RATIONALITY & COMPETITION
CRC TRR 190

Deterministic versus Stochastic Contracts in a Dynamic Principal-Agent Model

Thomas Mettral (HU Berlin)

Discussion Paper No. 93

April 24, 2018

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

Deterministic versus stochastic contracts in a dynamic principal-agent model

Thomas Mettral*

March 2018

Abstract

I show that deterministic dynamic contracts between a principal and an agent are always at least as profitable to the principal as stochastic ones, if the so-called first-order approach in dynamic mechanism design is satisfied. The principal commits, while the agent's type evolution follows a Markov process. My results demonstrate, even when allowing for potential correlation of stochastic contracts across periods that the usual restriction in the literature to deterministic contracts is admissible, as long as the first-order approach is valid.

JEL Code: D82, D86

Keywords: Contract theory, Principal-agent theory, Dynamic contracting

1 Introduction

In recent years, there has been an increased interest in dynamic mechanism design, e.g. Courty and Li (2000), Battaglini (2005), Pavan et al. (2009), Kapicka (2010), Gershkov and Perry (2012), Eső and Szentes (2013), Li and Shi (2013),

*Humboldt-Universität zu Berlin, Institute for Economic Theory 1, Spandauer Str. 1, D-10178 Berlin (Germany), thomas.schacherer@hu-berlin.de. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 is gratefully acknowledged.

Pavan et al. (2014), Battaglini and Lamba (2017), Deb and Said (2015) and Krähmer and Strausz (2015) discuss this issue. All these papers, however, restrict to deterministic mechanisms accepting that this assumption is often with loss of generality. Moreover, most of these papers use the local approach to characterize optimal mechanisms, the so-called first-order approach, which means that only local downward binding IC-constraints have to be taken into account.

Extending Strausz (2006) to a dynamic framework, I show that the ad hoc restriction to deterministic contracts is without loss valid if the first-order approach is valid.

The extension is not immediate, because stochastic mechanisms in a dynamic framework also allow for intertemporal correlation, an issue which in a static framework does not arise.[1]

2 Model

There are two players, a principal and an agent. In each period $t \in \mathcal{T} := \{1, \dots, T\}$, $T \geqslant 2$,[2] the agent consumes a quantity $q_t \in \mathbb{R}_+$ at some price $p_t \in \mathbb{R}$. This generates a per-period utility of $u(\theta_t, q_t) - p_t$ for the agent, where $\theta_t \in \Theta := \{\theta_N, \dots, \theta_0\} \subset \mathbb{R}$ represents agent's type in period $t \in \mathcal{T}$. I follow the standard assumptions in the literature that u is twice continuously differentiable in both arguments, increasing in both arguments, with $u(\cdot, 0) = 0$, is concave in q_t and satisfies the single crossing condition, i.e. marginal utility is higher for higher types. The principal produces q_t given a cost function $c(q_t)$. This function fulfills as well usual conditions. There are no fixed costs, it is twice continuously differentiable, increasing and convex. To guarantee an interior solution, I assume that marginal costs vanish at 0 and tend to infinity if the quantity tends to infinity.

In the first period, the principal commits to a long term contract to the agent who has the opportunity to accept or reject it. In every later period $t \in \mathcal{T} \backslash \{1\}$, he decides to continue or to terminate the relationship. Once the agent terminates

[1] Pavan et al. (2014) in Corollary 2 (iv) mention without formal proof that results of Strausz (2006) imply an optimality of deterministic contracts, but they neglect the possibility of intertemporal correlation.

[2] It is not important for the analysis if T is finite or not. The results still hold for $T = \infty$, the proofs become however more extensive.

the contract, he has no possibility to rejoin the contract.

2.1 Basic Assumptions

For notational convenience, I assume that agent's types are equidistant, i.e. $\Delta\theta := \theta_{i-1} - \theta_i > 0$ for all $i \in I\backslash\{0\}$, where $I := \{0, \dots, N\}$ is the set of all indices of types.[3] The initial type of the agent is chosen from a prior distribution $f(\theta_i) =: \mu_i \in]0, 1[$ for all $i \in I$, with $\sum_{i\in I} \mu_i = 1$, which is common knowledge. Its cumulative distribution function is therefore $F(\theta_i) = \sum_{j=i}^{N} \mu_j$, for all $i \in I$. In all later periods the type changes according to a Markov process. The probability that the agent's type changes from θ_i to θ_j is given through $f(\theta_j|\theta_i) =: \alpha_{ij} \in]0, 1[$, for all $i, j \in I$ and for every period $t \in \mathcal{T}$. This reflects the Markov property of independence regarding time and earlier types. It fulfills $\sum_{j=0}^{N} \alpha_{ij} = 1$, for all $i \in I$ and for simplicity, I assume full support of the conditional distribution, i.e. $\alpha_{ij} > 0$ for all $i, j \in I$. The corresponding cumulative distribution function F is given through $F(\theta_k|\theta_i) = \sum_{j=k}^{N} \alpha_{ij}$, for all $i, k \in I$. I also follow the usual convention of first order stochastic dominance, i.e. $F(\theta_k|\theta_i) \geqslant F(\theta_k|\theta_{i-1})$ or $0 \leq \Delta F(\theta_k|\theta_i) := F(\theta_k|\theta_i) - F(\theta_k|\theta_{i-1})$, for all $k \in I$ and all $i \in I\backslash\{0\}$.

In the following, I use the notation θ_t to characterize the agent's type in period $t \in \mathcal{T}$.[4] Moreover, let $\theta^t \in \Theta^t$ be the evolution vector $\theta^t := (\theta_1, \dots, \theta_t)$ of agent's types from period 1 up to period t, for all $t \in \mathcal{T}$. The whole type path is denoted by $\theta := \theta^T \in \Theta^T$. In addition, let $\Theta^{t+\tau}(\theta^t) := \{\vartheta^{t+\tau} \in \Theta^{t+\tau} : \vartheta_s = \theta_s,\ \forall\ 1 \leqslant s \leqslant t\}$, for all $t \in \mathcal{T}$, all $\theta^t \in \Theta^t$ and all $0 \leqslant \tau \leqslant T - t$. Furthermore, let $q^t := (q_1, \dots, q_t) \in \mathbb{R}^t_+$ be the vector of quantity realizations and $p^t := (p_1, \dots, p_t) \in \mathbb{R}^t$ the price-vector with $p_t = p(q_t)$, each from period 1 up to period $t \in \mathcal{T}$, where $q := q^T$, $p := p^T$ are the corresponding vectors over the whole time horizon $\mathcal{T}$. By the revelation principle, it suffices that q_t and p_t depend on the current report θ_t and earlier reports and realizations. Recursively, one can denote q_t as the occurred realization of $q(\theta_t|q^{t-1}, \theta^{t-1})$ for all $t \in \mathcal{T}$, whereby $q^0, \theta^0 \in \emptyset$.

[3] As in Strausz (2006), I assume a finite number of types to circumvent measure theoretical complications.

[4] The notation θ_t characterizes the stochastic process of agent's type which takes values in Θ, whereas θ_i specifies a possible event of agent's type in any period. Therefore, expressions like θ_1 are ambiguous, but it should become clear in the specific situation.

2.2 Stochastic contracts

In order to represent stochastic contracts, I distinguish between the realized quantity q_t and the random variable $q(\theta_t|h^{t-1})$, which depends on agent's report θ_t in the current period and the history h^{t-1} of previous reports θ^{t-1} and quantity realizations q^{t-1}. Here, I use $h^t := (\theta^t, q^t)$ the history of previous types and occurred realizations with $h^t \in H^t := \Theta^t \times \mathbb{R}_+^t$, for all $t \in \mathcal{T}$ and let $h^0 \in H^0 := \emptyset$. Therefore, $q(\theta_t|h^{t-1})$ defines on the image space $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ the implementation function

$$\begin{aligned} \xi(\cdot|h^{t-1}, \theta_t) : \qquad & \mathbb{R}_+ \longrightarrow [0,1], \\ \xi(q_t|h^{t-1}, \theta_t) \quad & = \quad \mathbb{P}(q \leqslant q_t|h^{t-1}, \theta_t), \end{aligned}$$

for all $q_t \in \mathbb{R}_+$.

Indeed, the principal can choose the weights of possible outcomes over $\mathbb{R}_+$ of the implementation function depending on the history of type reports θ^{t-1}, the current report θ_t and the history of previous realized quantities q^{t-1}. This, however, creates in addition to the reports of agent's type, a second uninformative channel for both, the agent and the principal.[5] Furthermore, it allows for interdependences between the random variables over several periods. I use the notation

$$\xi_{\theta^t}(q_t|q^{t-1}) := \xi(q_t|h^{t-1}, \theta_t), \tag{1}$$

which illustrates the dependence of ξ of current and previous reports. With Bayes' rule and the fact that q^{t-1} is independent of θ_t one obtains

$$\mathrm{d}\xi_{\theta^t}(q_t|q^{t-1}) \ldots \mathrm{d}\xi_{\theta^1}(q_1) = \mathrm{d}\xi_{\theta^t}(q^t),$$

for all $t \in \mathcal{T}$. Hence, ξ_θ reflects the implementation function of the whole allocation vector $q \in \mathbb{R}_+^T$.

[5] I assume that prices $p(q_t)$ are deterministic, which is due to quasi-linear utilities without loss of generality.

2.3 Agent's continuation utility

After signing the contract, the agent receives in every period $t \in \mathcal{T}$ a quantity $q_t \in \mathbb{R}_+$ chosen from a lottery for a price $p_t \in \mathbb{R}$. Moreover, he discounts future utilities by $\delta \in]0,1[$. Therefore, one can define his continuation utility recursively as

Definition 1. The agent's continuation utility under truth-telling in period $t \in \mathcal{T}$ is given through

$$U(\theta_t|h^{t-1}) \\ := \int_0^\infty \left(u(\theta_t, q_t) - p_t + \delta \sum_{\theta_{t+1} \in \Theta} f(\theta_{t+1}|\theta_t) U(\theta_{t+1}|h^{t-1}, \theta_t, q_t) \right) \mathrm{d}\xi_{\theta^t}(q_t|q^{t-1}).$$

2.4 Timing

The time structure is as follows. At the beginning, the agent learns his initial type $\theta_1 \in \Theta$. Then, the principal offers a contract $\{p, \xi_\theta\}$ or equivalently $\{U, \xi_\theta\}$, which incorporates in every period t all possible type reports θ_t of the agent and all possible histories $h^{t-1} \in H^{t-1}$. U represents the vector $U = (U(\theta_1|h^0), \ldots, U(\theta_T|h^{T-1}))$ of agent's continuation utility. After the contract proposal, the agent decides whether to accept or reject the offer. If he accepts, he gives in a report θ_1 and $\xi_{\theta^1}(q^1)$ is realized. In the beginning of every later period $t > 1$, the agent learns his new type drawn from $f(\theta_t|\theta_{t-1})$ and decides to continue or terminate the contract. If he continues, he gives in a new report θ_t and $\xi_{\theta^t}(q_t|q^{t-1})$ is realized.

Since in every period, the agent can terminate the contract, the principal has to take into account the IR-constraints in every period. If the agent terminates, he cannot resume to the contract, therefore the IR-constraint $\mathrm{IR}(\theta_t|h^{t-1})$ can be described as

$$U(\theta_t|h^{t-1}) \geqslant 0, \tag{2}$$

for all $\theta_t \in \Theta$, all $h^{t-1} \in H^{t-1}$ and all periods $t \in \mathcal{T}$.

For the IC-constraints, in every period $t \in \mathcal{T}$, the principal has to give incentives to the agent to report his true type $\theta_t \in \Theta$ instead of any other type $\vartheta_t \in \Theta$. Since the history-path h^{t-1} only depends on previous type reports and not on previous true types, the IC-constraint $\mathrm{IC}(\theta_t, \vartheta_t|h^{t-1})$ can be characterized by

$$\begin{aligned}U(\theta_t|h^{t-1}) \geqslant U(\vartheta_t|h^{t-1}) + \int_0^\infty \big(u(\theta_t, q_t) - u(\vartheta_t, q_t)\big)\mathrm{d}\xi_{(\theta^{t-1},\vartheta_t)}(q_t|q^{t-1}) \\ + \ \delta \sum_{\theta_{t+1}\in\Theta} \big(f(\theta_{t+1}|\theta_t) - f(\theta_{t+1}|\vartheta_t)\big) \int_0^\infty U(\theta_{t+1}|h^{t-1}, \vartheta_t, q_t)\mathrm{d}\xi_{(\theta^{t-1},\vartheta_t)}(q_t|q^{t-1}),\end{aligned} \tag{3}$$

for all $\theta_t, \vartheta_t \in \Theta$, all $h^{t-1} \in H^{t-1}$ and all periods $t \in \mathcal{T}$. Note that only one time deviations have to be considered since after any deviation to ϑ_t, the highest future continuation utility is given by $U(\theta_{t+1}|h^{t-1}, \vartheta_t, q_t)$ if all future IC-constraints are fulfilled.

Given these inequalities the principal's objective is to maximize her expected surplus, i.e.

$$\max_{\{U,\xi_\theta\}} \left\{ \sum_{\theta_1\in\Theta} f(\theta_1)\big(S(\theta_1) - U(\theta_1)\big) \right\}, \tag{4}$$

s.t. (2) and (3) are satisfied, whereby

$$S(\theta_t|h^{t-1}) := \int_0^\infty \left(s(\theta_t, q_t) + \delta \sum_{\theta_{t+1}\in\Theta} f(\theta_{t+1}|\theta_t) S(\theta_{t+1}|h^{t-1}, \theta_t, q_t) \right) \mathrm{d}\xi_{\theta^t}(q_t|q^{t-1}) \tag{5}$$

is the aggregated continuation surplus and $s(\theta_t, q_t) := u(\theta_t, q_t) - c(q_t)$ the per-period aggregated surplus in period t, for all $t \in \mathcal{T}$, with $S(\theta_{T+1}|h^T) := 0$, for all histories $h^T \in H^T$.

3 Optimal contracting under the first-order approach

As in Battaglini and Lamba (2017), I define the first-order approach as follows:

Definition 2. A contract is first-order optimal if and only if it is sufficient to consider the relaxed problem, including only $\{\mathrm{IR}(\theta_t = \theta_N|h^{t-1})\}_{t\in\mathcal{T}}$ and $\{\mathrm{IC}(\theta_t = \theta_i, \vartheta_t = \theta_{i+1}|h^{t-1})\}_{t\in\mathcal{T}}$, for all $i \in I\backslash\{N\}$, and the other constraints can be disregarded.

Following now the same arguments as in Battaglini and Lamba (2017), I get the following Lemma, which differs only to their result by allowing for stochastic contracts.

Lemma 1. *In the relaxed problem, the principal's objective* (4) *simplifies to*

$$\sum_{\theta_1 \in \Theta} f(\theta_1)\big(S(\theta_1) - U(\theta_1)\big) = \sum_{\theta \in \Theta^T} \prod_{s=0}^{T-1} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^T} V(\theta, q) \mathrm{d}\xi_\theta(q), \tag{6}$$

where $V(\theta, q) := \sum_{\tau=0}^{T-1} \delta^\tau v(\theta_{\tau+1}, q_{\tau+1})$ *captures the virtual surplus over the whole time horizon* $\mathcal{T}$ *depending on reported types* θ *and occurred realizations of quantities* q *and*

$$v(\theta_\tau, q_\tau) := s(\theta_\tau, q_\tau) - \frac{1 - F(\theta_1)}{f(\theta_1)} \prod_{s=1}^{\tau-1} \frac{\Delta F(\theta_{s+1}|\theta_s)}{f(\theta_{s+1}|\theta_s)} \Delta u(\theta_\tau, q_\tau)$$

denotes the virtual surplus in period $\tau \in \mathcal{T}$*.*

With this representation, principal's objective simplifies to a maximization problem of V with respect to ξ_θ, which allows for any kind of mixing across periods. Given that such a representation of principal's objective exists, the static proof of Strausz (2006) extends to dynamic environments, i.e. the principal gets the maximal profit if she maximizes V with respect to q for every given $\theta \in \Theta^T$. Hence, for any $\hat{q} \in \arg\max_{q \in \mathbb{R}_+^T} V(\theta, q)$, a contract with implementation function $\hat{\xi}_\theta(q)$ that is equal to 1 if $q \geqslant \hat{q}$ maximizes principal's objective, i.e.

$$\begin{aligned} & \sum_{\theta \in \Theta^T} \prod_{s=0}^{T-1} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^T} V(\theta, q) \mathrm{d}\xi_\theta(q) \\ \leqslant \quad & \sum_{\theta \in \Theta^T} \prod_{s=0}^{T-1} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^T} V(\theta, q) \mathrm{d}\hat{\xi}_\theta(q) \\ = \quad & \sum_{\theta \in \Theta^T} \prod_{s=0}^{T-1} f(\theta_{s+1}|\theta_s) V(\theta, \hat{q}). \end{aligned}$$

Hence, stochastic contracts are at most as profitable for the principal as deterministic contracts. This result is summarized in

Proposition 1. Consider a dynamic setting with $T < \infty$ periods in which the first-order approach holds. Then, deterministic contracts are always superior than stochastic contracts.

The idea of the proof is as follows. Since the principal has full commitment to her initially offered contract, she cannot react to history $h^{t-1} \in H^{t-1}$ in any later period $t \geqslant 2$. Therefore, the principal maximizes her expected discounted sum of virtual surpluses $V(\theta, q)$ with respect to $q \in \mathbb{R}_+^T$. Hence, she always prefers to choose such quantities that maximize the expectation of $V(\theta, q)$ like $\hat{q} \in \mathbb{R}_+^T$. If there are multiple maximizers, she could randomize between them, but still, the deterministic quantity $\hat{q}$ would provide at least the same surplus to the principal.

Battaglini and Lamba (2017), however, already mention that the first-order approach is often not justified, and they state the optimal deterministic contract in a specific but enlightening example, which is even optimal in the wider set of all stochastic contracts. In a more general setup, however, it could be with loss of generality to restrict to deterministic contracts only.

4 Conclusion

This paper shows that stochastic contracts do not yield higher profits to the principal in dynamic contracting, if the first-order approach is valid. In situations for which the first-order approach does not work, it remains an open question whether stochastic contracts could yield higher profits to the principal. However, a proper analysis of stochastic contracts in such environments is complicated, since already no characteristic result of optimal deterministic contracts exists when the first-order approach fails.

5 Appendix

To prove Lemma 1, I show first two necessary Lemmata:

Lemma 2. *If the first-order approach is valid, the agent's continuation utility $U(\theta_t|h^{t-1})$ has the explicit representation*

$$U(\theta_t = \theta_i|h^{t-1}) = \sum_{j=i+1}^{N} \sum_{\tau=0}^{T-t} \delta^\tau \sum_{\theta^{t+\tau} \in \Theta^{t+\tau}(\theta^{t-1}, \theta_j)} \prod_{s=t}^{t+\tau-1} \Delta F(\theta_{s+1}|\theta_s) \cdot \int_{\mathbb{R}_+^{\tau+1}} \Delta u(\theta_{t+\tau}, q_{t+\tau}) \, \mathrm{d}\xi_{\theta^{t+\tau}}(q_{t+\tau}, \ldots, q_t|q^{t-1}),$$

for all $i \in I$ and all $t \in \mathcal{T}$.

Proof of Lemma 2. Let $t \in \mathcal{T}$, and $h^{t-1} \in H^{t-1}$ be an arbitrary history-path. Under the first-order approach, the IR-constraint is always binding for θ_N, i.e.

$$U(\theta_t = \theta_N | h^{t-1}) = 0.$$

Moreover, the IC-constraints are downward binding, i.e.

$$U(\theta_t = \theta_i | h^{t-1}) = U(\theta_t = \theta_{i+1} | h^{t-1}) + \int_0^\infty \Delta u(\theta_t = \theta_{i+1}, q_t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_{i+1})}(q_t | q^{t-1})$$
$$+\delta \sum_{k=0}^{N} (\alpha_{ik} - \alpha_{(i+1)k}) \int_0^\infty U(\theta_{t+1} = \theta_k | h^{t-1}, \theta_t = \theta_{i+1}, q_t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_{i+1})}(q_t | q^{t-1}),$$

for all $i \in I \backslash \{N\}$. Plugging in recursively all binding IC-constraints for all $i < j < N$, and the binding IR-constraint for θ_N, one obtains

$$U(\theta_t = \theta_i | h^{t-1}) = \sum_{j=i+1}^{N} \int_0^\infty \Delta u(\theta_t = \theta_j, q_t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_j)}(q_t | q^{t-1})$$
$$+ \sum_{j=i+1}^{N} \delta \sum_{k=0}^{N} (\alpha_{(j-1)k} - \alpha_{jk}) \int_0^\infty U(\theta_{t+1} = \theta_k | h^{t-1}, \theta_t = \theta_j, q_t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_j)}(q_t | q^{t-1}),$$

for all $t \in \mathcal{T}$, and all histories $h^{t-1} \in H^{t-1}$, whereby $U(\theta_{T+1} | h^T) := 0$ for all histories $h^T \in H^T$. Now, I show the explicit representation of $U(\theta_t = \theta_i | h^{t-1})$ by means of backward induction. The basis for $t = T$ is given through the last equality. For the inductive step for $t+1$ to t, one has

$$U(\theta_t = \theta_i | h^{t-1}) = \sum_{j=i+1}^{N} \int_0^\infty \Delta u(\theta_t = \theta_j, q_t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_j)}(q_t | q^{t-1})$$
$$+ \sum_{j=i+1}^{N} \delta \sum_{k=0}^{N} (\alpha_{(j-1)k} - \alpha_{jk})$$
$$\cdot \int_0^\infty \sum_{l=k+1}^{N} \sum_{\tau=0}^{T-(t+1)} \delta^\tau \sum_{\theta^{t+\tau+1} \in \Theta^{t+\tau+1}(\theta^{t-1}, \theta_j, \theta_l)} \prod_{s=t+1}^{t+\tau} \Delta F(\theta_{s+1} | \theta_s)$$
$$\cdot \int_{\mathbb{R}_+^{\tau+1}} \Delta u(\theta_{t+\tau+1}, q_{t+\tau+1}) \, \mathrm{d}\xi_{\theta^{t+\tau+1}}(q_{t+\tau+1}, \ldots, q_{t+1} | q^t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_j)}(q_t | q^{t-1})$$

$$
\begin{aligned}
= \quad & \sum_{j=i+1}^{N} \int_0^\infty \Delta u(\theta_t = \theta_j, q_t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_j)}(q_t | q^{t-1}) \\
& + \sum_{j=i+1}^{N} \sum_{\tau=1}^{T-t} \delta^\tau \sum_{l=0}^{N} \Delta F(\theta_l | \theta_j) \sum_{\theta^{t+\tau} \in \Theta^{t+\tau}(\theta^{t-1}, \theta_j, \theta_l)} \prod_{s=t+1}^{t+\tau-1} \Delta F(\theta_{s+1} | \theta_s) \\
& \cdot \int_0^\infty \int_{\mathbb{R}_+^\tau} \Delta u(\theta_{t+\tau}, q_{t+\tau}) \ \mathrm{d}\xi_{\theta^{t+\tau}}(q_{t+\tau}, \ldots, q_{t+1} | q^t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_j)}(q_t | q^{t-1}) \\
= \quad & \sum_{j=i+1}^{N} \int_0^\infty \Delta u(\theta_t = \theta_j, q_t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_j)}(q_t | q^{t-1}) \\
& + \sum_{j=i+1}^{N} \sum_{\tau=1}^{T-t} \delta^\tau \sum_{\theta^{t+\tau} \in \Theta^{t+\tau}(\theta^{t-1}, \theta_j)} \Delta F(\theta_{t+1} | \theta_t) \prod_{s=t+1}^{t+\tau-1} \Delta F(\theta_{s+1} | \theta_s) \\
& \cdot \int_0^\infty \int_{\mathbb{R}_+^\tau} \Delta u(\theta_{t+\tau}, q_{t+\tau}) \ \mathrm{d}\xi_{\theta^{t+\tau}}(q_{t+\tau}, \ldots, q_{t+1} | q^t) \mathrm{d}\xi_{(\theta^{t-1}, \theta_t = \theta_j)}(q_t | q^{t-1}) \\
= \quad & \sum_{j=i+1}^{N} \sum_{\tau=0}^{T-t} \delta^\tau \sum_{\theta^{t+\tau} \in \Theta^{t+\tau}(\theta^{t-1}, \theta_j)} \prod_{s=t}^{t+\tau-1} \Delta F(\theta_{s+1} | \theta_s) \\
& \cdot \int_{\mathbb{R}_+^{\tau+1}} \Delta u(\theta_{t+\tau}, q_{t+\tau}) \ \mathrm{d}\xi_{\theta^{t+\tau}}(q_{t+\tau}, \ldots, q_t | q^{t-1}),
\end{aligned}
$$

for all $i \in I$.

$\square$

Lemma 3. *Under the first-order approach, the explicit representation of the continuation surplus $S(\theta_t | h^{t-1})$ is given through*

$$
\begin{aligned}
S(\theta_t | h^{t-1}) = & \sum_{\tau=0}^{T-t} \delta^\tau \sum_{\theta^{t+\tau} \in \Theta^{t+\tau}(\theta^t)} \prod_{s=t}^{t+\tau-1} f(\theta_{s+1} | \theta_s) \\
& \cdot \int_{\mathbb{R}_+^{\tau+1}} s(\theta_{t+\tau}, q_{t+\tau}) \ \mathrm{d}\xi_{\theta^{t+\tau}}(q_{t+\tau}, \ldots, q_t | q^{t-1}),
\end{aligned}
$$

for all $i \in I$, all $t \in \mathcal{T}$ and all histories $h^{t-1} \in H^{t-1}$.

Proof of Lemma 3. Using again backward induction, the basis for $t = T$ follows directly from equation (5). The Lemma is therefore shown with

$$\begin{aligned}
S(\theta_t|h^{t-1}) &= \int_0^\infty s(\theta_t, q_t)\mathrm{d}\xi_{\theta^t}(q_t|q_{t-1}) \\
&\quad +\delta \sum_{\theta_{t+1}\in\Theta} f(\theta_{t+1}|\theta_t) \sum_{\tau=0}^{T-(t+1)} \delta^\tau \sum_{\theta^{t+\tau+1}\in\Theta^{t+\tau+1}(\theta^{t+1})} \prod_{s=t+1}^{t+\tau} f(\theta_{s+1}|\theta_s) \\
&\quad \cdot \int_0^\infty \int_{\mathbb{R}_+^{\tau+1}} s(\theta_{t+\tau+1}, q_{t+\tau+1})\ \mathrm{d}\xi_{\theta^{t+\tau+1}}(q_{t+\tau+1},\dots,q_{t+1}|q^t)\mathrm{d}\xi_{\theta^t}(q_t|q_{t-1}) \\
&= \int_0^\infty s(\theta_t, q_t)\mathrm{d}\xi_{\theta^t}(q_t|q_{t-1}) \\
&\quad + \sum_{\tau=0}^{T-t-1} \delta^{\tau+1} \sum_{\theta^{t+\tau+1}\in\Theta^{t+\tau+1}(\theta^t)} \prod_{s=t}^{t+\tau} f(\theta_{s+1}|\theta_s) \\
&\quad \cdot \int_{\mathbb{R}_+^{\tau+2}} s(\theta_{t+\tau+1}, q_{t+\tau+1})\ \mathrm{d}\xi_{\theta^{t+\tau+1}}(q_{t+\tau+1},\dots,q_t|q^{t-1}) \\
&= \int_0^\infty s(\theta_t, q_t)\mathrm{d}\xi_{\theta^t}(q_t|q_{t-1}) \\
&\quad + \sum_{\tau=1}^{T-t} \delta^\tau \sum_{\theta^{t+\tau}\in\Theta^{t+\tau}(\theta^t)} \prod_{s=t}^{t+\tau-1} f(\theta_{s+1}|\theta_s) \\
&\quad \cdot \int_{\mathbb{R}_+^{\tau+1}} s(\theta_{t+\tau}, q_{t+\tau})\ \mathrm{d}\xi_{\theta^{t+\tau}}(q_{t+\tau},\dots,q_t|q^{t-1}).
\end{aligned}$$

□

Proof of Lemma 1. Now, it is easy to deduce Lemma 1 from Lemmata 2 and 3 by inserting $U(\theta_t = \theta_i|h^{t-1})$ and $S(\theta_t|h^{t-1})$ for $t = 1$ into principal's maximization problem:

$$\begin{aligned}
&\sum_{i=0}^N \mu_i\big(S(\theta_1 = \theta_i) - U(\theta_1 = \theta_i)\big) \\
&= \sum_{i=0}^N \mu_i \left(\sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta^{\tau+1}\in\Theta^{\tau+1}(\theta_i)} \prod_{s=1}^{\tau} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+1}} s(\theta_{\tau+1}, q_{\tau+1})\ \mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}) \right. \\
&\quad \left. - \sum_{j=i+1}^N \sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta^{\tau+1}\in\Theta^{\tau+1}(\theta_j)} \prod_{s=1}^{\tau} \Delta F(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+1}} \Delta u(\theta_{\tau+1}, q_{\tau+1})\ \mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}) \right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\tau=0}^{T-1} \delta^\tau \sum_{i=0}^{N} \mu_i \Bigg(\sum_{\theta^{\tau+1} \in \Theta^{\tau+1}(\theta_i)} \prod_{s=1}^{\tau} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+1}} s(\theta_{\tau+1}, q_{\tau+1})\ \mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}) \\
&\quad - \frac{1-F(\theta_i)}{\mu_i} \sum_{\theta^{\tau+1} \in \Theta^{\tau+1}(\theta_i)} \prod_{s=1}^{\tau} \Delta F(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+1}} \Delta u(\theta_{\tau+1}, q_{\tau+1})\ \mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}) \Bigg) \\
&= \sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta_1 \in \Theta} f(\theta_1) \sum_{\theta^{\tau+1} \in \Theta^{\tau+1}(\theta_1)} \prod_{s=1}^{\tau} f(\theta_{s+1}|\theta_s) \cdot \\
&\quad \int_{\mathbb{R}_+^{\tau+1}} \left(s(\theta_{\tau+1}, q_{\tau+1}) - \frac{1-F(\theta_1)}{f(\theta_1)} \prod_{s=1}^{\tau} \frac{\Delta F(\theta_{s+1}|\theta_s)}{f(\theta_{s+1}|\theta_s)} \Delta u(\theta_{\tau+1}, q_{\tau+1}) \right) \mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}) \\
&= \sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta^{\tau+1} \in \Theta^{\tau+1}} \prod_{s=0}^{\tau} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+1}} v(\theta_{\tau+1}, q_{\tau+1}) \mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}) \\
&= \sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta^{\tau+2} \in \Theta^{\tau+2}} \prod_{s=0}^{\tau+1} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+2}} v(\theta_{\tau+1}, q_{\tau+1}) \mathrm{d}\xi_{\theta^{\tau+2}}(q^{\tau+2}) \\
&= \ldots \\
&= \sum_{\theta \in \Theta^T} \prod_{s=0}^{T-1} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^T} V(\theta, q) \mathrm{d}\xi_\theta(q).
\end{aligned}
$$

□

References

Battaglini, M. (2005). Long-term contracting with markovian consumers. *American Economic Review*, 95(3):637–658.

Battaglini, M. and Lamba, R. (2017). Optimal dynamic contracting: the first-order approach and beyond. Cornell University and Pennsylvania State University.

Courty, P. and Li, H. (2000). Sequential screening. *Review of Economic Studies*, 67(4):697–717.

Deb, R. and Said, M. (2015). Dynamic screening with limited commitment. *Journal of Economic Theory*, 159(Part B):891–928.

Eső, P. and Szentes, B. (2013). Dynamic contracting with adverse selection: an irrelevance result. *Unpublished manuscript LSE, Oxford.*

Gershkov, A. and Perry, M. (2012). Dynamic contracts with moral hazard and adverse selection. *Review of Economic Studies*, 79(1):268–306.

Kapicka, M. (2010). Efficient allocations in dynamic private information economies with persistent shocks: a first order approach. *Unpublished paper, University of California, Santa Barbara.[265].*

Krähmer, D. and Strausz, R. (2015). Ex post information rents in sequential screening. *Games and Economic Behavior*, 90:257–273.

Li, H. and Shi, X. (2013). Discriminatory information disclosure. *American Economic Review*, 107(11):3363–3385.

Pavan, A., Segal, I., and Toikka, J. (2009). Dynamic mechanism design: incentive compatibility, profit maximization and information disclosure. Technical report, Discussion Paper, Center for Mathematical Studies in Economics and Management Science.

Pavan, A., Segal, I., and Toikka, J. (2014). Dynamic mechanism design: a myersonian approach. *Econometrica*, 82(2):601–653.

Strausz, R. (2006). Deterministic versus stochastic mechanisms in principal-agent models. *Journal of Economic Theory*, 128(1):306–314.