UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

# RISCO MORAL DINÂMICO COM APRENDIZADO SOBRE A FUNÇÃO DE PRODUÇÃO

# DYNAMIC MORAL HAZARD WITH LEARNING ABOUT THE PRODUCTION FUNCTION

**Maurício Massao Soares Matsumoto**

**Orientador: Prof. Dr. Gabriel de Abreu Madeira**

**Versão Corrigida**

**SÃO PAULO**
**2014**

Prof. Dr. Marco Antonio Zago
Reitor da Universidade de São Paulo

Prof. Dr. Reinaldo Guerreiro
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Joaquim José Martins Guilhoto
Chefe do Departamento de Economia

Prof. Dr. Marcio Issao Nakane
Coordenador do Programa de Pós-Graduação em Economia

MAURÍCIO MASSAO SOARES MATSUMOTO

# RISCO MORAL DINÂMICO COM APRENDIZADO SOBRE A FUNÇÃO DE PRODUÇÃO

# DYNAMIC MORAL HAZARD WITH LEARNING ABOUT THE PRODUCTION FUNCTION

Dissertação apresentada ao Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito para a obtenção do título de Mestre em Ciências.

**Orientador: Prof. Dr. Gabriel de Abreu Madeira**

**SÃO PAULO**

**2014**

**FICHA CATALOGRÁFICA**
Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Matsumoto, Maurício Massao Soares
Dynamic moral hazard with learning about the production function / Maurício Massao Soares Matsumoto. -- São Paulo, 2014.
37 p.

Dissertação (Mestrado) – Universidade de São Paulo, 2014.
Orientador: Gabriel de Abreu Madeira.

1. Contratos 2. Incentivos 3. Aprendizagem 4. Programação linear I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

CDD – 346.02

Aos meus pais, Haroldo e Marisa.

# AGRADECIMENTOS

Aos meus pais, que sempre fizeram tudo que podiam por mim e pela minha irmã. Obrigado pelo apoio sempre incondicional, e por me ensinarem o valor dos estudos. À minha irmã Mônica, que sempre admirei e que me incentivou a voltar para a vida acadêmica.

À minha turma do mestrado, em que tive o prazer de descobrir colegas inteligentes e dedicados, com quem aprendi muito, e que também se tornaram grandes amigos. O percurso foi muito mais divertido e gratificante na companhia de todos vocês. Em especial aos que se tornaram mais próximos pelos tantos momentos, ideias e cervejas compartilhados: Bel, Bruna, Dimas, Fabio, Gabriel, João, Julia, Leo, Leopoldo, Luísa, Natalia, Pedro, Perez. As boas lembranças são muitas: das discussões de listas de exercício na salinha às idas à praia, das monitorias compartilhadas aos blocos de carnaval, das corridas na raia aos seminários que organizávamos entre nós, a viagem para Boston, os churrascos e jantares na casa da Bel. Entre tantas outras.

Também às outras turmas de mestrado e doutorado com quem convivi em meu tempo na FEA, veteranos e bixos. Em especial: Carnaúba, Laura, Leo Rosa, Rafael Proença e Renata.

Aos professores do IPE-USP, com quem aprendi muito dentro e fora da sala de aula, às vezes até no bar. Em especial, agradeço ao meu orientador Gabriel Madeira, pelas inúmeras conversas (acadêmicas ou não) que me ensinaram muito, por acreditar em mim e me dar tantas oportunidades. Agradeço também a Bruno Giovanetti, Marcos Rangel, Marcos Nakaguma, Mauro Rodrigues, Rafael Coutinho e Ricardo Madeira, não apenas pelas interações acadêmicas, mas também por terem me ajudado nas decisões de vida e carreira.

Por fim, agradeço à FIPE, CNPq e FAPESP pelo fundamental apoio financeiro.

# RESUMO

Neste trabalho, propomos uma estratégia numérica para lidar com modelos de risco moral dinâmico com aprendizado sobre a função de produção. Pela complexidade do problema, soluções analíticas na literatura têm sido limitadas em seu escopo. A contribuição é metodológica: através de métodos computacionais, o problema pode ser estudado sob poucas hipóteses a respeito de formas funcionais. Partindo de um mecanismo geral, reformulamos o problema como um mecanismo compatível em incentivos, e então mostramos como este pode ser resolvido por indução retroativa por meio de uma sequência de programas lineares. Aplicamos o método a alguns casos de interesse, e confirmamos a conclusão da literatura de que a incerteza sobre a função de produção aumenta a volatilidade da utilidade do agente para prevenir manipulação de crenças.

## ABSTRACT

In this work we propose a flexible numerical approach to deal with models of dynamic moral hazard with simultaneous learning about the production function. Because of the complexity of the problem, analytical solutions have so far been limited in scope. The contribution is methodological: through computation, the problem can be studied under few assumptions about functional forms. We depart from a general mechanism, reformulate it as an incentive compatible mechanism, and show how it can be solved by backward induction through a sequence of linear programs. We apply our method to a few cases of interest, and confirm that uncertainty about the production function increases the volatility of the agent's utility in order to prevent belief manipulation, as found in the literature.

**SUMÁRIO**

# LISTA DE FIGURAS

# 1 Introduction

In the standard moral hazard setting, it is assumed that principal and agent know precisely how effort affects the distribution of output. However, in many contractual arrangements in practice, the production function is not fully known *a priori*, and so the issue of providing incentives is entangled with a learning process.

A few examples are: a firm hiring a worker whose ability to perform the contracted task is unknown to both parties; investors contracting a manager for a start-up operating in a market with unknown profitability; a government contracting a firm to provide a public service with unknown demand or operating costs.

We propose to study situations in which the production function is stochastic and unknown. In standard moral hazard models , output is uncertain given the level of effort exerted by the agent, but its probability distribution is publicly known. The new ingredient here is allowing the distribution itself to be unknown, and so principal and agent need to learn about it as the relationship unfolds. An interesting consequence of this environment is that the agent can attempt to manipulate the principal's learning process, in order to achieve private informational rents in subsequent periods.

This learning process is only possible in a dynamic setting, where past histories can be observed. If the contract is static, the problem is uninteresting because the uncertainty about the production function is bundled with the uncertainty of output.

Problems of learning and incentive provisioning are treated separately by most of the economic literature. In the dynamic moral hazard literature, it is usually assumed that the production function is publicly known, e.g. (ROGERSON, 1985; PHELAN; TOWNSEND, 1991). In this case, the problem is greatly simplified by using a recursive formulation, since past history can be summarized by a single state variable: the agent's continuation utility. This has been originally demonstrated in (SPEAR; SRIVASTAVA, 1987). When learning is introduced, however, this recursive formulation is no longer possible and the problem becomes fully history dependent.

There have been a few attempts to model learning and moral hazard simultaneously, under no long-term commitment such as in (LAFFONT; TIROLE, 1988; HOLMSTROM, 1999). (HOLMSTROM, 1999) shows that if wages are set competitively by the market in every period, the agent is willing to make effort in order to build his reputation (output is a signal about ability) and increase his future expected wage, even if no contingent contracts are written (that is, no explicit incentives are given). He analyzes these "implicit contracts" in a few tractable cases to argue that inefficiencies due to moral hazard may be alleviated, but not offset, by reputation concerns. This paper inaugurated the "career concerns" literature.

(LAFFONT; TIROLE, 1988) studies a hidden action environment with adverse selection where the principal learns about the agent's type as the contract unfolds. The agent avoids making effort because it raises the principal's expectations on future output, what has become know as the "ratchet effect".

In a more recent literature such as (DEMARZO; SANNIKOV, 2011; PRAT; JOVANOVIC, 2012; HE; WEI; YU, 2013), dynamic moral hazard with learning has been studied under full commitment, using a continuous-time principal-agent framework and making simplifying hypotheses to maintain tractability. In these papers, the production function is simplified as the sum of an unknown parameter (ability or productivity), a privately observed effort and an error term. This structure makes the problem tractable because productivity and effort become substitutes for production, but interesting cases such as learning whether effort matters or not for production[1] are excluded.

(PRAT; JOVANOVIC, 2012) finds that under a high level of uncertainty about the production function, providing incentives is easier through spot contracts (one period contracts) like in (HOLMSTROM, 1999), and under low levels of uncertainty full commitment contracts can improve ex-ante welfare. They show that under uncertainty about the production function, the agent's utility has to be more volatile to prevent "belief manipulation" when implementing the first best level of effort. While (PRAT; JOVANOVIC, 2012) studies the optimal contract that implements high effort, (HE; WEI; YU, 2013) explores a similar model further to characterize the optimal effort policy.

In a corporate finance setting, (DEMARZO; SANNIKOV, 2011) analyzes the problem of aligning incentives of the CEO of a firm with those of its shareholders, when productivity follows random fluctuations. They find that it is possible to implement the optimal contract by giving the agent a non-tradable share of equity, which offsets the gains

[1] We study this case in subsection 4.2.

from manipulating the principal's beliefs because he becomes a shareholder himself.

Related to this paper, (MANSO, 2011) also studies a dynamic moral hazard environment with learning. His paper focuses on the tension between exploitation and innovation (that is: whether to learn or not), while here we study the optimal compensation schedule under learning.

The literature so far is focused on both extremes of the spectrum of commitment. On one end there are discrete models with little or no commitment such as the "career concerns" literature. On the other end, there are full-commitment models in continuous time. An interesting research agenda would be to study in a uniform and accessible language the whole spectrum of commitment, and to investigate what happens to the dynamic moral hazard problem with learning under limited commitment.

Our contribution is to propose an alternative, computational approach to study the problem under full commitment. Different from the recent literature above mentioned, we study the problem using a discrete-time principal-agent model, and second we impose little structure on preferences and production functions. We build upon the linear programming framework introduced by (PRESCOTT; TOWNSEND, 1984) to propose a flexible method for dealing with dynamic moral hazard with learning about the production function.

This linear programming framework has been extensively used in the moral hazard literature, in papers such as (PAULSON; TOWNSEND; KARAIVANOV, 2006; KARAIVANOV; TOWNSEND, 2012; DOEPKE; TOWNSEND, 2006; PHELAN; TOWNSEND, 1991; MADEIRA; TOWNSEND, 2008; KILENTHONG; MADEIRA, 2009)[2]. Computing solutions to moral hazard problems has two main advantages: first, it allows the study of features of optimal contracts when an analytical solution is not available, such as in (PHELAN; TOWNSEND, 1991); second, it permits estimating parameters of the model and distinguishing across competing models using structural techniques when data on contracts is available, such as in (PAULSON; TOWNSEND; KARAIVANOV, 2006; KARAIVANOV; TOWNSEND, 2012).

This paper is organized as follows: in Section 2, we justify our methods: departing

[2] A very good introduction to the linear programming approach to moral hazard problems can be found in (PRESCOTT, 1999). For a hands-on tutorial on how to implement a static moral hazard linear program using Matlab, please refer to (KARAIVANOV, 2001)

from a general contracting game, we restate it as an incentive compatible mechanism, and show how it can be solved computationally by backward induction. Then, in Section 3, we show some validation exercises that have been performed, and in Section 4 we discuss applications of the method to cases of interest. Section 5 concludes.

# 2 Model and computational framework

The general model consists of a dynamic moral hazard environment with full commitment, to which we add a Bayesian learning process over the production function. Time is discrete and finite with T periods, and we assume that the underlying (unknown) production function is an i.i.d. process.

At the beginning of the contract, principal and agent share a prior about the production function[3]. By observing the history of output, both make inferences about the true production function using the Bayes Rule. Since the action undertaken by the agent is private information, the agent can manipulate the principal's learning process in order to extract future informational rents.

This section follows closely (DOEPKE; TOWNSEND, 2006), in order to demonstrate that the general mechanism can be reformulated as a direct mechanism, and then solved computationally using linear programming techniques.

## 2.1 General mechanism with full history dependence

Here we present the model in its most general form. Time is discrete, and the contractual relationship lasts $T$ periods[4]. In any period $t$, the timing of events is as follows (Figure 1): first, the principal sends the agent a message $m_t \in M$, which we can interpret as a recommendation for the agent's action. Then, the agent executes action $a_t \in A$, which is unobserved by the principal. Next, output $q_t \in Q$ is realized and publicly observed, and finally the principal pays the agent a compensation $c_t \in C$. The agent has no access to

[3]Although not studied here, an interesting future exercise would be to relax this assumption by introducing asymmetric priors

[4]We do not study the infinite-time version of the problem because, since a recursive formulation is not possible, we would not be able to solve it computationally.

credit or savings, and so he consumes exactly $c_t$. We impose that $A$,$M$,$Q$ and $C$ be finite, which can be interpreted as an approximation to continuous sets.

We assume full commitment in the contracting game: the principal is able to commit to follow the contracting rules established upon signing the contract, and the agent is able to commit not to quit. It is interesting to note that under full commitment the ex ante optimal contract can be ex post inefficient. This happens because in future periods, the principal is constrained to give incentives for past actions.

**Notation - past histories:**

It will be useful to have clear notations for past histories. Period $t$ history will be denoted by $h_t = \{m_t, a_t, q_t, c_t\}$, and $h^t = \{h_0, h_1, ..., h_t\}$ will be the the history up to period $t$. Information is asymmetric (hidden actions), so that the equivalent histories as observed by the principal are $s_t = \{m_t, q_t, c_t\}$ and $s^t = \{s_0, s_1, ..., s_t\}$. Since $s_t$ and $s^t$ are uniquely determined by $h_t$ and $h^t$, we will sometimes use the notation $s_t(h_t)$ or $s^t(h^t)$. For completeness, we define $s_0 = h_0 = \emptyset$.

**Strategies:**

The principal chooses outcome functions $\pi(m_t|s^{t-1})$ and $\pi(c_t|m_t, q_t, s^{t-1})$, which are probability distributions over messages and consumption plans for the agent. The agent chooses a strategy $\sigma(a_t|m_t, h^{t-1})$.

**Production function:**

For now, we model the production function in its most general form, but we will impose further restrictions in subsection 2.5, where we also describe the Bayesian learning process. Since the true underlying production function is unknown, we have a subjective probability distribution $f(q_t|a_t, h^{t-1})$ of output $q_t$ given the action undertaken $a_t$. The distribution is conditioned on the past history $h^{t-1}$, because principal and agent learn from observation.

**Probability of $h^t$:**

Given $\{\pi, \sigma\}$, we can define recursively the subjective probability of history $h^t$ happening:

$$p(h^t|\pi, \sigma) = p(h^{t-1}|\pi, \sigma).\pi(m_t|s^{t-1}\left(h^{t-1}\right)).\sigma(a_t|m_t, h^{t-1}).f(q_t|a_t, h^{t-1}).\pi(c_t|m_t, q_t, s^{t-1}\left(h^{t-1}\right))$$

And we define $p(h^0|\pi, \sigma) = 1$.

**Preferences:**

This allows us to calculate $U(\pi, \sigma)$ and $V(\pi, \sigma)$, the $t = 0$ expected utility of the agent and the principal, respectively:

$$U(\pi, \sigma) = \sum_{t=1}^{T} \beta^{t-1} \{ \sum_{H^t} p(h^t|\pi, \sigma).u(c_t, a_t) \}$$

$$V(\pi, \sigma) = \sum_{t=1}^{T} \alpha^{t-1} \left\{ \sum_{H^t} p \left( h^t|\pi, \sigma \right) . (q_t - c_t) \right\}$$

Note that the agent's utility is only required to be time-separable, and is very general otherwise. We have specialized the principal's utility function as a risk-neutral expected surplus, but the methods proposed here would equally apply to any other time-separable utility function.

In the same fashion, we define probabilities and continuation utilities conditional on realized histories $h^k$; we will denote those by $p(h^t|\pi, \sigma, h^k)$ and $U(\pi, \sigma|h^k)$.

Given an outcome function $\pi$, an *optimal strategy* $\sigma$ maximizes the agent's utility at every node $h^k$, that is:

$\forall \widehat{\sigma}, k, h^k$ :

$$U(\pi, \widehat{\sigma}|h^k) \leq U(\pi, \sigma|h^k) \tag{1}$$

Also, an outcome function $\pi$ combined with an optimal strategy $\sigma$ respect the participation constraint if they provide the agent with his reservation utility $\overline{U}$:

$$U(\pi, \sigma) \geq \overline{U} \tag{2}$$

**Problem 1.** *The principal's problem in the general mechanism is to maximize his utility $V(\pi, \sigma)$, by choice of an outcome function $\pi$, subject to the agent's optimizing behavior constraints* (1) *and the participation constraint* (2)

## 2.2 Incentive Compatible Mechanism

In this subsection, we restate the problem as an incentive compatible mechanism: that is, a mechanism in which the principal chooses the agent's actions directly, provided incentive compatibility constraints are respected. We then show, through Proposition 1, that it is possible to restrict ourselves to those mechanisms in order to achieve a solution to the general problem.

We now restrict the message space $M$ to be equal to $A$, the action space. The principal chooses outcome functions $\pi(a_t|s^{t-1})$ and $\pi(c_t|a_t, q_t, s^{t-1})$, which refer to the recommended action $a_t$, and the agent is restricted to follow the recommendation (to be "obedient"). Since $a_t$ is now a recommendation for action we will denote $h_t = \{a_t, \widehat{a}_t, q_t, c_t\}$, where $\widehat{a}_t$ is the true action undertaken by the agent.

Under the outcome function $\pi$, the $t = 1$ subjective probability $p(s^t|\pi)$ becomes:

$$p(s^t|\pi) = p(s^{t-1}|\pi).\pi(a_t|s^{t-1}).f(q_t|a_t, s^{t-1}).\pi(c_t|a_t, q_t, s^{t-1})$$

And so the agent's utility and the principal's surplus become:

$$U(\pi) = \sum_{t=1}^{T} \beta^{t-1} \left\{ \sum_{S^t} p\left(s^t|\pi\right).u\left(c_t, a_t\right) \right\}$$

$$V(\pi) = \sum_{t=1}^{T} \alpha^{t-1} \left\{ \sum_{S^t} \left(s^t|\pi\right).\left(q_t - c_t\right) \right\}$$

**Participation constraint:**

The principal is constrained to provide the agent with his reservation utility $\overline{U}$.

$$U(\pi) \geq \overline{U} \tag{3}$$

**Obedience constraint:**

In the notation of the previous subsection, the incentive compatible mechanism implies that the agent's optimal strategy $\sigma$ is a degenerate distribution with $\sigma\left(a_t|s^{t-1}, a^t\right) =$

$1$. That is: the agent executes the recommended action with probability one on the equilibrium path.

It should not be on the agent's interest to deviate from obedience. We define deviations as pure strategies $\delta_a(a_t, h^{t-1})$ that map histories and recommended actions into actions actually undertaken by the agent. Under a deviation $\delta_a$, probabilities are affected through the production function, as follows:

$$p(h^t|\pi, \delta_a) = p(h^{t-1}|\pi, \delta_a).\pi(a_t|s^{t-1}(h^{t-1})).\mathbb{I}\left(\widehat{a}_t = \delta_a(a_t, h^{t-1})\right).$$
$$.f(q_t|\delta_a(a_t, h^{t-1}), h^{t-1}).\pi(c_t|a_t, q_t, s^{t-1}(h^{t-1}))$$

The indicator function $\mathbb{I}$ is needed because $p(h^t|\pi, \delta_a) = 0$ if $\delta_a(a_k, h^{k-1}) \neq \widehat{a}_k$ for some $k \leq t$ (or, put more simply, $\mathbb{I}$ is the degenerate probability distribution over actions associated to the pure strategy $\delta_a$).

For the agent to be obedient, it must be optimal for him to do so at every node of the equilibrium path, and thus the following must hold:

$$\forall k, s^{k-1}, \delta_a : \sum_{t=k}^{T} \beta^{t-k}\{\sum_{H^t} p(h^t|\pi, \delta_a, s^{k-1}).u(c_t, \delta_a(a_t, h^{t-1}))\} \leq \sum_{t=k}^{T} \beta^{t-k}\{\sum_{S^t} p(s^t|\pi, s^{k-1}).u(c_t, a_t)\} \tag{4}$$

The inequalities (4) might seem less restrictive than their counterparts (1) in the general mechanism because we only allow deviations to be pure strategies instead of probability distributions as the strategies $\sigma$. However, if we allowed for random deviations, we would simply have linear combinations of the constraints (4), which would not alter the set of feasible outcome functions $\pi$.

**Problem 2.** *The principal's problem in the incentive compatible mechanism is to maximize his utility $V(\pi)$, by choice of an outcome function $\pi$, subject to the agent's participation constraint* (3) *and the obedience constraints* (4).

We argue that by solving the incentive compatible mechanism we will reach a solution to the general mechanism. This happens if the set of feasible allocations under both mechanisms is the same. If an allocation is feasible under the incentive compatible mechanism, it will be feasible under the general mechanism, since the former is more restrictive than the latter (under incentive compatibility the agent has fewer options, that is, he is restricted to obey). But the converse must be verified.

**Proposition 1.** *Every feasible allocation under the general mechanism is also feasible under the incentive compatible mechanism.*

*Proof.* (Outline) First, to be clear, an *allocation* is a probability distribution over actions, output and consumption that respects the constraints imposed by the subjective production function. A *feasible allocation* under the general mechanism or the incentive compatible mechanism is an allocation that results from an optimal contract in Problem 1 or 2, respectively.

Choose any feasible allocation under the general mechanism, and let $\{\pi, \sigma\}$ be the corresponding contract that implements it. To prove the proposition, we need to find a contract $\pi^*$ that implements the same allocation under the incentive compatible mechanism.

We define $\pi^*$ so that it reflects the equilibrium path from $\{\pi, \sigma\}$. That is, we impose $\pi^*\left(a_t|s^{t-1}\right) = p_{\pi,\sigma}\left(a_t|s^{t-1}\right)$ and $\pi^*\left(c_t|a_t, q_t, s^{t-1}\right) = p_{\pi,\sigma}\left(c_t|a_t, q_t, s^{t-1}\right)$, where $p_{\pi,\sigma}$ is the probability measure resulting from $\{\pi, \sigma\}$.

Using the fact that $\{\pi, \sigma\}$ respect the constraints in Problem 1, we can show that $\pi^*$ respects the constraints in Problem 2, and so the proposition holds.

□

Proposition 1 ensures us that an optimal contract resulting from Problem 2 corresponds to an optimal contract to Problem 1.

## 2.3 Continuation problem and backward induction

In this subsection, we show how the incentive compatible mechanism can be broken down period by period. We begin by formulating the principal's continuation problem at any node on the equilibrium path. Then, we show that this formulation allows the solution of the incentive compatible mechanism by backward induction.

A recursive formulation is not possible in this environment, because there is full history dependence and thus the state space grows with time without bound. But this does not mean that the problem cannot be broken down to smaller parts to be solved, which is the essence of backward induction in game theory.

The complication from usual backward induction in our case is the following: at any node of the equilibrium path, to calculate the optimal continuation contract, the principal needs to take into account all the possible deviations that the agent might have undertaken in previous periods and make sure they're not interesting from an ex-ante perspective. This creates a problem of time-consistency: the continuation contract might be inefficient, because it is constrained to provide incentives for past actions. Though, we have assumed this problem away in the previous sections, by imposing that the principal is fully committed to the initial contract. We also assume that the agent does not quit the contract, even if at some point his expected future compensation is very low because of the way the sequence of outputs unfolded. [5]

We begin by formulating the principal's continuation problem from period $k$ on ($1 < k < T$), at a node $s^{k-1}$ of the equilibrium path. The size of the vector of state variables grows quickly with time: we need to keep track of past histories and of utility promises on and off the equilibrium path. This formulation borrows from (FERNANDES; PHELAN, 2000), who proposed imposing bounds on utility off the equilibrium path on problems where privately observed histories have an influence over future preferences (like in our case).

The Problem's state variables in period $k$ are the past history $s^{k-1}$, and the vector of promised utilities $w_{k,s^{k-1}}$. The vector of promised utilities contains the utility delivered to the agent on the equilibrium path $w_{k,s^{k-1}}\left(s^{k-1}\right)$, but also the utilities delivered to the agent off the equilibrium path[6] $w_{k,s^{k-1}}\left(h^{k-1}\right)$ (that is, the agent who deviated previously

[5] It would be an interesting future exercise to model limited commitment from both parties.

[6] In a dynamic moral hazard model with no learning such as (PHELAN; TOWNSEND, 1991), the utility delivered on and off path is the same. This will become clearer in section 4 when we explore the spaces of feasible vectors of utility promises.

and is now at node $h^{k-1} \neq s^{k-1}$)[7]. It is necessary to keep track of the agent's utility off the equilibrium path to make sure that the ex-ante incentive structure is preserved when we solve the problem by backward induction.

Formally, we define the vector $w_{k,s^{k-1}}$ as a function mapping the space $H^{k-1}(s^{k-1})$ (histories that have $s^{k-1}$ as its counterpart observable by the principal[8]) into the real line. That is, $w_{k,s^{k-1}} : H^{k-1}(s^{k-1}) \to \mathbb{R}$. The space of feasible vectors of utility promises will be denoted $W_{k,s^{k-1}}$. It is important to note that under different past histories $s^{k-1}$, the space $W_{k,s^{k-1}}$ is different. This will become evident when we see the results in section 4.

**Problem 3.** *The principal's continuation problem at node $s^{k-1}$ of the equilibrium path is:*

$$V_k\left(s^{k-1}, w_{k,s^{k-1}}\right) = \max_{\pi} \sum_{t=k}^{T} \alpha^{t-k} \left[\sum_{S^t} p\left(s^t|\pi, s^{k-1}\right)(q_t - c_t)\right]$$

*subject to constraints* (5)*,* (6) *and* (7)*, defined below.*

**Promise-keeping constraint:**

The outcome function $\pi$ has to deliver the agent on the equilibrium path the promised utility [9]:

$$\sum_{t=k}^{T} \beta^{t-k} \left[\sum_{S^t} p\left(s^t|\pi, s^{k-1}\right) u\left(c_t, a_t\right)\right] = w_{k,s^{k-1}}\left(s^{k-1}\right) \tag{5}$$

**Threat-keeping constraints:**

The agent who reaches node $s^{k-1}$ having deviated previously (that is, the agent who is in node $h^{k-1}$ which has $s^{k-1}$ as its counterpart observable by the principal) should receive the promised utility under his maximizing strategy. Note that here we allow the agent who deviated in the past to deviate again. $\forall h^{k-1} \in H^{k-1}(s^{k-1})$, $h^{k-1} \neq s^{k-1}$:

[7] We will abuse on notation writing $s^{k-1}$ to denote the history $h^{k-1}$ in which $\widehat{a}_j = a_j$ for all $j = 1, ..., k-1$

[8] Abusing again on notation, $H^{k-1}\left(s^{k-1}\right)$ is the inverse of the function $s^{k-1}$ that maps full histories $h^{k-1}$ into their counterpart that is observable by the principal.

[9] If $k = 1$, then we should define this constraint as an inequality, and it will correspond to the participation constraint.

$$(6)\quad \underset{\delta_a}{max}\ \sum_{t=k}^{T}\beta^{t-k}\left[\sum_{H^t\left(s^{k-1}\right)} p\left(h^t|\pi,\delta_a,h^{k-1}\right)u\left(c_t,\delta_a\left(a_t,h^{t-1}\right)\right)\right]=w_{k,s^{k-1}}\left(h^{k-1}\right)$$

**Obedience constraints:**

At every node on the equilibrium path, the agent should find it at least as attractive to follow the principal's recommendation as to deviate. $\forall l\in\{k,...,T\},\ s^{l-1},\ \delta_a$:

(7)

$$\sum_{t=l}^{T}\beta^{t-l}\left[\sum_{H^t} p\left(h^t|\pi,\delta_a,s^{l-1}\right)u\left(c_t,\delta_a\left(a_t,h^{t-1}\right)\right)\right]\leq\sum_{t=l}^{T}\beta^{t-l}\left[\sum_{S^t} p\left(s^t|\pi,s^{l-1}\right)u\left(c_t,a_t\right)\right]$$

Next, we write the continuation problem in a backward induction formulation. First, let us define some useful concepts. Given an outcome function $\pi$ and an on-path continuation node $s^k$, we can define the vector of utility promises at $s^k$, $w_{k+1,s^k}(\pi):H^k\left(s^k\right)\to\mathbb{R}$. For the agent on the equilibrium path, the future utility at $s^k$ will then be:

$$w_{k+1,s^k}(\pi)\left(s^k\right)=\sum_{t=k+1}^{T}\beta^{t-k-1}\left[\sum_{S^t} p\left(s^t|\pi,s^k\right)u\left(c_t,a_t\right)\right]$$

And for the agent off-path at node $h^k\neq s^k$, his future utility will be:

$$w_{k+1,s^k}(\pi)\left(h^k\right)=\underset{\delta_a}{max}\ \sum_{t=k+1}^{T}\beta^{t-k-1}\left[\sum_{H^t\left(h^k\right)} p\left(h^t|\pi,\delta_a,h^k\right)u\left(c_t,\delta_a\left(a_t,h^{t-1}\right)\right)\right]$$

**Proposition 2.** *For any $\left(s^{k-1},w_{k,s^{k-1}}\right)$, there is an optimal contract $\pi^*$ to the problem $V_k\left(s^{k-1},w_{k,s^{k-1}}\right)$ such that, for every continuation node $s^k$, from $s^k$ onward $\pi^*$ is an optimal contract to the problem $V_{k+1}\left(s^k,w_{k+1,s^k}\left(\pi^*\right)\right)$.*

*Proof.* (Outline) To prove this proposition, we depart from an optimal outcome function $\pi_k$ that solves problem $V_k\left(s^{k-1},w_k\right)$, and an optimal $\pi_{k+1}$ that solves $V_{k+1}\left(s^k,w_{k+1,s^k}\left(\pi^*\right)\right)$

for some continuation node $s^k$. We then define an outcome function $\pi^*$ that is equal to $\pi_k$ everywhere except in the branch of the contract that starts in $s^k$. In this branch, $\pi^*$ is equal to $\pi_{k+1}$. This new outcome function $\pi^*$, from $s^k$ onward, is an optimal contract for $V_{k+1}\left(s^k, w_{k+1,s^k}\left(\pi^*\right)\right)$ by construction. $\pi^*$ is also an optimal contract for problem $V_k\left(s^{k-1}, w_k\right)$, because it respects constraints (5), (6) and (7) and delivers at least the same surplus as $\pi_k$ (because $V_{k+1}\left(s^k, w_{k+1,s^k}\left(\pi^*\right)\right)$ cannot be lower than the surplus delivered by $\pi_k$ from $s^k$ onward, otherwise $\pi_{k+1}$ would not be optimal).

By repeating this process for every continuation node $s^k$ (there is a finite number of them), we construct the optimal contract we are looking for.

$\square$

Proposition 2 allows us to restate the maximized utility of the principal as follows:

$$V_k\left(s^{k-1}, w_k\right) = \sum_{S^k} p\left(s^k|\pi^*, s^{k-1}\right)\left\{(q_k - c_k) + \alpha . V_{k+1}\left(s^k, w_{k+1,s^k}\left(\pi^*\right)\right)\right\}$$

It is then natural to think of the principal's optimization problem as choosing over outcome functions only for period $k$, $\pi\left(a_k|s^{k-1}\right)$ and $\pi(c_k|a_k, q_k, s^{k-1})$; and vectors of continuation utility promises $w_{k+1,s^k}$. It is possible to apply this procedure if the continuation value functions $V_{k+1}\left(s^k, w_{k+1,s^k}\left(\pi^*\right)\right)$ are known in advance, which is why the problem needs to be solved backwards.

To be clear, we state the continuation problem in its backward induction form:

**Problem 4.** *The principal's backward induction continuation problem at node $s^{k-1}$ of the equilibrium path is:*

$$V_k\left(s^{k-1}, w_k\right) = \underset{\pi\left(a_k|s^{k-1}\right), \pi\left(c_k|a_k, q_k, s^{k-1}\right), \left\{w_{k+1,s^k}\right\}_{s^k}}{max} \sum_{S^k} p\left(s^k|\pi, s^{k-1}\right)\left\{(q_k - c_k) + \alpha . V_{k+1}\left(s^k, w_{k+1,s^k}\right)\right\}$$

*subject to constraints* (8)*,* (9) *and* (10)*, defined below.*

Below we define the constraints mentioned:

**Promise-keeping constraint:**

The outcome function $\pi$ has to deliver the agent the promised utility:

(8)
$$\sum_{S^k} p\left(s^k|\pi, s^{k-1}\right)\left\{u\left(c_k, a_k\right)+\beta.w_{k+1,s^k}\left(s^k\right)\right\}=w_{k,s^{k-1}}\left(s^{k-1}\right)$$

**Threat-keeping constraint:**

The agent who reaches node $s^{k-1}$ having deviated previously should receive the promised utility. $\forall h^{k-1} \in H^{k-1}(s^{k-1})$, $h^{k-1} \neq s^{k-1}$:

(9)
$$\underset{\delta_a}{max} \sum_{H^k\left(s^{k-1}\right)} p\left(h^k|\pi, \delta_a, h^{k-1}\right)\left\{u\left(c_k, \delta_a\left(a_k, h^{k-1}\right)\right)+\beta.w_{k+1,s^k}\left(h^k\right)\right\}=w_k\left(h^{k-1}\right)$$

**Obedience constraint:**

The agent should find it at least as attractive to follow the principal's recommendation as to deviate in period $k$. $\forall \delta_a : A \to A$:

(10)
$$\sum_{H^k\left(s^{k-1}\right)} p\left(h^k|\pi, \delta_a, s^{k-1}\right)\left\{u\left(c_k, \delta_a\left(a_k, h^{k-1}\right)\right)+\beta.w_{k+1,s^k}\left(h^k\right)\right\} \leq w_{k,s^{k-1}}\left(s^{k-1}\right)$$

In practice, $W_{k,s^{k-1}}$ (the space of feasible vectors of continuation utilities) is not known in advance - it needs to be calculated previously, when we solve the next-period problem $V_{k+1}\left(s^k, w_{k+1,s^k}\right)$.

## 2.4 Linear program

Finally, in this subsection we formulate the problem using the linear programming framework first introduced by (PRESCOTT; TOWNSEND, 1984). The advantage of linear programs is that they are well understood and can be solved easily on a computer using off-the-shelf optimization packages[10].

Problem 4 is not a linear program, because choice variables $\pi\left(a_k|s^{k-1}\right)$, $\pi(c_k|a_k, q_k, s^{k-1})$ and $w_{k+1,s^k}$ are multiplied in the objective function. So we restate the problem in a slightly different manner: we allow for lotteries over promised utility vectors $w_{k+1,s^k}$ and make the probability distribution $\pi\left(w_{k+1,s^k}|s^k\right)$ the choice variable for the principal instead of the actual vectors $w_{k+1,s^k}$ [11]. We also discretize the utility spaces $W_{k,s^{k-1}}$ in order to have a finite number of variables. The program is made linear by choosing over joint probability distributions $\pi(a_t, q_t, c_t, w_{t+1,s^t}|s^{t-1}, w_t) \equiv \pi\left(s_t, w_{t+1,s^t}|s^{t-1}, w_{t,s^{t-1}}\right)$ instead of the conditional probabilities as before, and imposing that $\pi$ is a probability measure that respects the conditional probabilities given by $f\left(q_t|a_t, h^{t-1}\right)$.

For a detailed explanation of this linear programming framework, we refer the reader to the excellent introduction in (PRESCOTT, 1999), to its first application to dynamic moral hazard in (PHELAN; TOWNSEND, 1991) and to the original article (PRESCOTT; TOWNSEND, 1984).

**Problem 5.** *Linear program in period* t*:*

$$V_t\left(s^{t-1}, w_{t,s^{t-1}}\right) = \max_{\pi} \sum_{s_t, w_{t+1,s^t}} \pi(s_t, w_{t+1,s^t}|s^{t-1}, w_t)\left\{(q_t - c_t) + \beta V_{t+1}\left(s^t, w_{t+1,s^t}\right)\right\}$$

*subject to constraints* (11)*,* (12)*,* (13)*,* (14) *and* (15) *defined below.*

Below we define the mentioned constraints:

**Probability measure constraints[12]:**

$$\begin{gathered}\sum_{s_t, w_{t+1,s^t}} \pi\left(s_t, w_{t+1,s^t}|s^{t-1}, w_{t,s^{t-1}}\right) = 1 \\ \pi\left(s_t, w_{t+1,s^t}|s^{t-1}, w_{t,s^{t-1}}\right) \geq 0 \;, \forall s_t, w_{t+1,s^t}\end{gathered} \tag{11}$$

[10] In our case, we have used Matlab's built-in optimization toolbox.

[11] Note that this is in fact more general, since choosing a specific vector $w_{k+1,s^k}$ is equivalent to choosing $\pi\left(w_{k+1,s^k}|s^k\right) = 1$.

[12] Probability measure constraints are also needed in the formulations of the previous sections, but we make them explicit here to have the complete linear program as it is computed.

**Promise-keeping constraint:**

$$(12)\qquad \sum_{s_t, w_{t+1,s^t}} \pi\left(s_t, w_{t+1,s^t} | s^{t-1}, w_{t,s^{t-1}}\right)\left\{u\left(c_t, a_t\right) + \beta w_{t+1,s^t}\left(s^t\right)\right\} = w_{t,s^{t-1}}\left(s^{t-1}\right)$$

**Threat-keeping constraints**[13]**:**

(13) $\forall h^{k-1} \in H^{k-1}(s^{k-1}),\, h^{k-1} \neq s^{k-1}$

$$\underset{\delta_a}{max} \sum_{s_t, w_{t+1,s^t}} \pi\left(s_t, w_{t+1,s^t} | s^{t-1}, w_{t,s^{t-1}}\right) \frac{f\left(q_t | \delta_a\left(a_t\right), h^{t-1}\right)}{f\left(q_t | a_t, s^{t-1}\right)}.$$

$$.\left\{u\left(c_t, \delta_a\left(a_t\right)\right) + \beta w_{t+1,s^t}\left(\delta_a\left(a_t\right), q_t, h^{t-1}\right)\right\} = w_{t,s^{t-1}}\left(h^{t-1}\right)$$

**Subjective production function constraints:**

$\forall \overline{q}_t, \overline{a}_t$ :

(14)

$$\sum_{c_t, w_{t+1,s^t}} \pi\left(\overline{a}_t, \overline{q}_t, c_t, w_{t+1,s^t} | s^{t-1}, w_{t,s^{t-1}}\right) = f\left(\overline{q}_t | \overline{a}_t, s^{t-1}\right) . \sum_{q_t, c_t, w_{t+1,s^t}} \pi\left(\overline{a}_t, \overline{q}_t, c_t, w_{t+1,s^t} | s^{t-1}, w_{t,s^{t-1}}\right)$$

**Obedience constraint:**

$\forall a_t, \widehat{a}_t$ :

[13]In fact, the maximization over $\delta_a$ is not included in the program. Instead, for every $\delta_a$ we impose a common utility bound $w_{t,s^{t-1}}\left(h^{t-1}\right)$ on the utility of the deviating agent (an inequality constraint), and after solving the program with these constraints we calculate under which $\delta_a$ the utility of the deviating agent was highest, and the actual value of this maximal utility. We then substitute $w_{t,s^{t-1}}\left(h^{t-1}\right)$ by this actual value. This is important because when solving the previous period we need to know the value of the utility actually delivered to the deviating agent, and not only an upper bound.

$$\text{(15)} \quad \sum_{s_t, w_{t+1,s^t}} \pi\left(s_t, w_{t+1,s^t} | s^{t-1}, w_{t,s^{t-1}}\right) \left\{u\left(c_t, a_t\right) + \beta w_{t+1,s^t}\left(s^t\right)\right\} \geq$$

$$\geq \sum_{s_t, w_{t+1,s^t}} \pi\left(s_t, w_{t+1,s^t} | s^{t-1}, w_{t,s^{t-1}}\right) \frac{f\left(q_t | \widehat{a}_t, s^{t-1}\right)}{f\left(q_t | a_t, s^{t-1}\right)} \left\{u\left(c_t, \widehat{a}_t\right) + \beta w_{t+1,s^t}\left(\widehat{a}_t, q_t, s^{t-1}\right)\right\}$$

Two observations are important to clarify how to implement this algorithm:

**First period (** $t = 1$ **)**

In the initial period, there is no possible prior deviation, so there is no need for threat keeping constraints[14]. The vector of utility promises reduces to a scalar, the outside option of the agent upon signing the contract. The promise-keeping constraint (12) is imposed as an inequality instead of an equality.

**Last period (** $t = T$ **)**

In the last period, incentives are static (there are no utility promises), which simplifies the surplus of the principal, the utility of the agent and the set of choice variables (there is no continuation utility vector).

## 2.5 Learning

In the linear program above we have intentionally left history-dependent subjective production functions $f\left(q_t | a_t, h^{t-1}\right)$ unmodeled. This indicates that this formulation can account for any kind of full history dependence of the production function.

Here we are interested in modeling a case in which the true distribution of output conditional on effort is invariant and independent across periods, but unknown to both the agent and the principal. The subjective probabilities are then history dependent (through the Bayesian updates). However, this framework could also account for other sorts of history dependence, such as correlation of output across periods, or dependence of current output on previous actions (see (FERNANDES; PHELAN, 2000) for example). This formulation can also encompass the case of no history dependence such as (PHELAN;

[14]In fact, this is already embedded in the notation, since there are no $h^0 \in H^0(s^0)$ such that $h^0 \neq s^0$ (remember we defined $h_0 = s_0 = \emptyset$).

TOWNSEND, 1991), although in a computationally inefficient way when compared to their algorithm. In section 3 we use this to test our code.

Also, different types of uncertainty can be modeled. A non-exhaustive list can be found below. The kind of uncertainty to be modeled is key to what class of real-world problems we are studying. A few questions that could be asked using this framework are:

- Does effort matter and how much? (does $f$ shift or change its dispersion with effort level?)
- What is the quality of the project? (what is the mean of $f$ for a given effort level?)
- How risky is the project? (how dispersed is $f$ for a given effort level?)

The specificity of the problem at hand then lies in how to model the conditional distribution of output, $f\left(q_t|a_t, h^{t-1}\right)$. In this paper we will make the following modeling choice:

There are $n$ possible states of nature, $\theta \in \Theta = \{1, \ldots, n\}$, each $\theta$ referring to a different (and publicly known) production function $f_\theta\left(q|a\right)$. $\theta$ is realized before the contract is signed, but neither the principal nor the agent can observe it. There is no ex-ante asymmetry of information: upon signing the contract, principal and agent share the same belief $\mu\left(h^0\right) = \left(\mu_1, ...\mu_n\right) \in \mathbb{R}^n$ about the state of the world, that is:

$$\mu_i\left(h^0\right) \equiv p(\theta = i)$$

As output is revealed, principal and agent update their beliefs about the state of nature using Bayes' rule:

$$\mu_i\left(h^{t-1}\right) \equiv p\left(\theta = i|h^{t-1}\right) \underset{\text{Bayes}}{=} \frac{p\left(\theta = i, h^{t-1}\right)}{p\left(h^{t-1}\right)} = \frac{p\left(\theta = i, h^{t-1}\right)}{\sum_{j=1}^{n} p\left(\theta = i, h^{t-1}\right)} = \tag{16}$$

$$\underset{\text{i.i.d.}}{=} \frac{p\left(\theta = i\right) . \prod_{s=1}^{t-1} f_i\left(q_s|a_s\right)}{\sum_{j=1}^{n} p\left(\theta = j\right) . \prod_{s=1}^{t-1} f_j\left(q_s|a_s\right)} = \frac{\mu_i\left(h^0\right) . \prod_{s=1}^{t-1} f_i\left(q_s|a_s\right)}{\sum_{j=1}^{n} \mu_j\left(h^0\right) . \prod_{s=1}^{t-1} f_j\left(q_s|a_s\right)}$$

At time $t \in \{1, ..., T\}$, the belief $\mu\left(h^{t-1}\right)$ allows us then to calculate the subjective production function:

$$f\left(q_t|a_t, h^{t-1}\right) = \sum_{i=1}^{n} \mu_i\left(h^{t-1}\right) f_i\left(q_t|a_t\right)$$

It is worth emphasizing that, since the principal does not observe actions, he will update his beliefs considering that the agent has followed his recommendations, whereas the agent will consider the actual levels of effort applied, and so beliefs can diverge from $t = 2$ on.

It is interesting to note from equation (16) that the order in which $(a_t, q_t)$ pairs occurred does not matter to the update of beliefs. This is a direct consequence of our assumption that the underlying production function is i.i.d.. This fact will allow us later to reduce the dimensionality of the computational problem to be solved.

## 2.6 Computation

The optimal contract is computed by backward induction. We start from the last period ($T$), solving one static linear program 5 for each possible history $s^{T-1}$ and candidate vector of utility promises $w_{T,s^{T-1}}$. For a portion of those linear programs, the computer will find a solution, which allows us to define the spaces of feasible vectors of utility promises $W_{T,s^{T-1}}$, as well as the associated value functions $V_T\left(s^{T-1}, w_{T,s^{T-1}}\right)$. With these values stored, we can proceed to calculate the linear programs for period $T-1$, and so on.

The computational burden of this algorithm increases exponentially with the number of periods and the size of grids. The number of periods to be solved impacts on the number of possible histories (and thus linear programs) that need to be considered, as well as the number of deviations that we need to keep track of (since multiple deviations are a concern). The number of possible deviations increases the dimensionality of the spaces of feasible utility promises $W_{t,s^{t-1}}$, which in turn increases the number vectors of utility promises to be tested in period $t$ (each vector requires solving a linear program), as well as the number of constraints (each threat-keeping constraint refers to a possible prior deviation). Also, this raises the number of variables of the linear programs in $t-1$ (remember that choice variables are probability distributions $\pi$ over values of action, consumption, output and *utility promises*).

Some simplifications are however possible to make the algorithm more computationally efficient:

- Not all the history $s^t$ matters: consumption doesn't affect beliefs. Considering only combinations of past actions and outputs as past histories greatly reduces the number of linear programs to be solved.

- Not all possible histories $s^t$ or $h^t$ matter: because of the i.i.d. hypothesis which we commented above, the order in which the histories occurred is irrelevant when updating beliefs. We can thus consider only one instance of all possible permutations of a given history, reducing further the number of linear programs to be solved. For example, if $A = \{a_L, a_H\}$ and $Q = \{q_L, q_H\}$, histories $s^2 = \{(a_H, q_L), (a_H, q_H)\}$ and $\hat{s}^2 = \{(a_H, q_L), (a_H, q_H)\}$ are equivalent in terms of what is learned from them, and so the continuation problems starting from those nodes are treated as one.

# 3 Implementation and Validation

The linear program formulated above has been implemented on Matlab. In order to validate it, we have performed three tests. If this algorithm is to be used more extensively, more validation work should be done to further understand the limitations of this computational approach.

The model proposed in the previous section was specialized as follows: the agent's instantaneous Bernoulli utility function was chosen as the following separable CRRA:

$$u(c, a) = \frac{-c^{-0.5}}{0.5} - g(a)$$

To limit the computational time required, we have chosen a minimalist approach to grids. There are two periods ($T = 2$), actions can take two values (low and high effort, $A = \{a_L, a_H\}$, with $g(a_L) = 1$ and $g(a_H) = 1.5$), as well as outputs (failure or success,

$Q = \{q_L = 0.5, q_H = 15\}$). The consumption grid $C$ was chosen to have 100 equally spaced points between 0.1 and 16. We have equalized the discount rates of the principal and the agent to be $\alpha = \beta = 0.95$.

Uncertainty about the production function was modeled differently in each case analyzed, so both $\Theta$ and the associated production functions $f_\theta(q|a)$ will be specified later.

## 3.1 Comparing first and second best outcomes

When the agent is risk-averse, moral hazard creates a tension between insurance and the provision of incentives, even without the learning process we propose here.

In the first best (that is, when the principal can contract effort directly), it is efficient to give the agent a non-contingent compensation. In the second best this is not the case, because under a non-contingent compensation the agent has no incentives to provide effort. The Pareto frontier of the second best should shrink as compared to the first best, because the problem is being constrained by incentive compatibility.

In this subsection, we test whether our algorithm reproduces these analytical results. To calculate the first best, it is sufficient to remove the informational constraints from the problem (that is: obedience and threat-keeping constraints).

Here we imposed $\Theta = \{1, 2\}$, and the candidate production functions as in Table 1.

| $f_1(q\|a)$ | $q_L$ | $q_H$ |
|---|---|---|
| $a_L$ | 0.8 | 0.2 |
| $a_H$ | 0.2 | 0.8 |

| $f_2(q\|a)$ | $q_L$ | $q_H$ |
|---|---|---|
| $a_L$ | 0.8 | 0.2 |
| $a_H$ | 0.8 | 0.2 |

Table 1: Candidate production functions $f_1$and $f_2$

The common prior imposed was $\mu_1 = \mu_2 = 0.5$.

In Figure 2, we can see observe two expected features. First, the range of agent's reservation utilities for which it is optimal to recommend high effort with probability one is narrower. This happens because the cost of providing incentives reduces the gains from high effort in the second best. Second, we can see that the Pareto frontier of the second best is everywhere below that of the first best, which is due to the efficiency cost from information asymmetry.

Then, in Figures 3 and 4, we see a few interesting features. First, both on the first best and in the second best, it is optimal to offer the agent a fixed compensation when low effort is recommended (points on the right of each figure, superposed). Second, when high effort is recommended (points on the left of each figure), in the first best we still have a fixed salary to the agent, while in the second best there is incentive provision through consumption: there is a "bonus" on consumption when high output $q_H$ occurs.

## 3.2 Impact of uncertainty on incentives

(PRAT; JOVANOVIC, 2012) study a problem that is similar to ours. In their case, the agent has unknown ability, which is a parameter of the production function. In a two-period model[15], they show that when the prior has more variance (that is, when the uncertainty about the production function is higher), it is necessary to provide more high-powered incentives in order to make the agent exert effort. Intuitively, this happens because when the prior is more uncertain there is more room for the agent to manipulate the principal's beliefs about the production function. This increases the information rents the agent can get in the second period after deviating in the first period. To curb this behavior, the principal needs to make the agent's compensation fluctuate more with output.

This simple model provides a lower bound on the volatility of the agent's expected utility in the first period when the contract implements high effort in both periods. Volatility increases with the variance of the prior over the production function. In this subsection, we calculate analytically this lower bound, and use it to test our algorithm.

We extend slightly the two-period model in (PRAT; JOVANOVIC, 2012) to include the possibility of nonlinear disutility of effort (the agent's Bernoulli utility function here is $[v(w) - g(a)]$, with $g$ increasing and possibly convex), and reach the following condition for the optimal contract that implements high effort in both periods:

[15] See appendix C of their paper

$$U\left(a_H, q_H\right) - U\left(a_H, q_L\right) \geq \frac{1}{\eta_\mu}\left[\frac{g(a_H) - g(a_L)}{a_H - a_L}\right]\left[\frac{\beta a_H Var_\mu(\eta)}{(\eta_\mu - a_H E_\mu[\eta^2])} + 1\right]$$

Where $\eta$ is (unknown) ability, $\mu$ is the prior about ability, $\eta_\mu \equiv E_\mu\left[\eta\right]$ , $Var_\mu(\eta)$ and $E_\mu[\eta^2]$ are moments calculated from the prior distribution $\mu$. $U\left(a, q\right)$ is the expected utility in $t = 1$ after history $(a, q)$ is realized.

Once again we have $\Theta = \{1, 2\}$, and now candidate production functions are as in Table 2.

| $f_1\left(q\|a\right)$ | $q_L$ | $q_H$ |
|---|---|---|
| $a_L$ | $1 - 0.1\eta_L$ | $0.1\eta_L$ |
| $a_H$ | $1 - \eta_L$ | $\eta_L$ |

| $f_2\left(q\|a\right)$ | $q_L$ | $q_H$ |
|---|---|---|
| $a_L$ | $1 - 0.1\eta_H$ | $0.1\eta_H$ |
| $a_H$ | $1 - \eta_H$ | $\eta_H$ |

Table 2: Candidate production functions $f_1$and $f_2$

We impose $\mu_1 = \mu_2 = 0.5$ as before. By modifying parameters $(\eta_L, \eta_H)$ we are able to vary the level of uncertainty in the prior, while keeping the mean production function unchanged.

For all the different parameters tested ($(\eta_L, \eta_H) \in \{(0.2, 0.8), (0.3, 0.7), (0.4, 0.6), (0.45, 0.55)\}$), the lower bound was respected as Figure 5 shows.

## 3.3 Removing uncertainty

In our setting, when there is no uncertainty about the production function (that is, when $\mu_\theta = 1$ for some $\theta \in \Theta$), we reach a special case that has already been studied by (PHELAN; TOWNSEND, 1991). It is then possible to formulate the problem recursively and computation can be greatly simplified, as they show. Although our method in this paper is not computationally efficient to deal with cases without uncertainty about the production function, it should yield the same results as the algorithm in (PHELAN; TOWNSEND, 1991).

To test whether this is the case, we have computed the same case as in Table 1, except this time with $\mu_1 = 1$ (the production function is $f_1(q|a)$ with certainty). We solved this case using the algorithm in (PHELAN; TOWNSEND, 1991) too.

In Figure 6, we can see that both algorithms deliver the same results[16].

# 4 Applications

The algorithm described allows us to compute the optimal contract in a variety of situations, and identify its key characteristics. First, in subsection 4.1 we explore a case in which the choice of candidate production functions simplifies the learning process, confirming the conclusion in (PRAT; JOVANOVIC, 2012) that learning harms incentives. Then, we proceed in subsection 4.2 to analyze a case in which both principal and agent do not know for sure whether effort matters or not for production.

## 4.1 Belief manipulation

In this section we apply our methods to the case in 3.2, choosing $(\eta_L, \eta_H) = (0.2, 0.8)$ and $\mu_1(s^0) = \mu_2(s^0) = 0.5$. We then have the candidate production functions as in Table 3.

| $f_1(q\|a)$ | $q_L$ | $q_H$ |
|---|---|---|
| $a_L$ | $0.98$ | $0.02$ |
| $a_H$ | $0.8$ | $0.2$ |

| $f_2(q\|a)$ | $q_L$ | $q_H$ |
|---|---|---|
| $a_L$ | $0.92$ | $0.08$ |
| $a_H$ | $0.2$ | $0.8$ |

Table 3: Candidate production functions $f_1$and $f_2$

In what follows we will say that the belief is more "optimistic" if $\mu_2$ is greater, since under $f_2$ the expected output conditional on high effort is greater than under $f_1$.

What is interesting about this case is that no divergence in beliefs is created between the principal and an agent who deviates in case high output happens, as is made clear

[16]In fact, if we compare contract details such as consumption and utility promises conditional on action and output, we find that both solutions are very similar. Those comparisons have been omitted for brevity. Small differences can occur because the algorithm in this paper needs to solve larger linear programs, that sometimes exceed the maximum number of iterations in Matlab and thus do not yield results.

in Table 4 with posterior beliefs. On the other hand, when high effort is recommended and low output occurs, the deviating agent becomes more optimistic than the principal ($\mu_2\left(\hat{h}^1\right) = 0.49$ while $\mu_2\left(s^1\right) = 0.2$). This simplifies the analysis, as will become clearer.

| | $(a_L, q_L)$ | $(a_L, q_H)$ | $(a_H, q_L)$ | $(a_H, q_H)$ |
|---|---|---|---|---|
| $\mu_2\left(s^1\right)$ | $0.49$ | $0.8$ | $0.2$ | $0.8$ |
| $\mu_2\left(\hat{h}^1\right)$ | $0.2$ | $0.8$ | $0.49$ | $0.8$ |

Table 4: Posterior conditional on $s^1 = (a_L, q_L)$ for prior $\mu_1\left(s^0\right) = \mu_2\left(s^0\right) = 0.5$. The second row indicates $\mu_2\left(\hat{h}^1\right)$, the posterior of the agent who deviated in the first period.

We want to understand the role played by learning on incentives, and so in this subsection we will frequently compare the solution to the case described above to its counterpart with no uncertainty about the production function. To pick a comparable case, we need to have the same ex-ante distribution of probabilities, so we impose the production function $\widetilde{f}\left(q|a\right)$ in Table 5.

| $\widetilde{f}\left(q\|a\right)$ | $q_L$ | $q_H$ |
|---|---|---|
| $a_L$ | $0.95$ | $0.05$ |
| $a_H$ | $0.5$ | $0.5$ |

Table 5: "Expected production function": production function with same ex-ante probabilities as the combination of the production functions in Table 3 with prior $\mu_1\left(s^0\right) = \mu_2\left(s^0\right) = 0.5$.

The first thing we highlight is the difference between the spaces of feasible utility promises in the second period in both cases considered. Without learning, it is not possible to deliver different utilities to the agents on and off the equilibrium path: since both have the same beliefs over future productivity, they solve the exact same problem and so there is no margin to punish the agent off-path while holding the on-path agent's utility constant. Neither is it possible to have different spaces of vectors of utility promises for different first-period histories. The spaces of feasible utility vectors $W_{2,s^1}$, for all histories $s^1$, are exactly the same, and lie in the diagonal of the Cartesian as can be seen in Figure 8.

With learning, on the other hand, the space of feasible utility promises is "thick" for low output shocks, but not for high output shocks, as can be seen in Figure 7. This is a consequence of the peculiar candidate production functions chosen in this example. In a more general case such as that in the next subsection, the space of feasible utility promises is "thick" for every first-period history because beliefs on and off-path always diverge, as can be seen in Figure 15.

The difference between Figures 7 and 8 carries, in fact, the essence of why it is possible to formulate the problem recursively in the case with no learning about the production function, while the same does not apply to the case with learning. With learning, future preferences are affected by past histories and so it is necessary to know in which node of the equilibrium path we are when we solve the continuation problem.

Then we proceed to examine the impact of learning on incentives. In Figures 9 and 10 we can see the optimism of the off-path agent regarding future utilities when a low output shock occurs. When high output occurs there is no distinction between obedient and deviating agents, but when low output occurs the deviating agent expects higher utility from the next period.

The same does not happen with utility from current consumption as we see in Figures 11 and 12, since both the deviating and the obedient agent have the same beliefs in $t = 1$.

This is why in this case, in order to implement high effort in the first period, the principal has to increase the volatility of the agent's utility, as (PRAT; JOVANOVIC, 2012) argues. Figure 13 makes this evident. We will see in section 4.2, that this appears to hold in more general settings.

It is then natural to ask whether this increased volatility comes at a cost on surplus. Figure 14 shows that the surplus loss associated to learning is non-negative and increases with the reservation utility of the agent (because it becomes more costly to provide incentives through consumption for higher utility values). When the probability of high effort being recommended $p\left(a_1 = a_H\right)$ decreases, so does the surplus loss, since incentives are not needed for low effort. Learning also impacts effort and the participation constraint, but this will become clearer in the next subsection.

## 4.2 Effort matters vs. effort does not matter

In this subsection, we explore in detail the computed optimal contract for a two-period environment in which the uncertainty about the production function is as in Table 1.

This means that if $\theta = 1$, effort matters: under the chosen parameters, $a_H$ would be recommended for a wide range of the agent's reservation utilities if $\mu_1(s^0) = 1$. On the other hand, if $\theta = 2$, effort is irrelevant to production: it would be optimal to provide the agent with a fixed salary if $\mu_2(s^0) = 1$, encouraging low effort.

We impose $\mu_1(s^0) = \mu_2(s^0) = 0.5$, so that learning can occur between the first and the second periods.

Given the prior and the production functions, we can easily calculate the posteriors conditional on each possible on-path first-period history $s^1$, as showed in Table 6. We can see in particular that, since under low effort $a_L$ both production functions $f_1$ and $f_2$ are identical, there is no learning: the posterior is identical to the prior.

| | $(a_L, q_L)$ | $(a_L, q_H)$ | $(a_H, q_L)$ | $(a_H, q_H)$ |
|---|---|---|---|---|
| $\mu_1(s^1)$ | $0.5$ | $0.5$ | $0.2$ | $0.8$ |
| $\mu_1\left(\hat{h}^1\right)$ | $0.2$ | $0.8$ | $0.5$ | $0.5$ |

Table 6: Posterior conditional on $s^1$ for $\mu_1(s^0) = \mu_2(s^0) = 0.5$. The second row indicates $\mu_1\left(\hat{h}^1\right)$, the posterior of the agent who deviated in the first period . Each column corresponds to a different pair of (recommended effort, realized output).

The divergence between beliefs on and off the equilibrium path is what makes this problem complex. As described earlier, the first step in solving the problem is to solve the linear programs in $t = 2$ and find the spaces of feasible utility promises for each possible history $s^1$. Figure 15 depicts the calculated spaces of feasible utility promises for this case. History dependence becomes clear, since we can see that the space associated to each possible history $s^1$ is different.

We can see that for the same on-path utility promise (x axis in each graph), there are multiple feasible values of off-path utility (y axis). Those values impact how tight the *threat-keeping* constraints in period 2 will be. An interesting question is then: what is the strategy used by the principal to make the off-path agent worse off while keeping the on-path agent's utility constant? The numerical solution allows us to explore this in detail. Figure 16 illustrates the contracts in two cross sections of the spaces of feasible utility promises (for first-period histories $s^1 = (a_H, q_H)$ and $s^1 = (a_H, q_L)$).

We can see in Figure 16 that the principal uses the difference in perceived probabilities (of high and low output) on and off path to punish the off-path agent, by changing the consumption plans (increasing or decreasing its volatility). Of course, tightening the threat-keeping constraint comes at a cost on surplus. Interesting to note is the ex-post inefficiency of such a schedule: once the second period is reached (and the principal knows the agent is on the equilibrium path), there is room for a Pareto improvement by renegotiating the contract, since the principal can increase his surplus at no utility cost to the agent on the equilibrium path. However, the ability to commit to non-renegotiation can bring ex-ante gains to the contractual relationship.

**Contract in** $t = 1$

With the solution to all $t = 2$ programs, it is then possible to compute the contract in $t = 1$. Through the choices of utility promises of the optimal contract, we can recover the whole equilibrium path.

Figure 17 shows us a few general features of the numerical solution. Worth noting are: the concavity of the value function (surplus, or "Pareto frontier"); the participation constraint does not bind for low levels of reservation utility of the agent; for high values of reservation utility it becomes too costly to implement high effort.

**Multiple deviations**

In Figure 18, we can see the level of effort in the second period on and off the equilibrium path when high effort is recommended in the first period ($a_1 = a_H$), for different realizations of the first-period output $q_1$. It is interesting to note that under learning the agent considers multiple deviations: for high reservation utility values, an agent who deviates in the first period and receives a good output shock will find it attractive to deviate again in the second period. This is because he will be more pessimistic about the production function in the second period than the principal. From Table 6, we see that for $s^1 = (a_H, q_H)$ the principal has posterior $\mu_1(s^1) = 0.8$ while the agent who deviated has posterior $\mu_1\left(\hat{h}^1\right) = 0.5$. That is, the principal believes effort is more important than the agent does.

Compared to the case in subsection 4.1, this is a new fact. There, the agent would only consider one-shot deviations, because under any first period history a deviating agent would be at least as optimistic as the agent on the equilibrium path[17]. Here, instead, the deviating agent can be more pessimistic than the agent on-path, and so incentives might be less than necessary in the second period and the deviating agent shirks again[18].

When confronted with the decision of whether to provide the level of effort recommended by the principal or not in $t = 1$, the agent analyzes his expected utility under each course of action. The role of the incentive constraints in the program is to make sure that following the recommended level of effort is at least as desirable as deviating.

We can see in Figure 19 (top panel) that the incentive constraint is always tight when high effort is recommended. This is expected, since providing incentives for high effort to a risk-averse agent is costly and so the principal avoids any slack in the incentive constraint.

The numerical solution allows us to go farther in this analysis: since we assumed the agent's utility is separable, first period incentives can be broken down into three components: the disutility of effort; the current utility from consumption; and the discounted second period utility. In Figure 19 (bottom panel) we can see how those three components change under recommended effort, and under deviation.

The difference in the disutility of effort for the obedient and the deviating agent is obvious - different effort levels imply different disutilities. When the principal wants to implement high effort, this gap in disutilities is what creates the need for incentives.

Incentives in $t = 1$ are provided in two ways: through current consumption, and through future utility promises.

To provide incentives through current consumption, the principal offers higher consumption when high output occurs than when low output occurs. Although obedient and deviating agents receive the same consumption level conditional on output, they differ in

[17] See appendix C in (PRAT; JOVANOVIC, 2012)

[18] When we graph second period average effort for the case in subsection 4.1 (the equivalent to Figure 18), we see that effort coincides for on path and off path agents. This graph has been omitted for brevity.

the probabilities of each output level occurring. Since the obedient agent expects high output with higher probability than the deviating agent, giving higher compensation if high output occurs creates an incentive for being obedient.

For the future utility component, incentives are more complex: not only do the obedient and deviating agents see different probabilities of each output occurring, but they also receive different utility levels conditional on output. This is because they will have divergent beliefs in the second period. Interestingly, there is a region in which the agent who deviates gets more future utility than the agent who is obedient (for low reservation utility values). This has to be compensated by the principal through more volatility of current consumption.

If we further decompose the agent's expected future utility (Figure 20), we can see that this case is more complex than that in subsection 4.1 (see Figure 9). Now the agent off-path expects different future utilities than the agent on-path both when high and low output occurs. For low reservation utilities of the agent, there is a region in which the deviating agent expects lower future utility if high output occurs in the first period than if low output occurs.

Ultimately, how the principal will make the optimal mix in $t = 1$ between providing incentives through consumption (front-loading) or through future utility (back-loading) is a quantitative issue of which is less costly.

**Comparison to the case with no uncertainty**

As in subsection 4.1, we compare the case with learning to its counterpart without learning, and check whether the same conclusions hold. Here, the production function that is ex ante equivalent is given by Table 7.

| $\widetilde{f}(q\|a)$ | $q_L$ | $q_H$ |
|---|---|---|
| $a_L$ | $0.8$ | $0.2$ |
| $a_H$ | $0.5$ | $0.5$ |

Table 7: "Expected production function": production function with same ex-ante probabilities as the combination of the production functions in Table 1 with prior $\mu_1(s^0) = \mu_2(s^0) = 0.5$.

When we look at the volatility of the agent's utility when high effort is recommended (Figure 21), we reach the same conclusion in this case: providing incentives requires higher volatility under learning.

In Figure 22, we note the loss in surplus from the uncertainty about the production function[19]. Also, interestingly, we can see the consequences of learning on the participation constraint, that becomes slack for higher reservation utilities of the agent when compared to the case with no learning. This is a consequence of the need for stronger incentives to implement high effort under learning: because there is a lower bound on the utility from consumption ($C$ is finite), higher volatility of consumption is only possible if average consumption increases.

Also, it is clear that the region where high effort is recommended is narrower under learning - it is more costly to implement high effort because of the need for stronger incentives.

# 5 Conclusion

In this paper we have proposed a computational method to deal with models of dynamic moral hazard with simultaneous learning about the production function. The contribution is methodological: through computation, the problem can be studied under few assumptions about functional forms. The literature has been dealing with this problem using continuous-time principal-agent models in which the production function has an unknown additive parameter, thus limited in scope.

We have formulated a general mechanism that can be tackled with our approach, and have shown how it can be reformulated as an incentive-compatible mechanism that can be solved by backward induction. Finally, we have written the backward induction problem in the linear programming framework of (PRESCOTT; TOWNSEND, 1984), which allows computation. The algorithm was simplified to increase computational performance: consumption histories were excluded from the space of state variables, and permutations

[19]However, there is a dip in surplus loss below zero which we believe can be attributed to the finiteness of the utility grid. This finiteness creates local risk-neutrality for the agent, which could permit increases in volatility at no cost on surplus.

in past histories (on and off the equilibrium path) were shown to be irrelevant. Those simplifications reduce the number of programs that need to be evaluated in each period, as well as the size of those linear programs.

Our algorithm was validated and applied to a few cases in a simple two-period, two-action, two-output, and two-production function environment. In one application, we have confirmed the findings of (PRAT; JOVANOVIC, 2012), that uncertainty about the production function increases the volatility of the agent's utility in order to compensate for gains from belief manipulation by the agent.

The numerical solution is very rich in details, and shows how complex the model becomes, as a result of diverging beliefs between the principal and the off-path agent. There is much room for future work in exploring different applications of this algorithm.

An existing gap in the literature, which we have not tackled, is to study in a uniform and accessible language the whole spectrum of commitment in the dynamic moral hazard problem with learning, and to investigate what happens under limited commitment. Up to now, on one end of the spectrum there are discrete models with little or no commitment in the "career concerns" literature. On the other end, there are full-commitment models in continuous time.

# References

DEMARZO, P.; SANNIKOV, Y. Learning , Termination , and Payout Policy in Dynamic Incentive Contracts. **Working Paper**, n. 0452686, 2011.

DOEPKE, M.; TOWNSEND, R. M. Dynamic mechanism design with hidden income and hidden actions. **Journal of Economic Theory**, v. 126, n. 1, p. 235–285, jan. 2006. ISSN 00220531. Disponível em: ⟨http://linkinghub.elsevier.com/retrieve/pii/S0022053104001954⟩.

FERNANDES, A.; PHELAN, C. A recursive formulation for repeated agency with history dependence. **Journal of Economic Theory**, v. 247, p. 223–247, 2000. Disponível em: ⟨http://www.sciencedirect.com/science/article/pii/S0022053199926194⟩.

HE, Z.; WEI, B.; YU, J. Optimal Long-term Contracting with Learning. 2013.

HOLMSTROM, B. Managerial Incentive Problems: A Dynamic Perspective. **Review of Economic Studies**, v. 66, n. 1, p. 169–182, jan. 1999. ISSN 0034-6527. Disponível em: ⟨http://restud.oxfordjournals.org/lookup/doi/10.1111/1467-937X.00083⟩.

KARAIVANOV, A. Computing Moral Hazard Programs With Lotteries Using Matlab. 2001.

KARAIVANOV, A.; TOWNSEND, R. M. Dynamic Financial Constraints : Distinguishing Mechanism Design from Exogenously Incomplete Regimes. 2012.

KILENTHONG, W. T.; MADEIRA, G. A. Observability and Endogenous Organizations. 2009.

LAFFONT, J.-J.; TIROLE, J. The Dynamics of Incentive Contracts. **Econometrica**, v. 56, n. 5, p. 1153–1175, 1988.

MADEIRA, G. D. A.; TOWNSEND, R. M. Endogenous groups and dynamic selection in mechanism design. **Journal of Economic Theory**, v. 142, p. 259–293, 2008.

MANSO, G. Motivating Innovation. **The Journal of Finance**, v. 66, n. 5, p. 1823–1860, out. 2011. ISSN 00221082. Disponível em: ⟨http://doi.wiley.com/10.1111/j.1540-6261.2011.01688.x⟩.

PAULSON, A. L.; TOWNSEND, R. M.; KARAIVANOV, A. Distinguishing Limited Liability from Moral Hazard in a Model of Entrepreneurship. **Journal of Political Economy**, v. 114, n. 1, p. 100–144, 2006.

PHELAN, C.; TOWNSEND, R. M. Computing Multi-Period, Information-Constrained Optima. **The Review of Economic Studies**, v. 58, n. 5, p. 853–881, 1991.

PRAT, J.; JOVANOVIC, B. Dynamic contracts when agent's quality is unknown. n. January, 2012. Disponível em: ⟨http://federation.ens.fr/ydepot/semin/texte1112/PRA2012DYN.pdf⟩.

PRESCOTT, E.; TOWNSEND, R. Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard. **Econometrica**, v. 52, n. 1, p. 21–46, 1984. Disponível em: ⟨http://www.jstor.org/stable/10.2307/1911459⟩.

PRESCOTT, E. S. A Primer on Moral-Hazard Models. **Federal Reserve Bank of Richmond Economic Quarterly**, v. 85, n. 1997, 1999.

ROGERSON, W. P. . Repeated Moral Hazard. **Econometrica**, v. 53, n. 1, p. 69–76, 1985.

SPEAR, S.; SRIVASTAVA, S. On repeated moral hazard with discounting. **The Review of Economic Studies**, v. 54, n. 4, p. 599–617, 1987. Disponível em: ⟨http://restud.oxfordjournals.org/content/54/4/599.short⟩.

# Figures

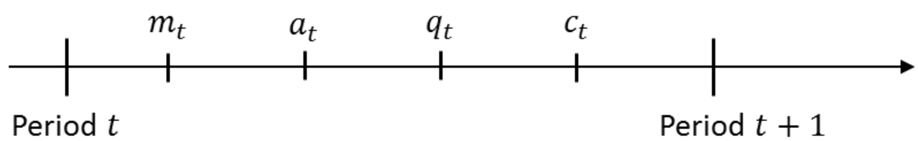


Figure 1 – Timing of events in period t.

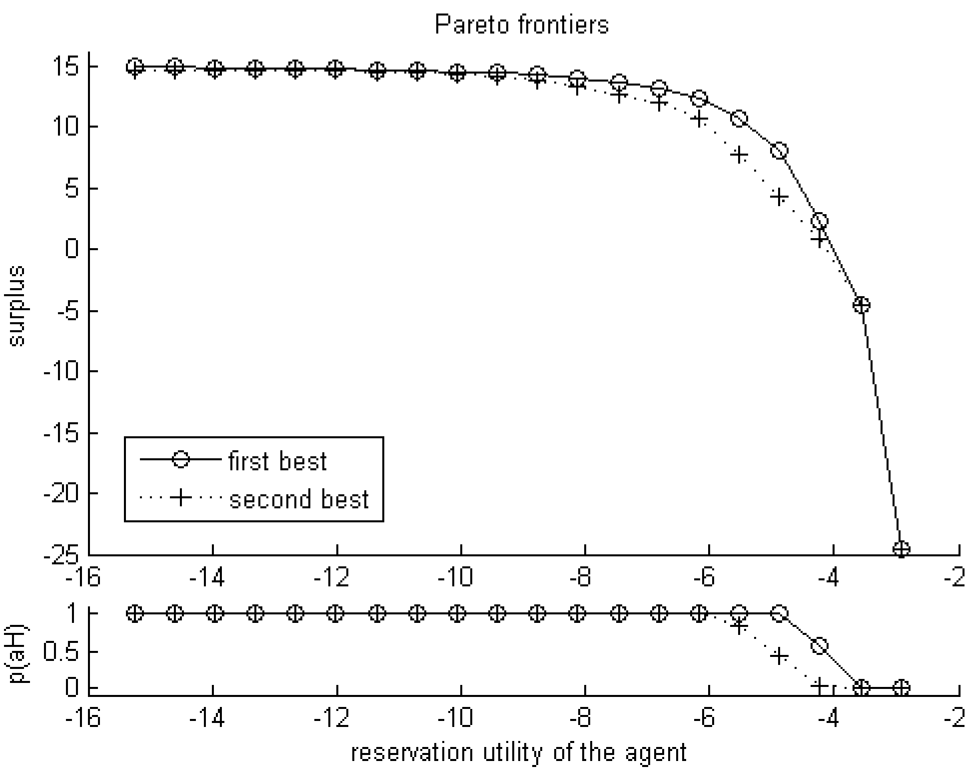


Figure 2 - Comparison of Pareto frontiers (top panel) and probability of recommending high effort (bottom panel) between the first and second best.

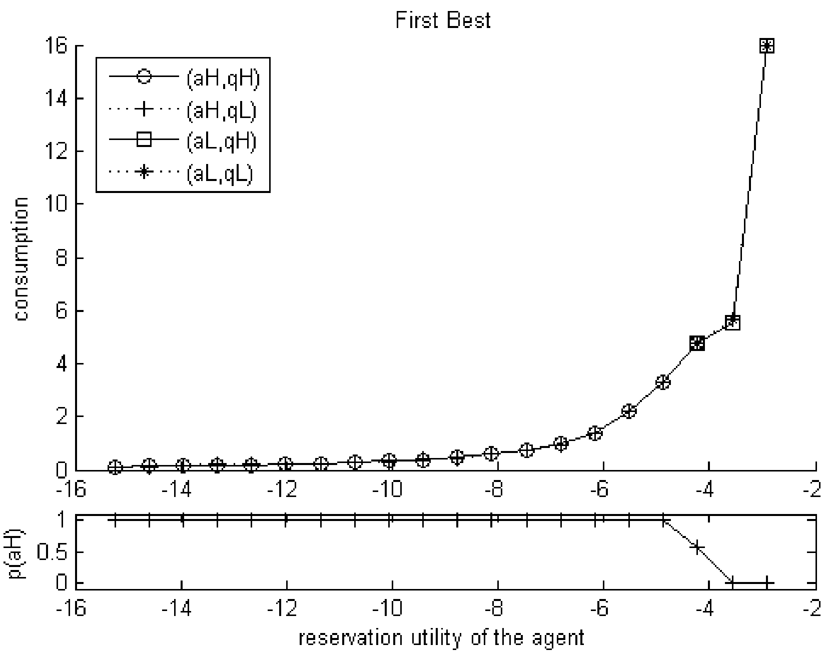


Figure 3 - Consumption plans (c1) under different (a1, q1) pairs - First best (top panel). For reference, the bottom panel shows the probability of recommending high effort in the first period.

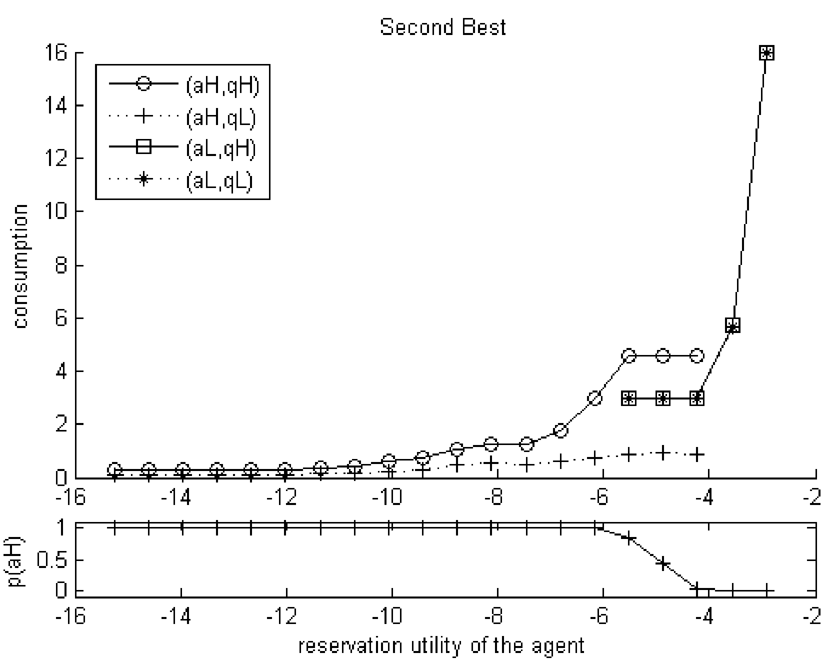


Figure 4 - Consumption plans (c1) under different (a1, q1) pairs - Second best (top panel). For reference, the bottom panel shows the probability of recommending high effort in the first period.

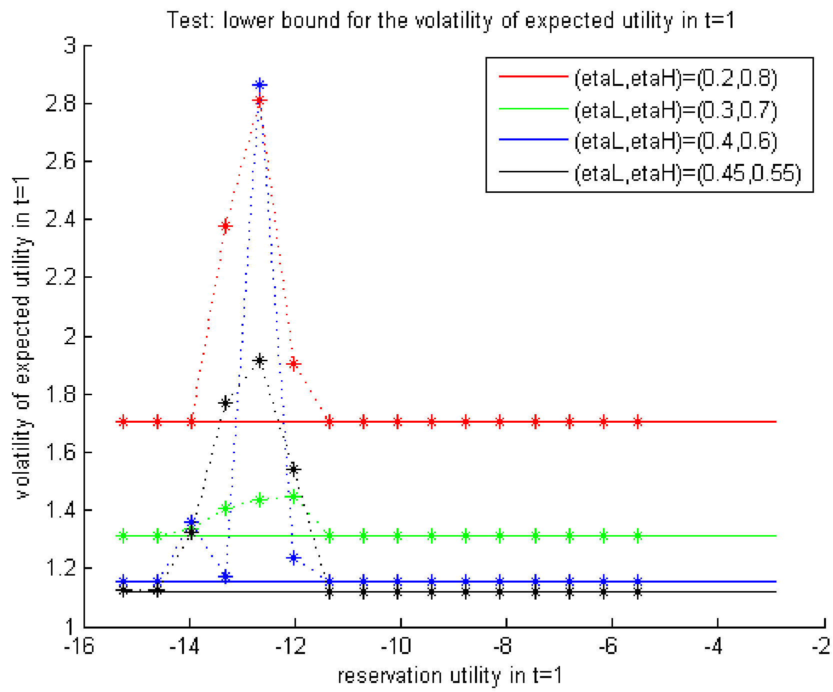


Figure 5 - Testing the lower bound on the volatility of the agent's utility. The figure shows the analytical lower bound (solid line) and the computed results (stars connected by dotted lines) for the volatility of the agent's utility.

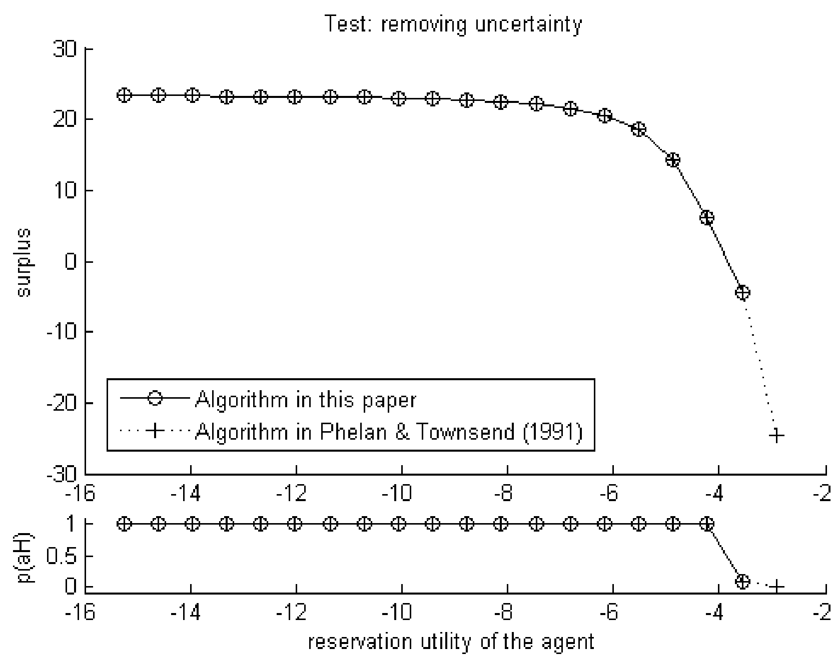


Figure 6 - Comparing surplus (top panel) and effort (bottom panel) in t=1 without uncertainty. Algorithm in this paper (solid line) vs. algorithm in Phelan and Townsend 1991 (dotted line).

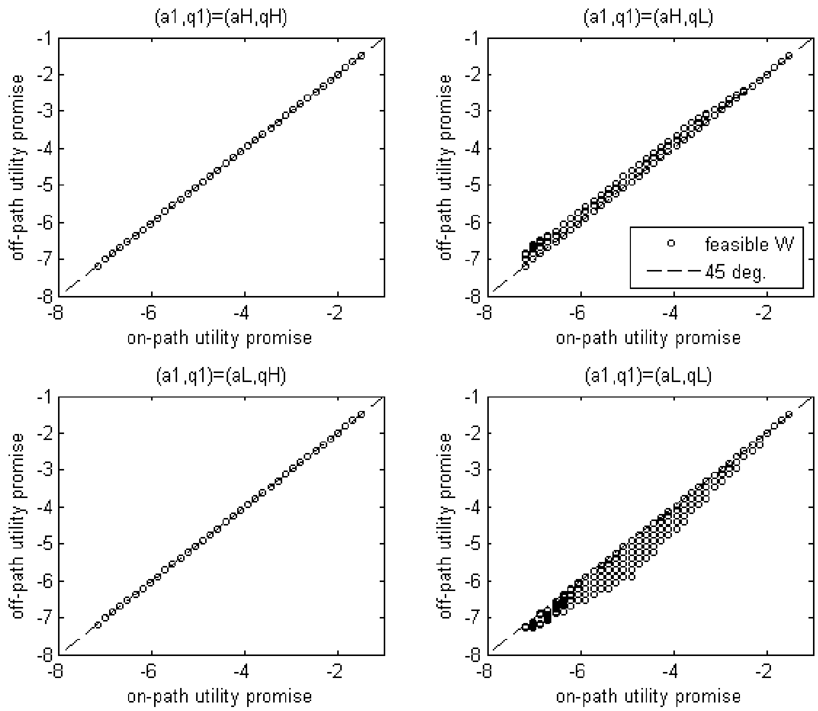


Figure 7 - Spaces of feasible vectors of utility promises in the second period (case with learning). Each panel corresponds to a different first period history s1.

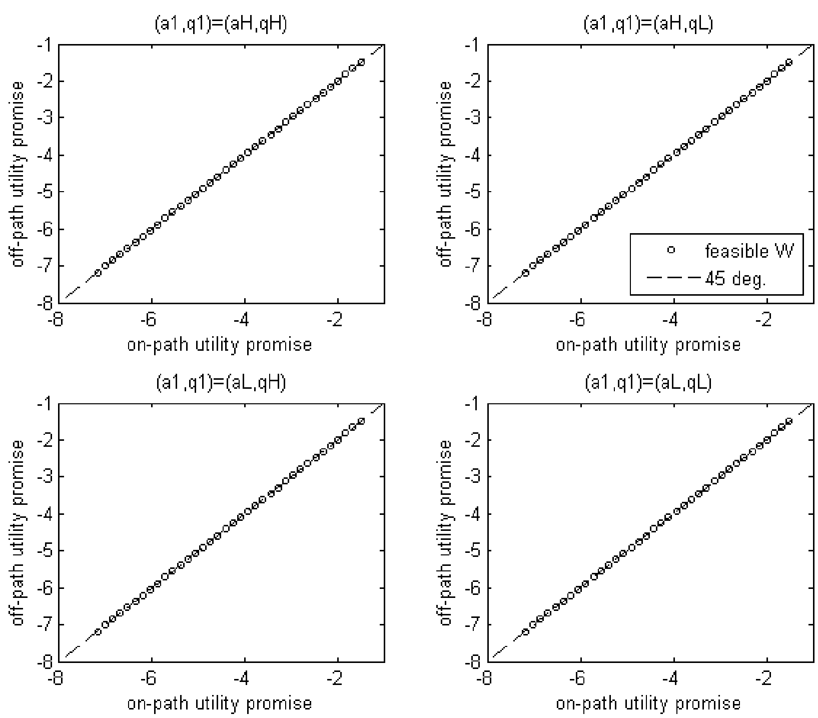


Figure 8 - Spaces of feasible vectors of utility promises in the second period (case without learning). Each panel corresponds to a different first period history s1.

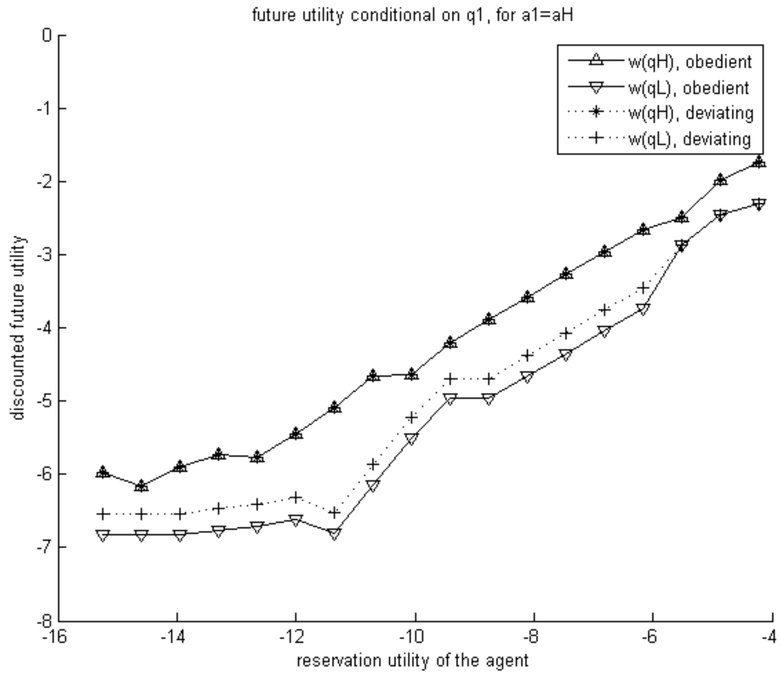


Figure 9 - Future utility conditional on output for obedient and deviating agents, when a1 = aH (case with learning).

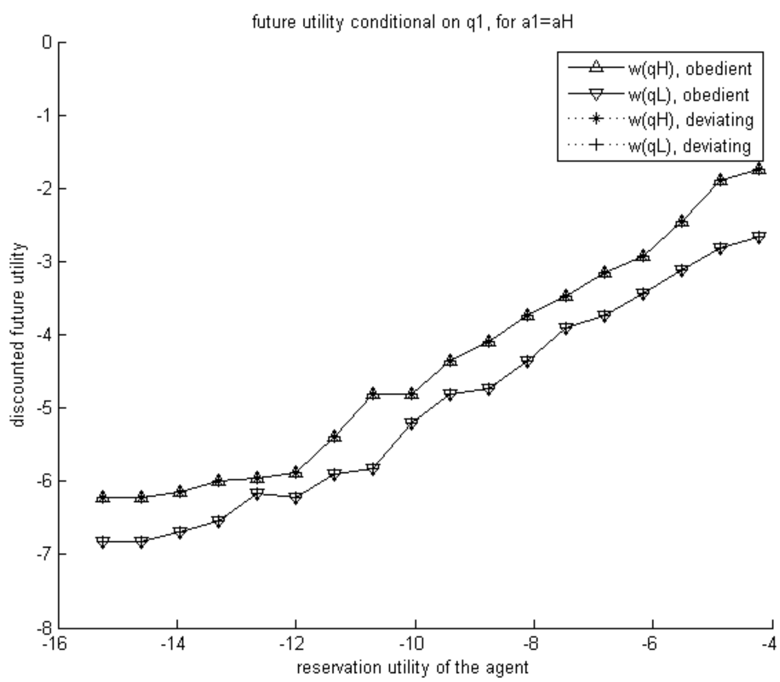


Figure 10 - Future utility conditional on output for obedient and deviating agents, for a1 = aH (case without learning).

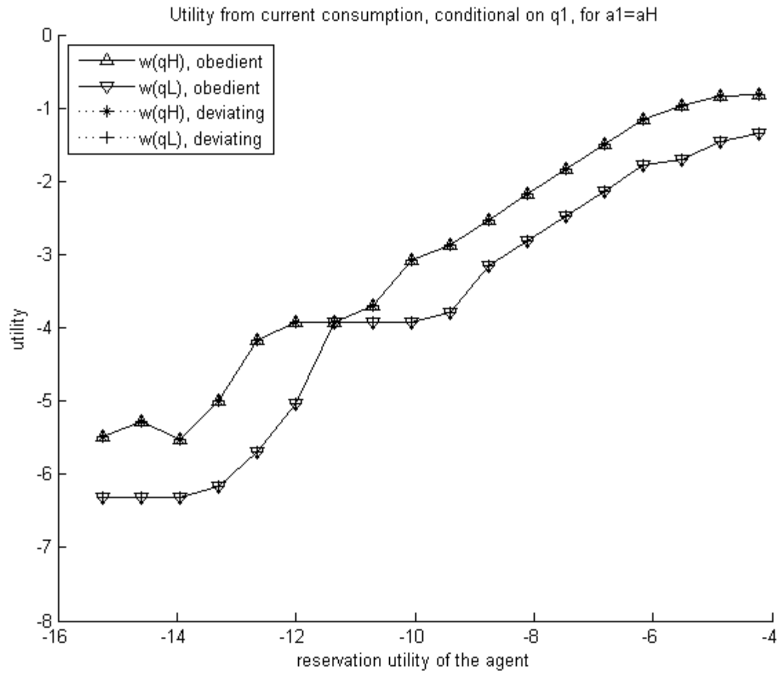


Figure 11 - Utility from current consumption conditional on output for obedient and deviating agents, for a1 = aH (case with learning).

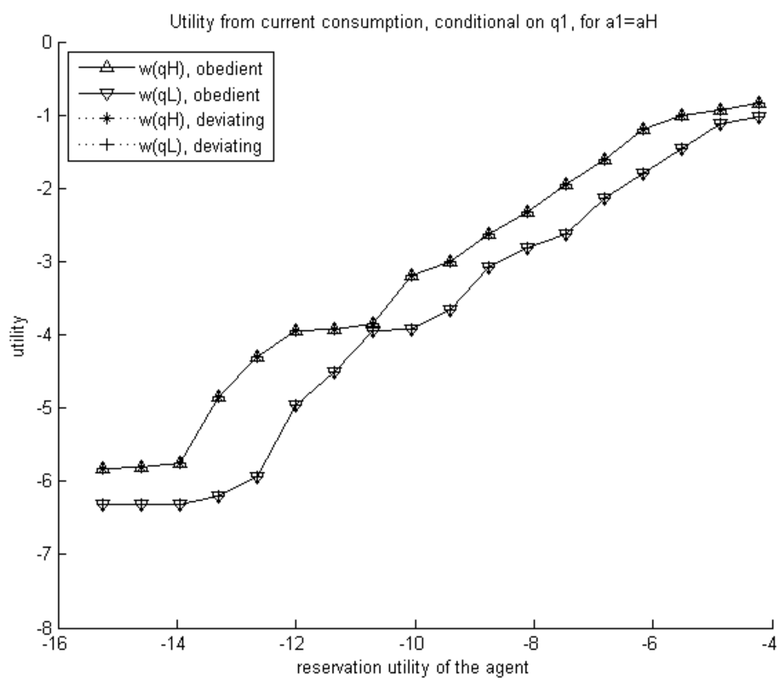


Figure 12 - Utility from current consumption conditional on output for obedient and deviating agents, for a1 = aH (case with no learning).

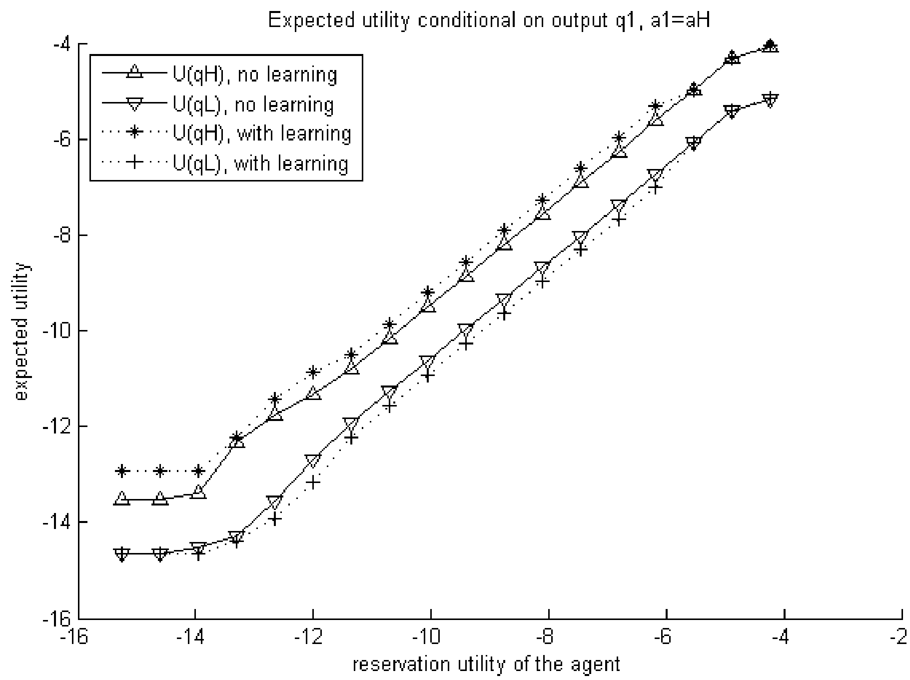


Figure 13 - Comparing the volatility of the agent's expected utility conditional on output in t=1, for a1=aH.

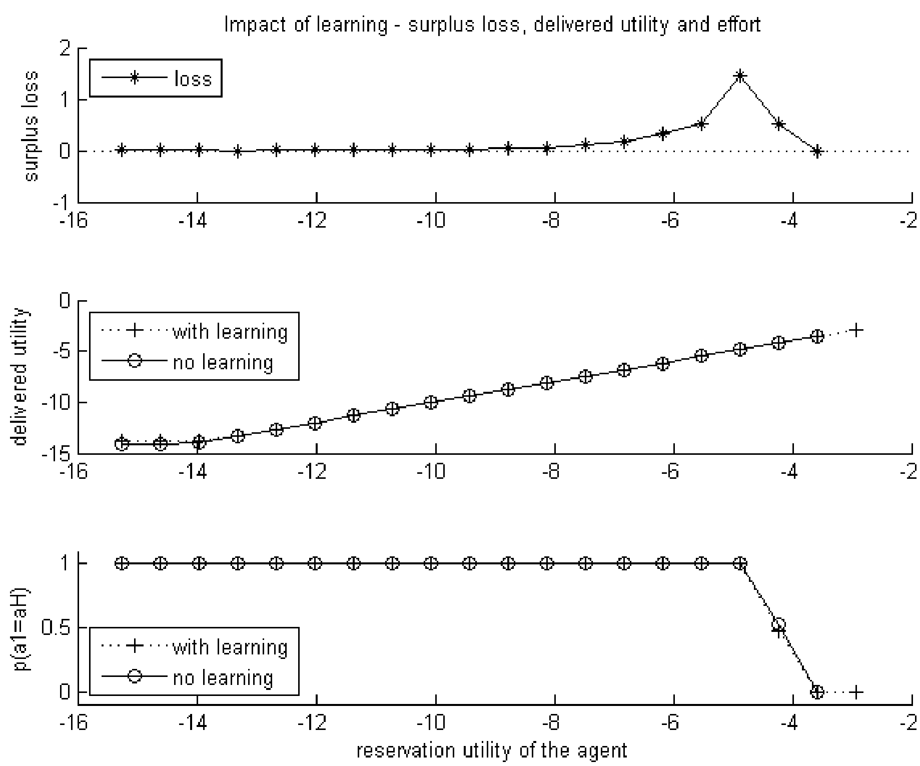


Figure 14 - Impact of learning on surplus (top), ex-ante utility (middle) and recommended effort (bottom).

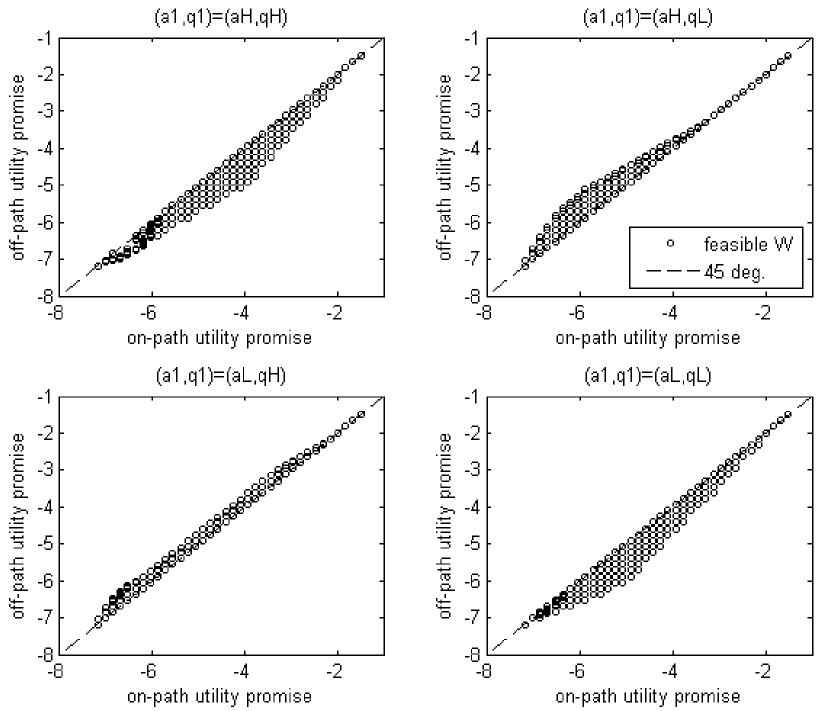


Figure 15 - Spaces of feasible vectors of utility promises in the second period. Each panel corresponds to a different first period history s1.

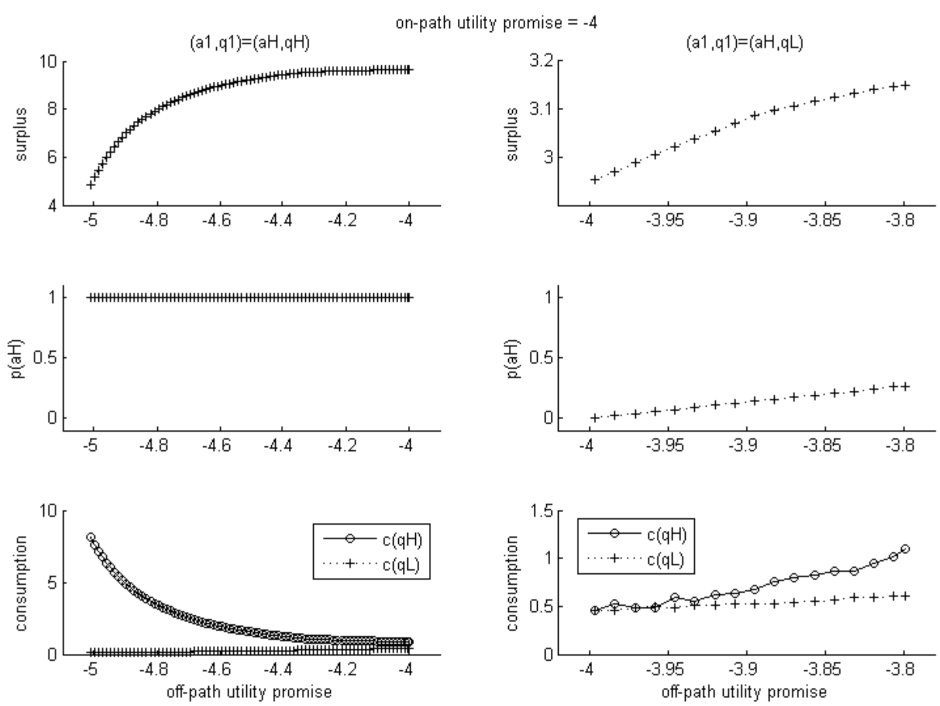


Figure 16 - Looking in detail at the contract in two cross sections of the spaces of feasible utility promises. Panels on the left: (a1, q1)=(aH, qH); panels on the right: (a1, q1)=(aH, qL).

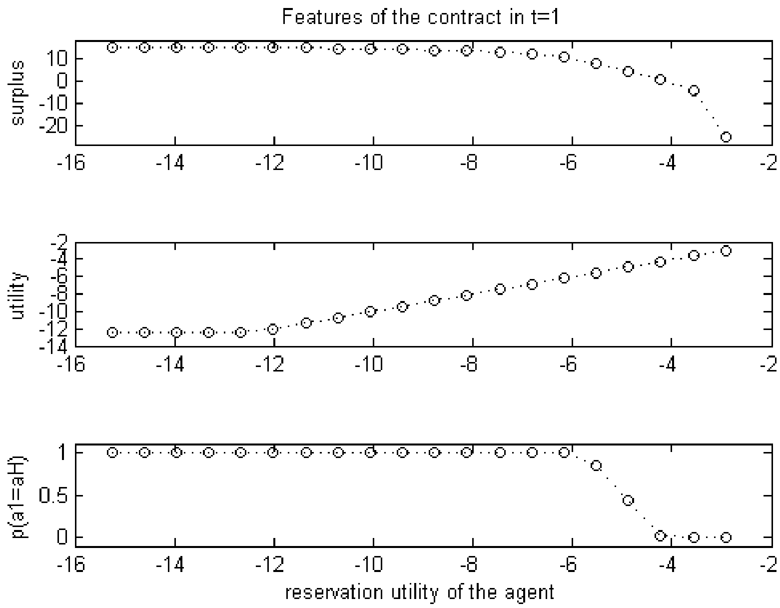


Figure 17 - Computed surplus, utility and probability of high effort in t=1.

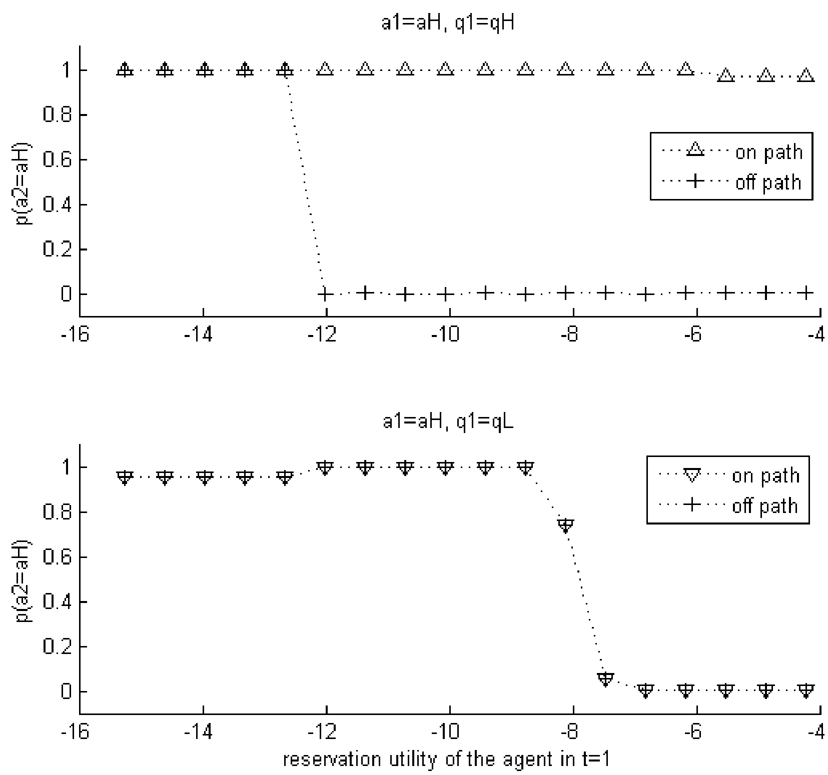


Figure 18 – Average effort in the second period conditional on high effort being recommended in t=1. Top panel: high output in the first period. Bottom panel: low output in the first period.

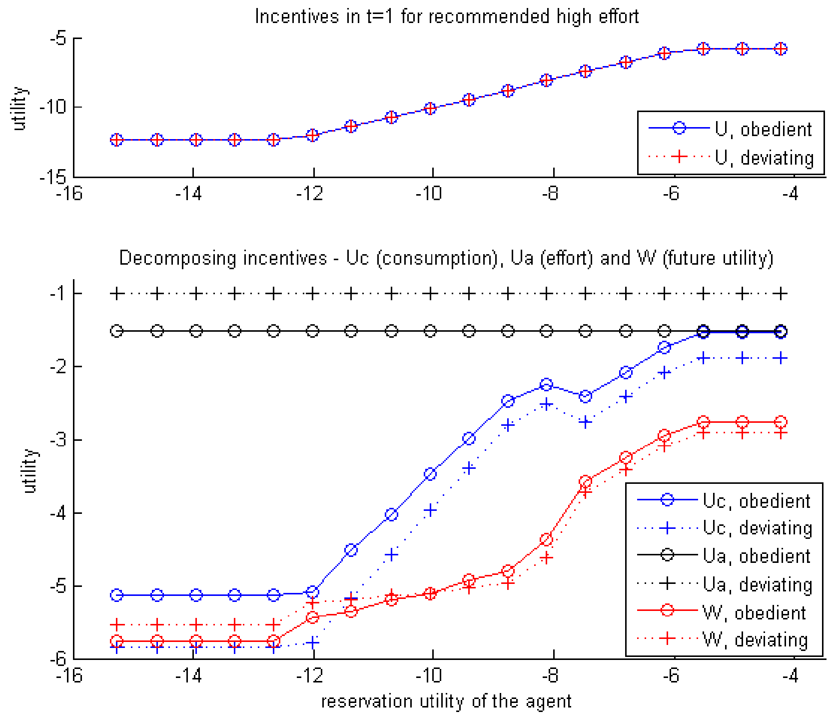


Figure 19 – Incentive contraint in t=1 (top) and decomposition of the incentive constraint in t=1 (bottom).

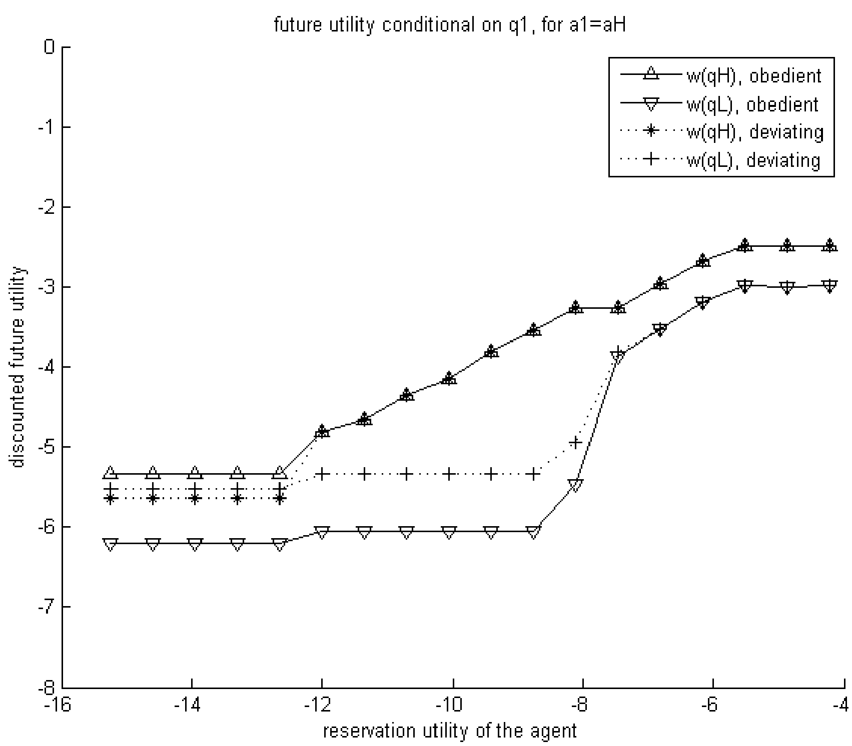


Figure 20 - Future utility promises conditional on q1, when high effort is recommended in the first period.

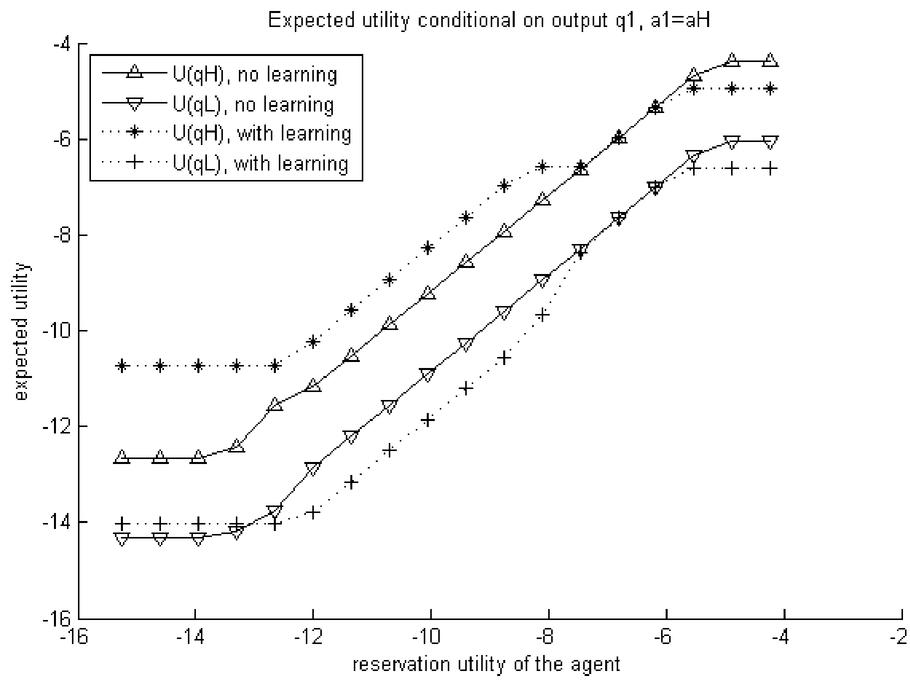


Figure 21 – Comparing the volatility of the agent's expected utility in t=1, cases with and without learning.

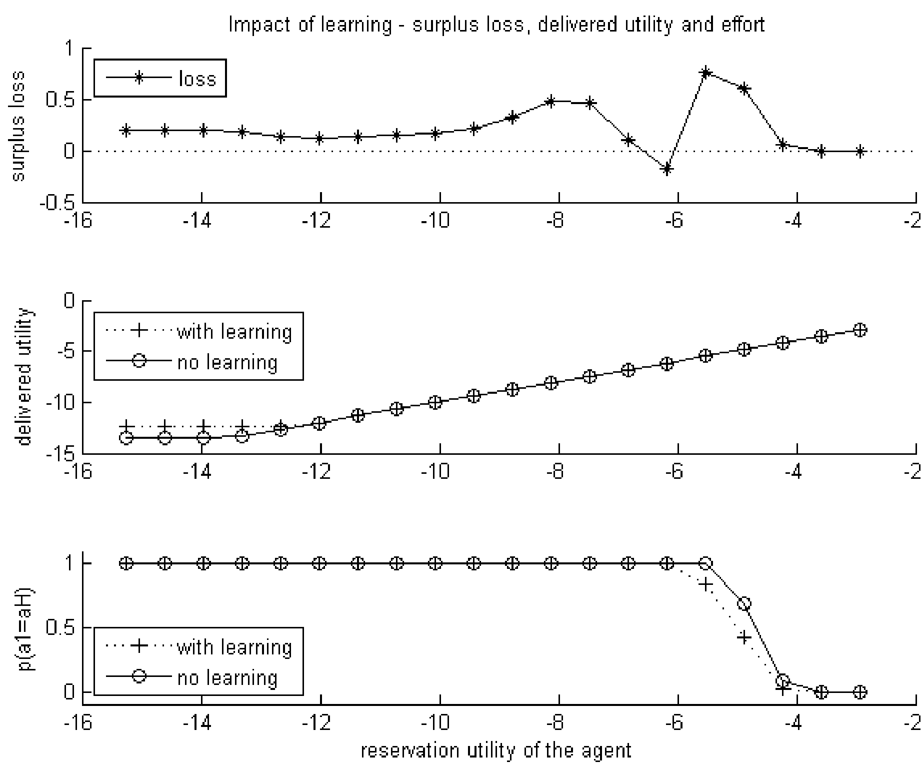


Figure 22 - Impact of learning on surplus (top), ex-ante utility (middle) and recommended effort (bottom).