BDPEMS
Berlin Doctoral Program in Economics and Management Science

Working Paper Series
2015-13

Deterministic versus stochastic contracts in a dynamic principal-agent model

Thomas Schacherer, Humboldt-Universität zu Berlin

Sponsored by

EINSTEIN Foundation.de
Einstein Stiftung Berlin
Einstein Foundation Berlin

BDPEMS
Spandauer Straße 1
10099 Berlin
Tel.: +49 (0)30 2093 5780
mail: bdpems@hu-berlin.de
http://www.bdpems.de

Deterministic versus stochastic contracts in a dynamic principal-agent model

Thomas Schacherer

August 12, 2015

Abstract

This paper studies stochastic dynamic contracting between a principal and an agent, whose type evolution follows a Markov process. I analyze contracts in which the agent can terminate the contract in every period whereas the principal has full-commitment to her offer. The principal tries to screen the true type of the agent to maximize her profit. Therefore, she wants to incentivize him to reveal his true type. I show that stochastic contracts can never bring about more profits than deterministic contracts for the principal if the first-order approach is valid. For this result, it is immaterial if stochastic contracts depend on earlier realizations of the contract or not.

1 Introduction

Many contractual relationships take place in a dynamic setting, involving a long-term interaction between a principal and an agent in which the agent has private information captured via his type.

In this paper, I study stochastic dynamic contracting with asymmetric information between a principal and an agent. The agent has private information about his type, which changes over time via a Markov process. I assume that only the agent knows his current type but neither he nor the principal knows future type realizations. I analyze contracts in which the agent can terminate the contract in every period, whereas the principal has full commitment to her offer. The principal tries to screen the true type of the agent to maximize her profit.

She wants to incentivize him to reveal his true type and to make sure that he continues the contract. Therefore, in every period, she has to take into account all constraints of incentive compatibility (IC) and individually rationality (IR). My main result is that deterministic contracts are always better if the IR-constraints only bind for the lowest types and if only downward binding IC-constraints have to be taken into account, which is the so-called first-order approach in the sense of Rogerson (1985).

There are a lot of imaginable situations, where a principal interacts with an agent over long period. For instance, one can imagine relationships between a financier and a firm, between an employer and an employee or between an insurance company and a holder. In all these examples, the agent's status (type) can change over time but it remains rather persistent. For instance, one could think of employer's productivity as the agent's type, which is persistent over a long time but can change exogenously, like after an accident. Stochastic contracts in which both parties play a lottery are rarely concluded. But this still raises the question, if they could yield a higher profit to the principal and if so, under which circumstances.

This is why most of the literature focuses only on deterministic contracts. The first paper which analyzes the situation of dynamic interaction between a principal and an agent is Baron and Besanko (1984). In their paper, they extend the static "envelope theorem" to a dynamic framework, for two special dynamic situations: First, when types are constant over time. In this situation, the contract is just a repetition of the static contract in every period. Second, when the type realizations are independently drawn. This yields that the contract achieves the first-best allocation in every period after the first period.

If one allows, however, for correlated and not perfectly persistent types, it is a priori not clear whether deterministic contracts could do better than stochastic ones. For instance, stochastic contracts could open another profitable channel to screen the agent's type. In a standard one period model, it suffices to restrict to deterministic contracts under very mild assumptions, which is shown in Strausz (2006). He shows that this is true as long as the solution of the first-order approach is monotonic, i.e. higher types obtain higher quantities. In a dynamic setting, however, monotonicity is a very strong requirement which fails in a lot of conceivable situations. Battaglini (2005) analyzes long-term contracts where the agent's type realization has only two outcomes following a Markov process, in which it is more probable that the two types will remain the same. He shows the "generalized no distortion at the top principle": Once the agent reports high type, he will be assigned with the first-best allocation in any further period, regardless of future type realizations. The reason for this is that today's report gives a signal to future realizations and the principal is able to extract expected rents initially. Battaglini and Lamba (2014) show that this happens for larger type spaces as well.

My setting of the model is mostly based on the model of Battaglini and Lamba (2014). Like in their paper, I assume that neither the agent nor the principal know future types, but one can extract information about the distribution of future types due to correlation of types between current and next period. To fulfill certain type persistence, I assume first-order stochastic dominance of the conditional distribution functions, which is a common assumption in the literature. This guarantees the "generalized no distortion at the top principle". After signing the offered contract, the agent decides to continue or to terminate the relationship. Once he terminates the contract, he has no possibility to renew the contract. Furthermore, I do not allow for renegotiation, which means that the principal has full commitment to her initially offered contract.

In this paper, I analyze whether stochastic contracts could yield a higher expected payoff to the principal. I show, however, that such contracts can never bring about more profit if the first-order approach is valid, that is if local-incentive constraints are sufficient for implementation. This result is a dynamic extension of Strausz (2006) and it is in accord with a remark in Pavan et al. (2014), where they claim it without a formal proof.

The paper is organized as follows: In Section 2, I introduce the setting of the model based on the notations of Battaglini and Lamba (2014) with stochastic implementation functions. In Section 3, I show explicitly, why stochastic contracts do not yield higher payoffs for the principal. Finally, in Section 4, I present my conclusions.

2 Model

There are two players, a principal and an agent. The principal offers a contract over a finite time horizon $T \geqslant 2$. The set of the time horizon is denoted by $\mathcal{T} := \{1, \ldots, T\}$ [1]. If the contract is concluded, the principal sells in every period $t \in \mathcal{T}$ a potentially stochastic quantity of a good, which depend on current and previous type reports as well as on realizations of previous quantity schedules. In the first period, the agent has the opportunity to accept or reject the contract. In every later period $t \in \mathcal{T}\backslash\{1\}$, he decides to continue or to terminate the relationship. Once the agent terminates the contract, he has no possibility to rejoin the contract. Furthermore, I do not allow for renegotiation, hence the principal has full commitment to the initially offered contract.

2.1 Basic Assumptions

Let $\Theta := \{\theta_N, \ldots, \theta_0\} \subset \mathbb{R}$ be the agent's type space with $\theta_{i-1} - \theta_i > 0$ for all $i \in I\backslash\{0\}$, where $I := \{0, \ldots, N\}$ is the set of all indices of types. 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. $F(\theta_k|\theta_i) \geqslant F(\theta_k|\theta_{i-1})$, for all $k \in I$ and all $i \in I\backslash\{0\}$. I define for all $k \in I$ and all $i \in I\backslash\{0\}$ the nonnegative expression

$$\Delta F(\theta_k|\theta_i) := F(\theta_k|\theta_i) - F(\theta_k|\theta_{i-1}).$$

In the situation of two types, the nonnegativity guarantees that it is more likely

[1] 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.

to remain the same type. Therefore, the assumption of first order stochastic dominance can be interpreted as a generalization, which captures a certain persistence feature of types.

In the following, I use the notation θ_t to characterize the agent's type in period $t \in \mathcal{T}$ [2]. Moreover, let $\theta^t \in \Theta^t$ be the evolution vector $\theta^t := (\theta_1, \ldots, \theta_t)$ 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 $q^t := (q_1, \ldots, q_t) \in \mathbb{R}^t_+$ be the vector of quantity realizations and $p^t := (p_1, \ldots, p_t) \in \mathbb{R}^t$ the price-vector, each from period 1 up to period $t \in \mathcal{T}$, where $p_t = p(q_t)$ and $q := q^T$, $p := p^T$ the corresponding vectors over the whole time horizon $\mathcal{T}$. It is necessary to take into account that both 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 $q^0, \theta^0 \in \emptyset$.

2.2 Stochastic contracts

Since I allow for 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

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

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

[2] 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.

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. Furthermore, it allows for interdependences between the random variables over several periods. Only in the initial period $t = 1$, the implementation function $\xi(q_1|h^0, \theta_1) = \xi(q_1|\theta_1)$ does not depend on other realized quantities. 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 $\mathrm{d}\xi_{\theta^t}(q_t|q^{t-1})\mathrm{d}\xi_{\theta^{t-1}}(q_{t-1}|q^{t-2}) = \mathrm{d}\xi_{\theta^t}(q_{t-1}, q_t|q^{t-2})$ for all $t \in \mathcal{T}\backslash\{1\}$ and hence

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

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

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 $p_t \in \mathbb{R}$. This generates a per-period utility of $u(\theta_t, q_t) - p_t$ for the agent. As in the literature, I assume that u satisfies several assumptions. It 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. These conditions are summarized in

Assumption 1. The per-period utility function u satisfies

- $u(\cdot, 0) = 0$,
- $\partial u/\partial q > 0$,
- $\partial u/\partial \theta > 0$,
- $\partial^2 u/\partial q \partial \theta > 0$,
- $\partial u/\partial q^2 \leqslant 0$.

The agent discounts future utilities by $\delta \in]0,1[$. Therefore, one can define the agent's continuation utility recursively as

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

$$U(\theta_t|h^{t-1}) := \int_0^\infty \left(u(\theta_t, q_t) - p_t + \delta \sum_{\theta_{t+1}\in\Theta} f(\theta_{t+1}|\theta_t) U(\theta_{t+1}|h^{t-1}, \theta_t, q_t) \right) \mathrm{d}\xi_{\theta^t}(q_t|q^{t-1}), \tag{2}$$

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

It captures the expected utility of current utility and the discounted expectation of future continuation utilities that considers that the type changes with probability $f(\theta_{t+1}|\theta_t)$.

2.4 Principal's offer

In every period $t \in \mathcal{T}$, the principal produces q_t given a cost function $c(q_t)$. This function fulfills some 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. These conditions are stated in

Assumption 2. The cost function c satisfies

- $c(0) = 0$,
- $c'(q) > 0$, for all $q > 0$,
- $c'(0) = 0$,
- $\lim_{q\to\infty} c'(q) = \infty$,
- $c''(q) > 0$.

The principal offers a stochastic contract $\{p, \xi_\theta\}$ in the initial period, which includes a sequence of prices and implementation functions that yield the realizations of the quantities q. Equivalently to $\{p, \xi_\theta\}$, the principal can offer $\{U, \xi_\theta\}$, where U represents the vector $U = (U(\theta_1|h^0), \ldots, U(\theta_T|h^{T-1}))$ of agent's continuation utility.

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 $\{U, \xi_\theta\}$ which incorporates in every period t all possible type reports θ_t of the agent and all possible histories $h^{t-1} \in H^{t-1}$. 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 $\text{IR}(\theta_t|h^{t-1})$ can be described as

$$U(\theta_t|h^{t-1}) \geqslant 0, \tag{3}$$

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 $\text{IC}(\theta_t, \vartheta_t|h^{t-1})$ can be characterized by

$$\begin{aligned} U(\theta_t|h^{t-1}) \geqslant U(\vartheta_t|h^{t-1}) + \int_0^\infty \big(u(\theta_t, q_t) - u(\vartheta_t, q_t)\big)\mathrm{d}\xi_{(\theta^{t-1},\vartheta_t)}(q_t|q^{t-1}) \\ + \ \delta \sum_{\theta_{t+1}\in\Theta} \big(f(\theta_{t+1}|\theta_t) - f(\theta_{t+1}|\vartheta_t)\big) \int_0^\infty U(\theta_{t+1}|h^{t-1}, \vartheta_t, q_t)\mathrm{d}\xi_{(\theta^{t-1},\vartheta_t)}(q_t|q^{t-1}), \end{aligned} \tag{4}$$

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.

$$\max_{\{U,\xi_\theta\}} \left\{ \sum_{\theta_1\in\Theta} f(\theta_1)\big(S(\theta_1) - U(\theta_1)\big) \right\}, \tag{5}$$

s.t. (3) and (4) are satisfied, whereby

$$S(\theta_t|h^{t-1}) := \int_0^\infty \left(s(\theta_t, q_t) + \delta \sum_{\theta_{t+1}\in\Theta} f(\theta_{t+1}|\theta_t) S(\theta_{t+1}|h^{t-1}, \theta_t, q_t) \right) \mathrm{d}\xi_{\theta^t}(q_t|q^{t-1}) \tag{6}$$

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 $h^T \in H^T$.

The purpose of this paper is to show that optimal deterministic contracts which fulfill all individually rationality (IR) and incentive compatibility (IC) constraints of the agent are still optimal under the first-order approach if one allows for stochastic contracts.

3 Optimal contracting under the first-order approach

One first consideration is a situation in which in every period $t \in \mathcal{T}$ the IR-constraint only binds for the lowest type $\theta_t = \theta_N$ and the IC-constraints only bind downward, i.e. type $\theta_t = \theta_i$ is indifferent between reporting θ_i and $\vartheta_t = \theta_{i+1}$. If it is sufficient to involve only these constraints into the contract, then it is first-order optimal and the first-order approach holds. In this framework, first-order optimality is characterized similarly to Battaglini and Lamba (2014) by

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

In a dynamic setting, it is a priori not clear why deterministic contracts should be superior to stochastic ones. However, it turns out that deterministic contracts are superior to any stochastic contract as long as the first-order approach holds. Therefore, the results of Battaglini and Lamba (2014) are still optimal in the set of stochastic contracts. In order to show that this is in fact the case, I first derive an explicit representation of principal's maximization problem (5) with respect to $\{\mathrm{IR}(\theta_t = \theta_N|h^{t-1})\}_{t\in\mathcal{T}}$ and $\{\mathrm{IC}(\theta_t = \theta_i, \vartheta_t = \theta_{i+1}|h^{t-1})\}_{t\in\mathcal{T}}$, for all $i \in I\backslash\{N\}$.

For all $i \in I\backslash\{0\}$, I use the notation $\Delta u(\theta_t = \theta_i, q(\theta_t = \theta_i|q^{t-1}, \theta^{t-1})) := u(\theta_t = \theta_{i-1}, q(\theta_t = \theta_i|q^{t-1}, \theta^{t-1})) - u(\theta_t = \theta_i, q(\theta_t = \theta_i|q^{t-1}, \theta^{t-1}))$, to characterize the net-utility of type θ_{i-1} compared to type θ_i if both report θ_i in period $t \in \mathcal{T}$. If the current type is not specified, this net-utility is denoted by $\Delta u(\theta_t, q_t)$. The virtual surplus in period $\tau \in \mathcal{T}$ is defined by

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

This allows to state the following result:

Lemma 1. *Involving all IR- and IC-constraints of the first-order approach, principal's objective* (5) *has the explicit representation*

$$\sum_{\theta_1 \in \Theta} f(\theta_1)\big(S(\theta_1) - U(\theta_1)\big) =$$

$$\sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta^{\tau+1} \in \Theta^{\tau+1}} f(\theta_1) \prod_{s=1}^{\tau} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+1}} v(\theta_{\tau+1}, q_{\tau+1}) \mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}). \tag{7}$$

To prove this Lemma, I show first two other 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*

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

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

A similar result for the continuation surplus is given through

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

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

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

Now, with Lemmata 2 and 3 it is straightforward to simplify principal's objective (7) 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.
Then, it is possible to further simplify Lemma 1. Using $\sum_{\theta^s\in\Theta^s(\theta^{s-1})} f(\theta_s|\theta_{s-1}) = 1$, and $\int_0^\infty \mathrm{d}\xi_{\theta^s}(q_s|q^{s-1}) = 1$ for all $s \in \mathcal{T}$, one obtains iteratively

$$\begin{aligned}
& \sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta^{\tau+1}\in\Theta^{\tau+1}} \prod_{s=0}^{\tau} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+1}} v(\theta_{\tau+1}, q_{\tau+1})\mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}) \\
= & \sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta^{\tau+1}\in\Theta^{\tau+1}} \prod_{s=0}^{\tau} f(\theta_{s+1}|\theta_s) \left(\sum_{\theta_{\tau+2}\in\Theta^{\tau+2}(\theta^{\tau+1})} f(\theta_{\tau+2}|\theta_{\tau+1}) \right) \int_{\mathbb{R}_+^{\tau+1}} v(\theta_{\tau+1}, q_{\tau+1})\mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}) \\
= & \sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta^{\tau+2}\in\Theta^{\tau+2}} \prod_{s=0}^{\tau+1} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+1}} v(\theta_{\tau+1}, q_{\tau+1}) \left(\int_{\mathbb{R}_+} \mathrm{d}\xi_{\theta^{\tau+2}}(q_{\tau+2}|q^{\tau+1}) \right) \mathrm{d}\xi_{\theta^{\tau+1}}(q^{\tau+1}) \\
= & \sum_{\tau=0}^{T-1} \delta^\tau \sum_{\theta^{\tau+2}\in\Theta^{\tau+2}} \prod_{s=0}^{\tau+1} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^{\tau+2}} v(\theta_{\tau+1}, q_{\tau+1})\mathrm{d}\xi_{\theta^{\tau+2}}(q^{\tau+2}) \\
= & \dots \\
= & \sum_{\theta\in\Theta^T} \prod_{s=0}^{T-1} f(\theta_{s+1}|\theta_s) \int_{\mathbb{R}_+^T} \sum_{\tau=0}^{T-1} \delta^\tau v(\theta_{\tau+1}, q_{\tau+1})\mathrm{d}\xi_\theta(q).
\end{aligned}$$

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

Similar to the static situation in Strausz (2006), 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.

4 Conclusion

This paper shows that stochastic contracts do not yield higher profits to the principal in dynamic contracting, if one assumes full-commitment of the principal and that the first-order approach is valid. Therefore, in this class of contracts, it is sufficient to restrict to deterministic contracts.

The first step in solving this is to transform the principal's objective (5) in terms of the virtual surplus function v, shown in Lemma 1. The crucial step in the proof is the possibility of creation of the function V, which captures the discounted sum of virtual surpluses over the whole time horizon. Once one has obtained this objective function, one can use the same argument as in Strausz (2006). The difficulty is rather in achieving the function V. The fact that the first-order approach is valid is essential in this case. Since the IC-constraints always only bind in one direction, namely to the lower neighbouring type, and the IR-constraints only bind for the lowest types, one obtains iteratively that dynamic virtual surplus function.

Consequently, it is not possible to obtain such a function V in general, when the first-order approach fails, since the set of binding constraints could depend on the quantity realizations of previous periods and therefore, it is impossible to extend the proof to such situations.

5 Appendix

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

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

for all $i \in I$. Here, I use

$$\mathrm{d}\xi_{\theta^{t+\tau}}(q_{t+\tau},\dots,q_t|q^{t-1}) = \mathrm{d}\xi_{\theta^{t+\tau}}(q_{t+\tau},\dots,q_{t+1}|q^t)\mathrm{d}\xi_{\theta^{t+1}}(q_{t+1},q_t|q^{t-1})$$

for the probability measures of the conditional implementation functions.

$\square$

Proof of Lemma 3. Again, I will show this statement with backward induction. The basis for $t = T$ follows directly from equation (6). The Lemma is therefore shown with

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

$\square$

Proof of Lemma 1. Now it is easy to deduce Lemma 1 from Lemmata 2 and 3:

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

□

References

Baron, D. P. and Besanko, D. (1984). Regulation and Information in a Continuing Relationship. *Information Economics and Policy*, 1(3):267–302.

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

Battaglini, M. and Lamba, R. (2014). Optimal Dynamic Contracting: the First-Order Approach and Beyond. *Working Paper*.

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

Rogerson, W. P. (1985). The first-order approach to principal-agent problems. *Econometrica: Journal of the Econometric Society*, pages 1357–1367.

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