types, the contract exhibits downward output distortions below the Baron–Myerson level that would have been achieved had enforcement been costless.

1. Introduction

We consider an infinitely repeated principal–agent relationship. The agent has private information on his cost parameter. This cost is persistent over time and drawn once for all from a continuous distribution. The contract can be breached at any point in time by either party. The principal may choose not to pay once delivered the good. The agent may choose not to deliver the requested quantity/quality of the good to be traded. Enforcement of the contract is restricted to the use of limited penalties for breach on the party who does not fulfill his obligations. On top, future trades are disrupted following a breach by either party.

Similar settings have already been studied in the literature. In a framework where no formal contract can ever be enforced, Levin (2003) considered a symmetric information setting (with either complete information or types which are i.i.d. over time). He showed that the optimal relational contract is stationary and may entail output and wage compression. Baron and Besanko (1984) studied a repeated contracting environment with full commitment and perfect enforcement. They showed that there is no loss of generality in considering stationary contracts and that the optimal long-term contract is indeed the replica of the well-known static optimum found in Baron and Myerson (1982). Allowing for limited enforcement but formal contracting, Martimort et al. (2017) considered the case of a two-type discrete distribution. We showed there that the optimal contract is actually non-stationary when enforcement constraints are binding and that long-run distortions encapsulate the shadow cost of limited enforcement. Kwon (2016) derives the optimal relational contract with persistent adverse selection in Levin (2003)'s environment and shows that it is no longer stationary. The non-stationarity of relational contracts may

* Corresponding author.
E-mail address: Aggey.Semenov@uottawa.ca (A. Semenov).

http://dx.doi.org/10.1016/j.econlet.2017.07.007

also come from learning persistent types as in a model of the labor market proposed by Yang (2013).

Unfortunately, in some structured environments, such non-stationary profiles of payments and outputs may be hard to implement. To illustrate, consider the ongoing relationships between an upstream manufacturer (the principal) and a collection of downstream retailers (her agents). If the retailers initiate contracts with the principal at different points in time, then even if each of those relationships might *a priori* be ruled under different time-varying contracts, a prohibition against discriminatory contracts at each point in time would force the use of stationary contracts throughout the whole vertical structure.[1] Similarly, when there are possibilities for the downstream retailers to further trade and subcontract with each other, arbitrage opportunities would make it difficult to credibly enforce trades at different prices with different retailers providing the same quality good. As an empirical matter, stationarity in vertical relationships is not uncommon. Lafontaine and Shaw (1999) found in their study of vertical franchise contracts, for example, that royalty rates and franchise fees do not change over time.

In this note, we characterize the optimal enforceable stationary contract and analyze its main properties. In a model with continuous types, we show that, under a mild assumption on the distribution of types that generalizes the well-known monotonicity of the hazard rate property, that the optimal contract, when constrained by enforcement, has two main features. First, bunching arises for the most efficient types who all produce the same quantity and receive the same payment. Second, less efficient types are separated, though at outputs below the Baron–Myerson allocation that would be achieved had enforcement been costless. Although apparently similar to Levin (2003)'s result on wage/output compression, our results are significantly different and hinge on a more subtle trade-off. Indeed, in Levin (2003), the fact that there is symmetric information between the principal and the agent means that the enforcement constraint can only be satisfied provided that current payments are small compared with the continuation value of the relationship. Since types are i.i.d. over time, this continuation value is "fixed" and the enforcement constraint is akin to an exogenous upper bound on payments (as in Thomas (2002)). Instead, in our screening scenario, payments play a second role as a screening instrument. Higher payments help to induce information revelation from the most efficient types of the agent. With imperfect enforcement, the principal mediates the desire to raise payments for screening reasons against the incentives to renege on such large promises. Wages and output compressions follow from this trade-off.

2. Model

We consider an infinitely-repeated trading environment between a principal (she, the buyer) and an agent (he, the seller). Time is indexed by $\tau = 0, 1, \ldots, \infty$, and $\delta < 1$ is the common discount factor. In each period, a quantity q_τ can be traded. Before (resp. after) trade takes place, the principal makes some payment $t_{1,\tau}$ (resp. $t_{2,\tau}$). The agent has private information on his cost parameter θ which is drawn once for all before the relationship starts. A contract is thus an array $\mathcal{C} = \{(t_{1,\tau}(\theta), t_{2,\tau}(\theta), q_\tau(\theta))_{\theta\in\Theta}\}_{\tau=0,\ldots,\infty}$ that stipulates for each trading period payments (respectively before and after current trade) $t_{1,\tau}(\theta)$, $t_{2,\tau}(\theta)$ and a quantity $q_\tau(\theta)$ that are contingent on the agent's report θ on his cost parameter.

The principal's and the agent's utility functions are respectively given by:

$$(1-\delta)\sum_{\tau=0}^{\infty}\delta^\tau(S(q_\tau(\theta)) - t_{1,\tau}(\theta) - t_{2,\tau}(\theta)) \text{ and}$$

$$(1-\delta)\sum_{\tau=0}^{\infty}\delta^\tau(t_{1,\tau}(\theta) + t_{2,\tau}(\theta) - \theta q_\tau(\theta)).$$

We assume that $S'(\cdot) > 0 > S''(\cdot)$ with $S(0) = 0$ and where $S'(0)$ is also large enough to avoid corner solutions and avoid shut-down of the least-efficient types. The principal only knows the cumulative distribution $F(\theta)$ whose support is $\Theta = [\underline{\theta}, \bar{\theta}]$ and whose positive and atomless density is denoted by $f(\cdot) = F'(\cdot)$.

We consider stationary contracts of the form: $q_\tau(\theta) = q(\theta)$, $t_{1,\tau}(\theta) = t_1(\theta)$ and $t_{2,\tau}(\theta) = t_2(\theta)$ for all τ. The timing of the contracting game unfolds as follows.

1. At date $\tau = 0^-$ the agent learns his cost parameter θ. The principal offers a contract $\mathcal{C} = \{t_1(\theta), t_2(\theta), q(\theta)\}_{\theta\in\Theta}$. The agent accepts or rejects $\mathcal{C}$. If the agent rejects, then both parties get their reservation values that are normalized at zero. If the agent accepts, he reports having a type $\hat{\theta}$.
2. At any date $\tau \geq 0$, trade takes place. First, the principal pays an advance payment $t_1(\hat{\theta})$. Second, the agent produces $q(\hat{\theta})$. If he does not deliver this requested quantity, the contract is breached and the agent must pay the penalty L. If $q(\hat{\theta})$ is delivered, the principal pays the after-sale payment $t_2(\hat{\theta})$. If she does not, the contract is again breached and the principal pays the penalty K. Following breach by either party, the contract is terminated.[2]

3. Enforcement and incentive compatibility

Denote by $U(\theta)$ the agent's average per-period payoff with a stationary contract, i.e., $U(\theta) = t_1(\theta) + t_2(\theta) - \theta q(\theta)$ where $t(\theta) = t_1(\theta) + t_2(\theta)$. By the Revelation Principle it is without the loss of generality to focus on contracts that induce incentive compatible, stationary allocations $(q(\theta), U(\theta))$. A standard argument characterizes incentive compatible allocations in this environment[3]:

Lemma 1. *An allocation $(q(\theta), U(\theta))_{\theta\in\Theta}$ is incentive compatible if and only if $U(\theta)$ is absolutely continuous, convex and satisfies at any point of differentiability (i.e., almost everywhere)*

$$\dot{U}(\theta) = -q(\theta), \tag{3.1}$$

$$q(\theta) \text{ is non-negative and non-increasing.} \tag{3.2}$$

Eq. (3.1) implies that U is a non-increasing function. Hence, a contract induces participation for all types if it does so for the least-efficient one, namely:

$$U(\bar{\theta}) \geq 0. \tag{3.3}$$

[1] Antitrust authorities have repeatedly addressed price discrimination in intermediate-good markets with suspicion. In the U.S., the Robinson–Patman was enacted to put on equal foot small businesses and large buyers in intermediate-good markets. Although this more complex motivation is not part of our model, it justifies our focus on non-discriminatory contracts.

[2] The no-renegotiation assumption can be justified on several grounds. First, it allows us to compute an upper bound on the possible gains from trade that can be achieved under asymmetric information and limited enforcement. Second, as an empirical matter, it is not uncommon for contractual breaches to lead to the termination of a productive relationship without any attempt at renegotiation. For example, such termination is commonly observed in construction contracts. This behavior may represent an equilibrium of a richer reputation game which we choose to leave unmodeled in this note. Finally, as a practical matter, the assumption of commitment to no-renegotiation is a standard assumption in the law and economics literature (see Edlin, 1998 and Shavell, 2004, p. 315), and an obvious starting point for richer studies.

[3] The proof is standard and is thus omitted. See for example Laffont and Martimort (2002).