Economic Quarterly—Volume 96, Number 4—Fourth Quarter 2010—Pages 339–372

Hidden Effort, Learning by Doing, and Wage Dynamics

Arantxa Jarque

Many occupations are subject to learning by doing: Effort at the workplace early in the career of a worker results in higher productivity later on.[1] In such occupations, if effort at work is unobservable, a moral hazard problem arises as well. The combination of these two characteristics of effort implies that employers need to provide incentives for the employee to work hard, possibly in the form of pay-for-performance,[2] while taking into account at the same time the optimal path of human capital accumulation over the duration of the contract.

The recent crisis had a big impact on the labor market with high job-destruction rates. If firm-specific human capital accumulation is important, the effect of these separations on welfare may come from several channels. A direct channel is through the loss of human capital prompted by the exogenous separation, as well as the loss in welfare from the decrease in wealth because of unemployment spells of workers. A less direct channel, but potentially an important one, is the change in the cost of providing incentives when the (exogenous to the incentive provision) separation rate increases. However, we are far from being able to understand and measure the importance of this

I would like to thank Huberto Ennis, Juan Carlos Hatchondo, Tim Hursey, and Pierre Sarte for helpful comments, as well as Nadezhda Malysheva for great research assistance. Andreas Hornstein provided many editorial suggestions that helped shape the final version of this article. All remaining errors are mine. The views presented in this article do not necessarily represent those of the Federal Reserve Bank of Richmond or the Federal Reserve System. E-mail: arantxa.jarque@rich.frb.org.

[1] See Arrow (1962), Lucas (1988), and Heckman, Lochner, and Taber (1998) for a complete discussion of this issue, as well as alternative specifications of learning by doing.

[2] Lemieux, MacLeod, and Parent (2009) report that, for a Panel Study of Income Dynamics sample of male household heads aged 18–65 working in private sector wage and salary jobs, the incidence of pay-for-performance jobs was about 38 percent in the late 1970s and increased to about 45 percent in the 1990s. They define pay-for-performance jobs as employment relationships in which part of the worker's total compensation includes a variable pay component (bonus, commission, piece rate). Any worker who reports overtime pay is considered to be in a non-pay-for-performance job. See also MacLeod and Parent (1999).

cost, since little is known so far about the structure of incentive provision in the presence of learning by doing.[3] This article constitutes a modest first step in this direction: Abstracting from separations and in a partial equilibrium setting, this article studies the time allocation of incentives and human capital accumulation in the optimal contract. This simplified analysis should be a helpful benchmark in future studies of the fully fledged model with separations and general equilibrium.

We modify the standard repeated moral hazard (RMH) framework from Rogerson (1985a) to include learning by doing. In the standard framework, a risk-neutral employer, the principal, designs a contract to provide incentives for a risk-averse employee, the agent, to exert effort in running the technology of the firm. Both the principal and the agent commit to a long-term contract. The agent's effort is private information and it affects the results of the firm stochastically: The probability distribution over the results of the firm (the agent's "productivity") in a given period is determined by the effort choice of the agent *in that same period only*. We introduce the following modification to this standard framework: We specify learning by doing by assuming that the probability distribution over the results of the firm in each period is determined by *the sum of past undepreciated efforts* of the agent, as opposed to his current effort only. In other words, the agent's productivity is determined by his "accumulated human capital." More human capital implies higher expected output, although all possible output levels may realize under any level of human capital. In this specification, the agent determines his human capital deterministically by choosing effort each period. Lower depreciation of past effort is interpreted as "more persistence" of effort.

We present a model of two periods. The first period represents the junior years, when the worker has just been hired and has little experience. The second period represents the mature worker years, when human capital has been potentially accumulated and there are no more years ahead in which to exploit the productivity of the worker. A contract contingent on the observed performance of the agent is designed by the principal to implement the path of human capital accumulation that maximizes the principal's expected profit (expected output minus expected payments to the agent).

In our analysis, we find the following two main implications of the presence of learning by doing. First, the principal does not find it optimal to require a high level of human capital in the last period of the contract, since there is not much time left to exploit the productivity of the worker. Hence, the more experienced workers are not the most productive ones, since they optimally are asked to let their human capital depreciate. This implies that workers exert

[3] The only articles dealing with effort persistence in a repeated moral hazard problem are, to our knowledge, Fernandes and Phelan (2000), Mukoyama and Şahin (2005), Kwon (2006), and Jarque (2010).

the most effort in their junior years, and the least in their pre-retirement years. In a comparison with the standard RMH problem, we find that the frontloading of effort, as well as the low requirement at the end of the worker's career, differ markedly from the optimal path of effort in a context without learning by doing. Second, and in spite of this difference in effort requirements over the contract length, we find that learning by doing does not imply a change in the properties of consumption paths; hence, the properties of consumption paths found by previous studies, such as Phelan (1994), remain true in this context (see also Ales and Maziero [2009]).

It is worth noting that in our analysis we assume perfect commitment to the contract both from the employer and the employee, and we do not allow for separations to be part of the contract. This means we need to abstract from the usual career concerns that have been explored in the literature (see Gibbons and Murphy [1992]). The implications of the hidden human capital accumulation that we model here should be viewed as complementary to the implications of career concerns.

As pointed out above, the problem studied here differs from the standard RMH in that the contingent contract needs to take into account the persistent effects of effort on productivity. On the technical side, this highly complicates solving for the optimal contract. The fact that both past and current effort choices are not observable means that, at the start of every period, the principal does not know the preferences of the agent over continuation contracts (that is, the principal does not know the true productivity of the agent for a given choice of effort today). Jarque (2010) deals with this difficulty and presents a class of problems with persistence for which a simple solution can be found. The article studies a general framework in which past effort choices affect current output, as opposed to other forms of persistence that one may consider, such as through output autocorrelation (see, for example, Kapička [2008]). The learning-by-doing problem that we are interested in, hence, constitutes a fitting application of the results in Jarque (2010). We adapt the assumptions in Jarque (2010) to a finite horizon and we show how this specification of learning by doing greatly simplifies the analysis of the optimal contract.

In Section 1 we introduce the common assumptions throughout the article. Section 2 presents, as a benchmark, the case in which the principal can directly observe the level of effort chosen by the agent every period, and hence can control his human capital at all times. For reference, we also discuss the case in which the effort of the agent does not have a persistent effect in time. The analytical properties of the problem are discussed in both cases. Then we analyze the main case of interest of this article, in which effort is unobservable and contracts that specify payments contingent on the observable performance of the agent are needed to implement the desired sequence of human capital accumulation. In Section 3, we discuss the case without persistence—a standard two-period repeated moral hazard problem. In Section 4 we discuss the technical difficulties of allowing for effort persistence in problems of repeated moral hazard, and the solutions provided in the literature. Section 5 presents the framework of hidden human capital accumulation, a particular case of effort persistence. As the main result, we provide conditions under which the problem with hidden human capital can be analyzed by studying a related auxiliary problem that is formally a standard repeated moral hazard problem. Hence, the discussion of the properties of the standard case in Section 3 becomes useful when deriving the properties of the case with persistence. The numerical solution to an example is presented in Section 6, together with a comparison to the standard RMH without learning by doing, and a discussion of the main lessons about the effects of hidden human capital accumulation on wage dynamics. Section 7 concludes.

1. DESCRIPTION OF THE ENVIRONMENT

The results in this article apply to contracts of finite length T; however, in order to keep the exposition and the notation as simple as possible, we discuss here the case of a two-period contract, $T = 2$. We assume that both parties commit to staying in the contract for the two periods. For tractability, we assume that the principal has perfect control over the savings of the agent. They both discount the future at a rate β. We assume that the principal is risk neutral and the agent is risk averse, with additively separable utility that is linear in effort.

Assumption 1 *The agent's utility is given by $U(c_t, e_t) = u(c_t) - ve_t$, where u is twice continuously differentiable and strictly concave and c_t and e_t denote consumption and effort at time t, respectively.*

There is a finite set of possible outcomes in each period, $Y = \{y_L, y_H\}$. Histories of outcomes are assumed to be observable to both the principal and the agent. We assume both consumption and effort lie in a compact set: $c_t \in [0, y_t]$ and $e_t \in E = \left[\underline{e}, \overline{e}\right]$ for all t.

We model the hidden accumulation of human capital by assuming that the effect of effort is "persistent" over time, in a learning-by-doing fashion. That is, we depart from the standard RMH framework, which assumes that the probability distribution over possible outcomes realizations at t depends only on e_t. In our human capital accumulation framework, the probability distribution at t depends on *all* past efforts up to time t. Assumption 2 states this formally for the two-period problem.

Assumption 2 *The agent affects the probability distribution over outcomes according to the following function:*

$$\Pr(y_t = y_H | s_t) \equiv \pi(s_t),$$

where

$$s_1 = e_1, \quad (1)$$
$$s_2 = \rho s_1 + e_2, \quad (2)$$

and $\pi(s)$ *is continuous, differentiable, concave, and* $\rho \in (0, 1)$.

In the human capital accumulation language, we could equivalently write the law of motion for human capital as

$$s_1 = e_1$$
$$s_2 = (1 - \delta) s_1 + e_2,$$

where $\delta = 1 - \rho$ would represent the depreciation rate. Then,

$$f(s_t) = \begin{cases} y_H & \text{with probability } \pi(s_t) \\ y_L & \text{with probability } 1 - \pi(s_t) \end{cases}$$

could be interpreted as the production function or technology of the firm.

In the rest of the article, we loosely refer to Assumption 2 as effort being "persistent," we refer to s_t as the accumulated human capital at time t, and we refer to ρ as the persistence rate.

The strategy of the principal consists of a sequence of consumption transfers to the agent contingent on the history of outcome realizations, $\mathbf{c} = \{c_i, c_{ij}\}_{i,j=L,H}$, to which the principal commits when offering the contract at time 0. The agent's strategy is a sequence of period best-response effort choices that maximize his expected utility from t on, given the past history of output: $\mathbf{e} = \{e_1, e_{2i}\}_{i=L,H}$. At the beginning of each period, the agent chooses the level of current effort, e_t. Then output y_t is realized according to the distribution determined by all effort choices up to time t. Finally, the corresponding amount of consumption is given to the agent.

A contract is a pair of contingent sequences $\mathbf{c}$ and $\mathbf{e}$. For the analysis in the rest of the article, it will be useful to follow Grossman and Hart (1983) in using utility levels $u_i = u(c_i)$ and $u_{ij} = u(c_{ij})$ as choice variables.[4] To denote the domain for this new choice variable, we need to introduce the following set notation:

$$\mathcal{U}_i = \{u | u = u(c_i) \text{ for some } c_i \in [0, y_i], i = L, H\}$$
$$\mathcal{U}_{ij} = \{u | u = u(c_{ij}) \text{ for some } c_{ij} \in [0, y_j]\, i, j = L, H\}.$$

[4] If the reader is knowledgeable about contract theory, he or she may notice that this is not a simple change of notation. In fact, when computing the solution to numerical examples (see Section 6), we will follow the two-step procedure proposed in Grossman and Hart (1983). This procedure consists of splitting the expected profit-maximization problem of the principal in two steps: (1) cost minimization of implementing a given effort level (on a grid of efforts), and (2) choosing the effort on the grid that implies the highest expected profit for the principal. Using utility as the choice variable, it is easy to show that under the assumptions of this article there will exist a unique minimum in the cost minimization problem.

The contingent sequence of utility is then denoted $\mathbf{u} = \{u_i, u_{ij}\}_{i,j=L,H}$, and we assume that $u_i \in \mathcal{U}_i$ and $u_{ij} \in \mathcal{U}_{ij}$.

In order to keep the expressions in the article as simple as possible, and abusing notation slightly, we also introduce some notation shortcuts. We denote $c_i = u^{-1}(u_i)$ for all i. We also write $\Pr(y_t = y_H|s_t)$ as $\pi_H(s_t)$ and $\Pr(y_t = y_L|s_t)$ as $\pi_L(s_t)$.

The expected profit of the principal, denoted by $V(\mathbf{u}, \mathbf{e})$, depends on the contract as follows:

$$V(\mathbf{u}, \mathbf{e}) \equiv \sum_{i=L,H} \left\{ \pi_i(s_1) \left[y_i - c_i + \beta \sum_{j=L,H} \pi_j(s_{2i}) \left(y_j - c_{ij} \right) \right] \right\},$$

where s_t changes with e_t as detailed in (1). In the same way, we can write the agent's expected utility of accepting to participate in the contract as

$$W_0(\mathbf{u}, \mathbf{e}) = \sum_{i=L,H} \pi_i(s_1) \left[u_i + \beta \sum_{j=L,H} \pi_i(s_{2i}) u_{ij} - ve_{2i} \right] - ve_1. \quad (3)$$

Within this environment we are now ready to set up the problem of finding the optimal contract that will provide the right incentives for human capital accumulation at the least expected cost. Before analyzing the hidden human capital accumulation case, however, we go through a series of related and simpler cases that will serve in clarifying the main case of interest.

2. OBSERVABLE EFFORT

The case of observable effort is often referred to in the literature as first-best (FB) since it represents the maximum joint utility achievable in the contractual relationship between the principal and the agent. This is because, if effort is observable, the principal can directly control the choice of effort of the agent and, hence, there is no need for incentives. This implies that there is no need to impose risk on the agent, which results in lower expected transfers from the principal to the agent. Although we are interested in the case of unobservable effort, it is useful to also analyze this simpler benchmark to learn about the differences between the problem with effort persistence (human capital accumulation) and the standard RMH problem (in which human capital fully depreciates every period).

We will refer to the problem of the principal when effort is observable as problem FB:

$$\max_{(\mathbf{u}, \mathbf{e})} V(\mathbf{u}, \mathbf{e})$$
$$s.to$$

$$\mathbf{e} \in \left[\underline{e}, \bar{e}\right]^3 \quad \text{(ED)}$$
$$u_i \in \mathcal{U}_i,\ u_{ij} \in \mathcal{U}_{ij}\ \ \forall i\ \forall i, j \quad \text{(CD)}$$

$$w_0 \leq W_0\left(\mathbf{u}, \mathbf{e}\right). \quad \text{(PC)}$$

The solution to problem FB is a contract that consists of a pair of contingent sequences of utility and effort that maximize the expected profit of the principal subject to the participation constraint (PC)—which assures that the agent expects as much utility from accepting the contract than staying out—and the domain constraints for consumption (CD) and effort (ED). Characterizing the solution to this problem when considering all the possible combinations of (ED) and (CD) binding constraints is very lengthy and tedious. In the interest of space, we choose to discuss here only the case in which neither of the constraints in (CD) or (ED) bind.

What are the properties of consumption and effort in the optimal contract? We learn them from looking at the first-order conditions of the problem. Let $\lambda \geq 0$ be the multiplier of the (PC).[5] We have:

$$\begin{aligned}
(u_i) &: \quad \frac{1}{u'\left(c_i\right)} = \lambda, \quad \text{for } i = L, H \\
\left(u_{ij}\right) &: \quad \frac{1}{u'\left(c_{ij}\right)} = \lambda, \quad \text{for } i, j = L, H \\
(e_1) &: \quad \left[\pi'\left(s_1\right) + \beta\rho\pi'\left(s_{2i}\right)\right]\left(y_H - y_L\right) = v\lambda \\
\left(e_{2_i}\right) &: \quad \pi'\left(s_{2i}\right)\left(y_H - y_L\right) = v\lambda, \quad \text{for } i = L, H. \qquad (4)
\end{aligned}$$

We analyze in turn the case with and without persistence.

Full Depreciation

First we analyze the observable effort version of a standard two-period RMH problem (see, for example, Rogerson [1985a]). This case is nested in the common framework presented above, for a value of the persistence parameter $\rho = 0$. In this case, effort does not have a persistent effect on the output distribution, that is, there is no learning by doing. Hence, we can say that the human capital of the agent fully depreciates every period.

Here and throughout the rest of the paper, we use stars to denote the solutions to the problems. When necessary, we index the solutions by two arguments: the first one takes a value P if $\rho > 0$ (persistence) and a value NP if $\rho = 0$ (no persistence). The second one takes a value FB if we are in the case of observable effort and a value SB if we are in the case of

[5] Standard arguments for $\lambda > 0$ hold in this setup with persistence. The basic intuition is that $V^*\left(\mathbf{c}, \mathbf{e}; w_0\right)$ is strictly decreasing in w_0.

unobservable effort. Hence, here we denote the solution to problem FB when $\rho = 0$ as $\mathbf{u}^* (NP, FB)$ and $\mathbf{e}^* (NP, FB)$. Note that, whenever it does not lead to confusion, we do not include these arguments to keep the notation light.

Since the right-hand sides of all the first-order conditions for utility are equal to λ, we conclude that the level of utility, and hence consumption, should be the same independent of the output realizations and the period: $u_i^* = u_{ij}^* = u^*$ for all i, j. The first-order conditions of effort, in turn, imply that effort requirements are independent of output realizations and the period: $e_1^* = e_{2i}^* = e^*$ for all i. It is easy to see that, given these properties of consumption and effort, the (PC) in problem FB simplifies to

$$w_0 = (1 + \beta) \left(u^* - v e^* \right).$$

Hence, we can solve for the level of utility in the solution to the FB problem:

$$u^* \equiv \frac{w_0 + v (1 + \beta) e^*}{1 + \beta}. \tag{5}$$

Let $c^* \equiv u^{-1} (u^*)$. Let $\pi'_j (e_2)$ denote the derivative of $\pi_j (e_2)$. Noting that $\pi'_H (e) = -\pi'_L (e)$, we can combine the first-order conditions for consumption and effort to get

$$u' \left(c^* \right) \pi'_H \left(e^* \right) (y_H - y_L) = v \; \forall t. \tag{6}$$

That is, the optimal effort level is such that the marginal benefit from increased effort (the marginal increase in expected output times the marginal utility of output) equals the marginal utility cost of effort.

The following properties summarize our conclusions about the FB problem with nonpersistent effort:

1A. We have that $c_1^* = c_2^* = c^*$.

2A. We have $e_1^* = e_2^* = e^*$.

The main property of the optimal *consumption* sequence of the FB contract in the standard RMH problem is that the contract insures the agent completely against consumption fluctuations whenever feasible. The intuition for this result is straightforward: Since the agent has concave utility in consumption, this is the cheapest way of providing the agent with his outside utility. The main property of the optimal *effort* sequence of the FB contract in the standard RMH problem is a constant effort requirement over time. The tradeoff between increasing the disutility suffered by the agent and increasing the expected output is exactly the same in each period, and hence the solution is the same each time.

It is worth noting that the solution in the observable-effort case coincides with that of a repeated static problem ("spot" contract) in which neither the agent nor the principal commit to the two-period contract, and the outside

Table 1 Parameters of the Numerical Example

υ	Marginal effort disutility	5.00
β	Discount factor	0.65
y_H	Output realization, high state	30.00
y_L	Output realization, low state	20.00
w_0	Outside utility	6.55

utility of the agent is $\frac{w_0}{2}$ each period. Hence, commitment has no value in the case of observable effort and no persistence.

An example

Throughout this article, we illustrate the properties of each particular case of the environment presented by solving a particular numerical example. This makes it easy to compare across the different cases presented. The common parameters of the example are listed in Table 1.

We also assume $u(c) = 2\sqrt{c}$ and a probability function

$$\pi(s) = \sqrt{s}, \tag{7}$$

as well as $\underline{e} = 0.01$ and $\overline{e} = 0.99$.

We now solve for c^* and e^*. Since we are in the case of full depreciation of human capital, we use $\rho = 0$ and the formulas derived above. For our example, we have that (6) becomes

$$\begin{aligned} \frac{1}{2\sqrt{e^*}}(30 - 20) &= 5\sqrt{c^*} \\ \frac{1}{2\sqrt{e^*}} &= \sqrt{c^*} \\ c^* &= \frac{0.25}{e^*}. \end{aligned}$$

Together with (5) this gives us the solutions listed in Table 2.

Observable Human Capital Accumulation

We now turn to analyzing the case in which the effects of effort are persistent in time, with $\rho > 0$. That is, we analyze the optimal contract in the presence of human capital accumulation, or learning by doing.

We established above that the main property of the optimal consumption sequence of the FB contract in the standard RMH problem is that the contract insures the agent completely against consumption fluctuations. Here we will learn that this property remains true in the case with effort persistence. The main property of the optimal effort sequence of the FB contract in the standard

RMH problem is also a constant effort requirement over time. We will learn that when effort is persistent this property no longer holds: Effort requirements will vary over time even in the observable effort benchmark.

We now proceed to derive these results by formally analyzing the problem of the principal FB for the case of $\rho > 0$. She chooses an optimal contract: a pair of contingent sequences $\mathbf{u}^*(P, FB)$ and $\mathbf{e}^*(P, FB)$ that solve problem FB, i.e., they maximize the expected profit of the principal subject to (PC) and the domain constraints (CD) and (ED). We initially discuss the case in which neither the (CD) nor the (ED) constraint bind. However, the lower (ED) constraint (the non-negativity constraint on effort) may bind, with persistence, in not-so-trivial cases. Because of its relevance, the case of this constraint binding will be discussed in turn.

We can derive the properties of the solution by analyzing the first-order conditions in (4) for the case of $\rho > 0$. The first thing to note is that, as in the case without persistence, neither consumption nor effort are contingent on output realizations. However, effort recommendations will depend on the time period. We can use the (PC) here as well to derive the optimal level of utility:

$$u^* \equiv \frac{w_0 + v\left(e_1^* + \beta e_2^*\right)}{1+\beta}.$$

The optimal level of consumption will be $c^* \equiv u^{-1}(u^*)$. We can substitute the first-order condition for effort e_2 into that for e_1, as well as the expression of λ from the consumption first-order conditions, to get an expression for the tradeoff determining the choice of e_1:

$$u'\left(c^*\right)\pi'_H\left(s_1^*\right)\left(y_H - y_L\right) = v\left(1-\beta\rho\right). \tag{8}$$

Comparing this to the tradeoff determining the choice of e_2,

$$u'\left(c^*\right)\pi'_H\left(s_{2i}^*\right)\left(y_H - y_L\right) = v, \tag{9}$$

we learn that the marginal cost of increasing effort in the first period is different (smaller) than that in the second period. The optimal choice takes into account that any effort e_1 exerted in the first period persists into the second one, i.e., it "saves" the agent the equivalent of the discounted disutility of effort of exerting ρe_1 in the second period. This difference in the effective cost of effort that appears because of persistence implies that the principal sets the effort requirements in a way that implies a higher probability of observing y_H in the first period than in the second. We can see exactly how this difference is determined by using the first-order conditions of effort to get the following relationship:

$$\frac{\pi'_H\left(s_1^*\right)}{1-\beta\rho} = \pi'_H\left(s_2^*\right). \tag{10}$$

This implies $s_1 > s_2$ since $1 - \beta\rho$ is always between 0 and 1. From the accumulation of human capital in (1) we have that

$$\begin{aligned} e_1^* &= s_1^*, \\ e_2^* &= s_2^* - \rho s_1^*, \end{aligned} \tag{11}$$

which implies a higher effort in the first period than in the second, $e_1^* > e_2^*$.

The following properties summarize our conclusions about the case with persistence and observable effort:

1B. We have that $c_1^* = c_2^* = c^*$.

2B. We have that $e_1^* > e_2^*$.

That is, whenever c^* is feasible in both states, the principal provides complete consumption smoothing, both across states and across time. As for effort requirements, the principal decreases the requirement from the first to the second period. We repeat the intuition for this result: In the first period, the effort disutility incurred by the agent is a sort of "investment," since it improves the conditional distribution not only in the current period but also in the following one. At $t = 2$, however, there is no period to follow, so the marginal benefits of effort are not as high, while the marginal cost is the same as in the first period.[6]

An example

We now solve for the optimal contract with persistence and observable effort. For this case with accumulation of human capital, we use $\rho = 0.2$ and the formulas derived above. We list the solution in Table 2. Note that the level of $s_2^*(P, FB)$ in this case is 0.16, smaller than that of the second-period effort in the no-persistence case of the previous section, which was $e_2^*(NP, FB) = 0.17$. Comparing the equations that determine each ([6] for $e_2^*(NP, FB)$ and [9] for $s_2^*(P, FB)$), we can see that $c^*(P, FB) < c^*(NP, FB)$ implies $1/u'(c^*(P, FB)) > 1/u'(c^*(NP, FB))$, and hence $\pi'_H\left(s_2^*(P, FB)\right) > \pi'_H\left(e_2^*(NP, FB)\right)$. Given the concavity of $\pi(\cdot)$, it follows that $s_2^*(P, FB) < e_2^*(NP, FB)$.

The Nonnegativity Constraint on Effort

In light of this solution we can discuss the case of the lower constraint in (ED) binding. As an introduction to why this case is of particular relevance to the

[6] In a $T > 2$ framework with $s_0 = 0$, we would have that $e_1 \geq e_t$ for $t < T$, that $e_t = e_2$ for $t = 2, ..., T-1$, and $e_T \leq e_2$. Again, the intuition is that in all $t < T$, effort improves the conditional distribution not only in the current period, but also in the periods that follow. At $t = 1$, since $s_0 = 0$, effort is higher than in any other period.

Table 2 Solutions for the Numerical Example, FB Problem

FB Solutions	c_1^*	c_2^*	e_1^*	e_2^*	s_1^*	s_2^*
NP	5.82	5.82	0.17	0.17	0.17	0.17
P	4.95	4.95	0.22	0.12	0.22	0.16

problem with persistence, it is useful to consider the effect of changes in the persistence parameter, ρ, on the effort solution just presented. For a value of persistence $\rho = 0$, effort equals accumulated effort trivially, and its level is constant across periods. On the other hand, if we instead substitute a value of persistence $\rho = 1$, $(1 - \beta\rho)$ takes its minimum value in (10) and the solution implies the maximum difference between the level of s_1 and s_2, with s_1 much higher than s_2. However, carefully inspecting (11), we can already see that such high level of persistence cannot be compatible with an interior solution for effort in period 2: The principal would choose $e_2^* = 0$. Since $s_1^* > s_2^*$ for all values of $\rho > 0$, effort e_2^* may not be interior for other high enough values of ρ. In other words, persistence implies that, in many interesting cases, the lower domain constraint on effort (ED) cannot be safely ignored.

Constraint (ED) is represented by the following set of inequalities:

$$s_{2i} \leq \rho s_1 + \overline{e} + \underline{e}, \tag{12}$$

and

$$s_{2i} \geq \rho s_1 + \underline{e}. \tag{13}$$

Constraint (12) may be binding for some parametrizations. However, we choose not to discuss this case explicitly here because it is easy to impose ex ante conditions on the parameters that preclude it from binding; for example, for the specification of the probability in (7), it is easy to see that $s \geq \overline{e}$ is never chosen in the optimal contract. The lower bound on s represented in (13), however, is endogenous, and equation (13) cannot be checked without having the solution s_1^* in hand. Fortunately, in the case of observable effort that we are analyzing here, we are able to include constraint (13) explicitly in the maximization problem FB. This allows us to study how the solution properties differ from those in 1B and 2B discussed above when this constraint binds.

Let $\gamma_i \geq 0$ be the multiplier associated with constraint (13) in the version of problem FB for the case $\rho > 0$. We have that the first-order condition for e_{2i} is modified as follows:

$$(e_{2i}) : \pi'(s_{2i})(y_H - y_L) = v\lambda - \gamma_i, \quad \text{for } i = L, H. \tag{14}$$

Note that, again, the choice for effort in the second period is not contingent on the first-period outcome, so we have $\gamma_L = \gamma_H = \gamma$. Then we can substitute

(14) into the unmodified first-order condition for first-period effort, (e_1), to get a general version of equation (8) that allows for the lower domain constraint of effort to be binding:

$$\pi'(s_1)(y_H - y_L) = v\lambda(1 - \beta\rho) + \beta\rho\gamma. \tag{15}$$

From the Kuhn-Tucker conditions, we know that whenever $\gamma > 0$ we have $e_2^* = 0$ and, hence, $s_2^* = \rho s_1^*$.

An example

In some special cases, we can check ex ante whether $\gamma = 0$ is a feasible solution to the FB problem, and hence we can restrict ourselves to the simpler analysis without domain constraints. In particular, with the specification for the probability function in (7) that we are using for our example, equation (10) becomes

$$\frac{1}{(1 - \beta\rho)\, 2\sqrt{s_1^*}} = \frac{1}{2\sqrt{s_2^*}},$$

or, rewriting,

$$s_2^* = (1 - \beta\rho)^2\, s_1^*. \tag{16}$$

This is the relationship that should hold between the level of s_1^* and s_2^* whenever $\gamma = 0$. Hence, the domain condition $e_2 \geq 0$ is satisfied whenever $s_2^* \geq \rho s_1^*$, or, substituting s_2^* from (16), whenever

$$(1 - \beta\rho)^2 \geq \rho. \tag{17}$$

A closer inspection of condition (17) shows that, for $\beta \leq 0.5$, it is always satisfied. For higher β values, however, the condition is satisfied only for low enough ρ values, i.e., when effort is not "too persistent." In our example, for $\beta = 0.65$, we need to check whether (17) is satisfied: The left-hand side is equal to 0.76, which is clearly greater than the right-hand side, 0.2.

To summarize the findings of our analysis, we have shown that for the numerical example presented here, we can provide ex ante conditions (a functional form for the probability as in equations [7] and [17]) on the parameters of the problem that assure us that the domain constraints on (ED) do not bind. Under such restrictions, the characteristics of the solution to the first-best problem 1B and 2B presented earlier in this section are valid.

In relation to those characteristics, it is worth pointing out that the properties of effort requirements depend strongly on our assumption that the utility of the agent is linear in effort. Linearity implies that there is no tradeoff between the efficient accumulation path of human capital and smoothing effort disutility over time. In other words, the smoothing of effort requirements over the duration of the contract does not increase the overall utility of the agent, as is the case with consumption smoothing; hence, the principal only takes into account the effects that different accumulation paths have on the utility of the agent and his own profit through the changes in expected output over time. In the numerical example in Section 6, we will revisit the solution to the observable-effort case discussed here, and we will see the direct consequence of this: It is optimal to ask the agent to exert effort earlier rather than later in the contract, since effort that is done early improves the distribution over future output, holding constant the level of future effort.

3. UNOBSERVABLE EFFORT WITH FULL DEPRECIATION

When effort is not directly observable, the principal must rely on observed output realizations, which are imperfect signals about the effort level of the agent, in order to implement the desired sequence of human capital. Contrary to the case of observable effort, here consumption in a given period will need to vary with the output realization in order to provide incentives for the worker to choose the recommended level of effort.

Formally, the problem of the principal, which we will refer to as the second-best (SB), is:

$$\max_{(\mathbf{u},\mathbf{e})} V(\mathbf{u},\mathbf{e})$$
$$s.to$$

$$\mathbf{e} \in \left[\underline{e}, \overline{e}\right]^3 \qquad \text{(ED)}$$
$$u_i \in \mathcal{U}_i,\ u_{ij} \in \mathcal{U}_{ij}\ \ \forall i\ \forall i, j \qquad \text{(CD)}$$

$$U \leq W_0(\mathbf{u},\mathbf{e}) \qquad \text{(PC)}$$

$$W_0(\mathbf{u},\mathbf{e}) \geq W_0(\mathbf{u},\widehat{\mathbf{e}})\ \ \forall\widehat{\mathbf{e}} \neq \mathbf{e}. \qquad \text{(IC)}$$

The incentive constraint (IC) ensures that the expected utility that the agent gets from following the principal's recommendation is at least as large as that of any other effort sequence.

In order to illustrate clearly the differences that derive from the presence of effort persistence in this two-period problem, we analyze first the version without persistence ($\rho = 0$), that is, with full depreciation of human capital every period, or no learning by doing. Moreover, because the main result that we will derive when we study the case with $\rho > 0$ is that, in some cases, the properties for consumption in the optimal contract will be the same as those of the optimal contract in a framework without persistence, it is useful to analyze in detail the properties of the solution with observable effort.

Without persistence, the structure of the incentive constraints simplifies considerably. This influences the solution, but also the ways in which the problem can be studied. In particular, the standard RMH problem has a simple recursive formulation that is not available with persistence. In this section we provide an illustration of this difference. Then, we discuss the difficulties of introducing persistence, along with some potential solutions, in Section 4. In Section 5 we discuss our example with human capital accumulation, a particularly simple case with effort persistence for which a solution can easily be found.

A Simplified Incentive Compatibility Constraint

In the case without persistence the structure of the incentive constraints simplifies considerably. In particular, the expected utility of the agent in the second period is independent of the first-period effort choice. Define

$$W_{1i}(\mathbf{u}, \mathbf{e}) = \sum_{j=L,H} \pi_j(s_{2i}) u_{ij} - v e_{2i}, \quad \text{for } i = L, H, \tag{18}$$

as the expected utilities for the second period, contingent on the first-period realization. This expression for the continuation utility simplifies, when $\rho = 0$, to

$$W'_{1i}(\mathbf{u}, e_{2i}) \equiv \sum_{j=L,H} \pi_j(e_{2i}) u_{ij} - v e_{2i}, \quad \text{for } i = L, H. \tag{19}$$

(Note that, to distinguish the notation for continuation utilities here from those of the general case that allows for persistence in (18), we denote them here with a prime and we make explicit the independence of e_1.)

What is the simplification of the incentive constraints that follows from this independence? As it turns out, all the sequences that have the same choice of effort in the second period, regardless of the first-period effort choice, provide the agent with the same expected utility in the second period, conditional on the first-period output realization being the same. In other words, the deviations of the agent in the second period can be evaluated independently of the first-period effort choice, and also independently at each node following the first-period output realization. As a consequence, the number of relevant incentive constraints for the agent is drastically decreased.

To see this formally, denote by $w_{1i} \equiv W'_{1i}(\mathbf{u}, e_{2i})$ the continuation utilities evaluated at the effort requirement of the principal. Then all the incentive constraints that involve deviations *only* in the second period, or that have the same effort choice for the first period, simplify to

$$w_{1i} \geq W'_{1i}(\mathbf{u}, \widehat{e}_{2i}) \quad \forall \widehat{e}_{2i} \neq e_{2i} \quad \text{for } i = L, H. \tag{20}$$

We refer to equation (20) as the "second-period incentive constraints."[7]

[7] For a more concrete illustration, consider the case with discrete effort and $E = [e_L, e_H]$. Then the initial number of IC constraints would be seven, and they would simplify to three: one first-period constraint and two second-period constraints.

Now note that the independence of $W'_{1i}(\mathbf{u}, \widehat{e}_{2i})$ on e_1 also implies the following: Imposing the second-period incentive constraints in (20) serves to assure that all potential deviations $(\widehat{e}_1, \widehat{e}_{2L}, \widehat{e}_{2H})$ that consider effort choices in the second period that are not e_{2H} and e_{2L} are dominated by a strategy $(\widehat{e}_1, e_{2L}, e_{2H})$ that considers the same deviation in period 1 and none in the second period. Formally, what we are saying is that

$$\sum_{i=L,H} \pi_i(\widehat{e}_1)[u_i + \beta w_{1i}] - v\widehat{e}_1 \geq \sum_{i=L,H} \pi_i(\widehat{e}_1)\left[u_i + \beta W'_{1i}(\mathbf{u}, \widehat{e}_{2i})\right] - v\widehat{e}_1$$

trivially simplifies to the second-period incentive constraint in (20). This is useful because it means that when we are evaluating deviations in the first period we forget about potential deviations in the second period as well, and simply substitute w_{1i} into the second period utility:

$$\sum_{i=L,H} \pi_i(e_1)[u_i + \beta w_{1i}] - ve_1 \geq \sum_{i=L,H} \pi_i(\widehat{e}_1)[u_i + \beta w_{1i}] - v\widehat{e}_1. \quad (21)$$

We refer to these constraints as the "first-period incentive constraints."

The independence of second-period expected utility on first-period effort choice not only decreases the number of IC constraints that we need to consider, but also allows the problem of the principal to be analyzed period by period. This is precisely because all future period payoffs can be summarized through the promised utility w_{1i} without specifying the particular consumption transfers or effort recommendations that will deliver w_{1i} in the future. From a practical point of view, it is important to note that the range of values that w_{1i} can take is independent of the agent's action in the first period, and hence can be calculated by simply using the domain restrictions for consumption and second-period effort, together with the second-period IC in (20). This is a very useful feature when we want to compute the solution for a particular numerical example, as we will do in Section 6.

To summarize, the simplifications we just discussed are the reason why the recursive formulation first introduced by Spear and Srivastava (1987) is possible. In a finite two-period problem like the one presented here, this also means that we can solve the problem backward and characterize the properties of the solution. We proceed to do that now.

A Backward Induction Solution to the Optimal Contract

As a first step, we use the fact that incentives in the second period are independent of choices and utilities in the first period. This allows us to split the problem of the agent in the IC into two problems: a first-period problem and a second-period problem. The second-period problem, PIC_2, is

$$\max_{e_{2i} \in [\underline{e}, \overline{e}]} \sum_{j=L,H} \pi_j(e_{2i})\, u_{ij} - ve_{2i},$$

and the first-period problem, PIC_1, is

$$\max_{e_1 \in [\underline{e}, \bar{e}]} \sum_{i=L,H} \pi_i(e_1)(u_i + \beta w_{1i}) - v e_1,$$

where w_i is the expected utility for the second period in equilibrium.

If we want to characterize the optimal contract, first we need to transform these maximization problems into an equality constraint that we can include in the problem of the principal. Following the spirit of the first-order approach (see Rogerson [1985b]), we establish concavity of the maximization problems in PIC_1 and PIC_2. Then we can substitute them by their first-order conditions, which are necessary and sufficient for a maximum. In our two-outcome example, this concavity is fairly straightforward to guarantee. It is easy to see that, for any positive first-period effort recommendation to satisfy the original first-period IC in (21), we need $u_H + \beta w_H > u_L + \beta w_L$. Also, for any second-period positive effort recommendation to satisfy the second-period IC in (20), we need $u_{iH} > u_{iL}$. Since we have assumed that $\pi_H(\cdot)$ is a concave function of effort, concavity of the expected utility of the agent in effort follows.[8] Hence, we can substitute PIC_1 for its first-order condition,

$$(e_1): \sum_{i=L,H} \pi_i'(e_1)(u_i + \beta w_i) - v = 0, \tag{22}$$

and we can substitute PIC_2 by its corresponding first-order condition,

$$(e_{2i}): \sum_{j=L,H} \pi_j'(e_{2i}) u_{ij} - v = 0. \tag{23}$$

Using these in place of the original IC allows us to derive some properties for the optimal contract.

As a second step in characterizing the optimal contract, we appeal to the same logic that we spelled out to show the independence of second-period utility of the agent on his first-period actions, to argue that the same independence holds for the expected profit of the principal. The objective function in problem SB can be written as

$$V(\mathbf{u}, \mathbf{e}) = \sum_{i=L,H} \pi_i(e_1)\left[y_i - c_i + \beta V_{1i}(w_{1i})\right],$$

where

$$V_{1i}(w_{1i}) = \sum_{j=L,H} \pi_j(e_{2i})\left(y_j - c_{ij}\right).$$

Hence, to solve problem SB subject to (PC) and (22) and (23)—assuming the domain constraints are not binding—we can simply split the problem across

---

8 For a higher number of output levels, the conditions on the probability function that would assure concavity have not been determined (see Rogerson [1985b] and Jewitt [1988] for a discussion of these conditions in the context of a static contract).

the two periods and solve it backward using subgame perfection. First, we solve the second-period problem, P_{2i}, for an unspecified value of w_{1i}:

$$\begin{aligned} & \max_{u_{iL},u_{iH},e_{2i}} V_{1i}\left(w_{1i}\right) \\ & s.t.\ (23) \text{ and} \\ w_{1i} \quad = & \sum_{j=L,H} \pi_j\left(e_{2i}\right) u_{ij} - ve_{2i}. \end{aligned}$$

Let μ_i and λ_i be the multipliers of the first and second constraints, respectively. For each $i = L, H$, the first-order conditions with respect to utility are

$$\left(u_{ij}\right) : \frac{1}{u'\left(c_{ij}\right)} = \lambda_i + \mu_i \frac{\pi'_j\left(e_{2i}\right)}{\pi_j\left(e_{2i}\right)}, \quad j = L, H. \tag{24}$$

This condition will be familiar to the reader acquainted with basic contract theory: Since the second-period problem is, in fact, a static moral hazard, we find that this first-order condition links consumption to likelihood ratios in the same way as in a static contract (see Prescott [1999] for a review of this textbook case). The likelihood ratios capture the informational value of each possible output realization. The same static intuition prevails in the case for effort. The first-order conditions are

$$\left(e_{2i}\right) : \sum_{j=L,H} \pi'_j\left(e_{2i}\right)\left(y_j - c_{ij}\right) + \mu_i \sum_{j=L,H} \pi''_j\left(e_{2i}\right) u_{ij} = 0. \tag{25}$$

It is easier to see the intuition when we substitute $\pi'_L(e) = -\pi'_H(e)$ in the expression above and get

$$\left(e_{2i}\right) : \pi'_H\left(e_{2i}\right)\left[y_H - y_L - \left(c_{iH} - c_{iL}\right)\right] + \mu_i \pi''_H\left(e_{2i}\right)\left(u_{iH} - u_{iL}\right) = 0.$$

We see that the principal equates the marginal increase in the expected net profit that comes from a higher probability of y_H with the change in the marginal increase in expected compensation associated with it, given that $u_{iH} > u_{iL}$.

Note, however, that the solution for the second period is contingent on the value of w_{1i} (which plays the role of the period outside utility in a static problem). With the solution to the second-period problem in hand, we can calculate the value to the principal of promising a level of utility of w_{1i} to the agent for the second period. Hence, we know the value of $V_{1i}\left(w_{1i}\right)$ and we can substitute it in the first-period problem, P_1:

$$\begin{aligned} & \max_{\substack{u_L,u_H,e_1,\\ w_{1L},w_{1H}}} \sum_{i=L,H} \pi\left(e_1\right)\left[y_H - c_H + \beta V_{1i}\left(w_{1i}\right)\right] \\ & s.t.\ (22) \text{ and} \\ w_0 \quad \leq & \sum_{i=L,H} \pi_i\left(e_1\right)\left(u_i + \beta w_{1i}\right) - ve_1. \end{aligned}$$

Let μ and λ be the multipliers of the first and second constraints, respectively. The first-order conditions for consumption are

$$(u_i) : \frac{1}{u'(c_i)} = \lambda + \mu \frac{\pi_i'(e_1)}{\pi_i(e_1)}, \quad i = L, H. \tag{26}$$

These mirror the conditions in (24) for the second period: The ranking of consumption is again determined by the likelihood ratios, although the dispersion is potentially different and depends on the multiplier of the first-period incentive constraint, μ. The values of μ and μ_i, as well as λ and λ_i, are difficult to get for generic utility functions. (To see this, note that the first-order conditions give us information about $u'(c)$, while the constraints of the problems P_1 and P_{2i} are written in terms of $u(c)$; this makes for a highly nonlinear system of equations that seldom has an explicit solution.) This is why computing numerically the solution to particular problems is a popular strategy in dynamic contract theory.[9]

Recall that in this first period the principal has an extra choice variable with respect to problem P_{2i}: the contingent levels of expected utility of the agent in the second period, w_{1i}. The importance of the value of w_{1i} relative to that of u_i in the optimal contract is at the heart of dynamic incentives. We can explore the optimal tradeoff between the two variables by looking at the first-order condition for the continuation utility:

$$(w_{1i}) : V_{1i}'(w_{1i}) + \lambda + \mu \frac{\pi_i'(e_1)}{\pi_i(e_1)} = 0, \quad i = L, H. \tag{27}$$

To interpret this condition we need to figure out the derivative of the value function of the principal, $V_{1i}'(w_{1i})$. We do this by using the envelope theorem and the second-period problem P_{2i} that determines $V(w_{1i})$:

$$V_{1i}'(w_{1i}) = -\lambda_i.$$

Substituting this derivative into (27) we get

$$\lambda_i = \lambda + \mu \frac{\pi_i'(e_1)}{\pi_i(e_1)}, \quad i = L, H.$$

Note that this, combined with (26), implies $\lambda_i = \frac{1}{u'(c_i)}$. What does the λ_i multiplier represent in the second period? It is the shadow value of relaxing the "promise keeping" constraint of the principal in the second period. The principal has committed to deliver a level of expected utility of w_{1i}. How costly this is for him depends on the necessary spread of utilities in order to satisfy incentives in the second period. This can be seen formally by multiplying the first-order conditions for u_{ij} in (24) for each j times $\pi_j\left(e_{2j}\right)$, and then

[9] For details on these computations see, for example, Phelan and Townsend (1991) or Wang (1997).

summing the resulting equations for $j = L$ and $j = H$; doing this we get that

$$\frac{\pi\left(e_{2_i}\right)}{u'\left(c_{iH}\right)} + \frac{1-\pi\left(e_{2i}\right)}{u'\left(c_{iL}\right)} = \lambda_i .$$

The shadow value depends on the expected tradeoff between the marginal value to the principal of increasing consumption, -1, and the marginal increase in utility of spending this extra unit of consumption, $u'(c)$. Now we take this condition further: Since we had established that $\lambda_i = \frac{1}{u'(c_i)}$, we get the following relationship of the inverse of the marginal utility of consumption:

$$\frac{\pi\left(e_{2_i}\right)}{u'\left(c_{iH}\right)} + \frac{1-\pi\left(e_{2i}\right)}{u'\left(c_{iL}\right)} = \frac{1}{u'\left(c_i\right)}. \tag{28}$$

This is the so called "Rogerson condition," first derived in Rogerson (1985a). It summarizes how the optimal dynamic contract with commitment allocates incentives over time and histories. We now discuss its implications for the choices of effort and consumption.

Effort and Consumption Choices Over Time

To illustrate the implications of the Rogerson condition, consider, for the sake of comparison, a slightly different model to the one presented here: Everything else equal, assume no commitment to long-term contracts for both the principal and the agent. This is often referred to as "spot contracting." For the purpose of our comparison, set the per period outside utility for the agent to $\frac{w_0}{2}$ in both periods. It is easy to see that the solution to this problem without commitment is the repetition of the one-period optimal contract. This implies that the second-period consumptions would be independent of the first-period realizations, and hence identical to those in the first: $c_H = c_{LH} = c_{HH}$, as well as $c_L = c_{HL} = c_{LL}$. It is immediate that this solution to the spot contract violates (28).

How is the contract with commitment different than the repetition of the static contract? The main difference is that with commitment the contract exhibits memory, i.e., the level of consumption in the second period, contingent on a second-period realization, is different depending on the first-period realization. Why is it optimal for the contract with commitment to be different than the repetition of the static contract? Because it allows incentives to be provided in a more efficient way. The reason becomes clear if we consider how the principal can improve on the repetition of the static contract once he has commitment to a two-period contract. If the agent gets a y_H realization in the first period, his overall expected utility increases if he trades off some of the consumption that the static contract assigns him in the first period with some expected consumption in the second. Because c_H was high to start with, the decrease in his first-period utility from postponing some consumption translates into a bigger increase in expected utility in the second period, where he has positive probability of facing low consumption whenever y_L realizes. This means the principal can, with this deviation from the spot contract solution, keep some of the consumption for himself while leaving constant the expected utility following the high realization node in the first period, i.e., $u_H + \beta w_H$. In the same way, if the agent gets a y_L realization in the first period, he is better off by trading some expected utility in the second period for some consumption in the first, and this again saves resources for the principal. Hence, in the optimal contract, we have that $w_{1H} > w_{1L}$. It is worth noting that these optimal tradeoffs result in a violation of the Euler equation of the agent, which is incompatible with (28).[10]

The last first-order condition of problem P_1 left for analyzing is that of effort in the first period:

$$(e_1) : \sum_{i=L,H} \pi_i' (e_1) (y_i - c_i + \beta V_{1i} (w_{1i})) + \mu \sum_{i=L,H} \pi_i'' (e_1) (u_i + \beta w_{1i}) = 0.$$

This condition captures the same tradeoff discussed after deriving the second-period effort first-order conditions in (25). Of course, the values of the variables and multipliers will typically be different than in the second period, implying a different solution across periods. To gain some important insight in the properties of effort requirements over time, it is again useful to compare the effort solution here to that of the spot contract without commitment. It is easy to see that the repetition of the static contract would imply $e_1 = e_{2H} = e_{2L}$.[11] Here, instead, this is not the case. If we recall that the optimal contract implies $w_{1H} > w_{1L}$, a simple inspection of the second-period problem P_2 tells us that, for the principal, effort incentives are more expensive following a y_H realization than a y_L realization. The continuation utility w_{1i} plays the role of the outside utility in a static contract. It is immediate from the risk aversion of the agent that, for the same spread of utility that would satisfy the IC in (23), a higher level of outside utility translates into more consumption. Hence, the principal will optimally choose $e_{2H} < e_{2L}$. Moreover, in the second period the principal cannot provide incentives for effort as efficiently as in the first period, since the intertemporal tradeoff of consumption that we described above is not available (there are no future periods after $t = 2$). This will typically imply a lower effort requirement in the second period than in the first. We conclude that, in contrast with the first-best property summarized in 2A, effort requirements will fluctuate over time and across histories in the unobservable effort case in order to provide incentives more efficiently.

The solution to this version of our numerical example is presented in Table 3 and Figures 1 and 2. We defer the discussion of this solution example until

---

[10] This follows from Jensen's inequality and the convexity of $1/u'(c)$. For details, see Rogerson (1985a).

[11] Simply set $w_{1H} = w_{1L}$ and note that $\pi_H' (e_1) = -\pi_L' (e_1)$.

Section 6, where we compare the solution to the unobserved effort case both with full depreciation and without.

4. DEALING WITH PERSISTENCE

The simplifications outlined in the previous section, when effort is not persistent, do not hold for the general case of $\rho > 0$. Before we go on to analyze a particular case of human capital accumulation in Section 5 and illustrate the differences, we discuss here the main particularities that persistence of effort introduces in the analysis of the optimal contract.

Two main differences with respect to the standard framework appear when effort is persistent. First, it is no longer the case that a given choice for effort in the second period provides the agent with the same expected utility w_{1i} regardless of his first-period effort choice e_1. It follows that the number of relevant incentive constraints is much higher in the problem with persistence. Second, the problem of the principal cannot, in general, be written in the usual recursive form in which the promised utility w_{1i} summarizes all relevant information about past periods. The relevant summary variable is the original $W_{1i}(\mathbf{u}, \mathbf{e})$, which depends on both the first- and the second-period effort choices. The dependence of $W_{1i}(\mathbf{u}, \mathbf{e})$ on e_1 complicates the calculation of its possible values. In particular, this state variable is not a number (like w_{1i} was) but a function: The principal needs to take into account all possible choices for e_1, including those off the equilibrium path. Finally, the conditions for concavity of the agent's problem in the IC are difficult to establish, even in the two-outcome case presented here.

These issues have so far been addressed in the literature with two main strategies. The first strategy limits the effort choices to a two-point set, and includes explicitly in the problem of the principal the complete list of relevant incentive constraints for all possible combinations of effort choices. The second strategy allows for a continuum of effort choices, but puts restrictions on the functional form of $\pi(e_1, e_2)$ in order to simplify the set of constraints. These approaches are now discussed in some detail.

A Hands-On Analysis of the Joint Deviations Problem

Within the first approach, the main contribution is Fernandes and Phelan (2000). They provide a tractable setup in which an augmented recursive formulation of the problem of the principal is possible. Intuitively, this formulation has an increased number of state variables with respect to the recursive formulation of the moral hazard problem without persistence first presented in Spear and Srivastava (1987). The simplified framework that allows for the recursive formulation limits the effort choices and the output realizations to two. Also, the contract lasts for an infinite number of periods but persistence lasts only for one period; that is, effort at time t affects only the probability distribution over outcomes at time t and $t+1$. The recursive formulation of the problem of the principal has three state variables, one of which is the standard promised utility on Spear and Srivastava's formulation. The two extra states allow the principal to keep track of the marginal disutility of effort for the agent across periods, as well as the set of utilities achievable by the agent off the equilibrium path.

Still within the first approach, Mukoyama and Şahin (2005) limit the effort choices and the output values to two and analyze a two-period problem. They assume that high effort is optimal every period. They are able to provide analytical conditions on the conditional probability function under which the implications of persistence are drastically different than those of no persistence: When the first-period effort affects the second-period probability in a sufficiently stronger way than the second-period effort, the optimal contract exhibits perfect insurance in the initial period. Using a recursive formulation in the spirit of Fernandes and Phelan (2000), Mukoyama and Şahin also analyze a three-period problem numerically.

Kwon (2006) uses a very similar framework with discrete effort choices (0 or 1), also assuming that high effort is implemented every period. He imposes concavity of $\pi(\cdot)$ on the sum of past effort choices, so past effort is more effective than current effort. These assumptions allow him to analyze a $T > 2$ period problem that shares the same perfect insurance characteristic as in Mukoyama and Şahin (2005).

A Particularly Simple Case of Persistence

The second approach, presented in Jarque (2010), allows for a continuum of effort choices but assumes that the conditional probability depends on past effort choices only through the sum of undepreciated effort in the same manner as stated in Assumption 2. Note that, even for a concave probability function $\pi(s)$, Assumption 2 implies that past effort is less effective than current effort in contrast to what was assumed in Mukoyama and Şahin (2005) or Kwon (2006). The article shows that, for a subset of problems with this particular form of persistence, the computation of the optimal contract simplifies considerably. For these problems, an auxiliary standard repeated moral hazard problem without persistence can be used to recover the solution to the optimal contract. The linearity in effort of both variable s (which determines the probability distribution) and the utility of the agent dramatically simplifies the structure of the joint deviations across periods; in practice, we can think of s as the choice variable, and the structure of the resulting transformed problem is (under some conditions) equivalent to that of a standard repeated moral hazard.

In the next section, a finite version of the model in Jarque (2010) is presented and this result is explained in detail. The finite version allows for the numerical computation of the optimal contract in an example in which the stochastic structure is interpreted as unobservable human capital accumulation.

5. HIDDEN HUMAN CAPITAL ACCUMULATION

The problem of the principal is again as in problem SB, but now we consider the case $\rho > 0$. We argued in Section 4 that this case is more complicated because of the dependence of second-period utility and optimal actions of the agent on first-period choices. In order to go around some of these difficulties, here we adapt to our two-period finite example the strategy presented in Jarque (2010) for solving problems with persistence. Following this work we will show that, under our assumptions, the structure of the problem simplifies to that of the standard repeated moral hazard presented above, *provided the domain constraints in (ED) do not bind*. This is an important qualification since, as we learned when analyzing the case of observable human capital accumulation in Section 2, in the presence of persistence the effort domain constraints in (ED) will sometimes bind, especially for high values of the persistence parameter ρ. To deal with this issue, we follow the approach in Jarque (2010): First, we find a candidate solution assuming that the constraint in (ED) does not bind. Then we need to check numerically that this constraint is indeed satisfied to be sure that we have found a true solution. Unfortunately, a general analysis of the optimization problem of the principal including the inequality constraints for effort (again, as in Section 2) is more difficult with unobserved effort. Hence, finding the properties of the general case when constraint (ED) binds remains a question for future research.

Rewriting the Problem

Jarque (2010) shows that, whenever the effort domain constraint (ED) is not binding, we can find the solution to the problem with persistence using a related RMH problem without persistence as an auxiliary problem. The key observation for that result is that we can write the expected utility of the agent, $W_0(\mathbf{u}, \mathbf{e})$, as a function of the s variable only. This is convenient because s is the variable that effectively determines the probability distribution over outcomes each period; different combinations of effort choices that give rise to the same s are equivalent both for the principal and for the agent. Hence, once we rewrite the problem with s as the choice variable, there is no need to consider joint deviations across periods, the recursive structure is recovered, and we can solve for the optimal contract as we do with a standard repeated moral hazard.

Let $\widetilde{W}_0(\mathbf{u}, \mathbf{s}) = W_0(\mathbf{u}, \mathbf{e})$ for all the pairs of **s** and **e** sequences such that **s** results from effort choices in **e** according to the law of accumulation of human capital in (1). Writing the effort in the second period as

$$e_{2i} = s_{2i} - \rho s_1,$$

we have

$$\begin{aligned}\widetilde{W}_0(\mathbf{u}, \mathbf{s}) &= \sum_{i=L,H} \pi_i(s_1) u_i - v s_1 \\ &\quad + \beta \sum_{i=L,H} \pi_i(s_1) \left[\sum_{j=L,H} \pi_j(s_{2i}) u_{ij} - v(s_{2i} - \rho s_1) \right].\end{aligned}$$

Note that we have explicitly written the utility accrued in the first period in the first row of this expression, and that of the second period in the second row. With utility spelled out this way it is easy to see that, although s_1 is all accumulated in the first period, it appears both in the first- and second-period utility. Also, since s_1 is not contingent on any realization, it appears in the second period both after observing a first-period y_H and a first-period y_L. Hence, we can group the s_1 terms of the second period together with those of the first, to get an expression of the form

$$\begin{aligned}\widetilde{W}_0(\mathbf{u}, \mathbf{s}) &= \sum_{i=L,H} \pi_i(s_1) u_i - v(1 - \beta\rho) s_1 \\ &\quad + \beta \sum_{i=L,H} \pi_i(s_1) \left[\sum_{j=L,H} \pi_j(s_{2i}) u_{ij} - v s_{2i} \right]. \qquad (29)\end{aligned}$$

This allows us to interpret s as the variable being chosen by the agent. In the first period, we can interpret $v(1 - \beta\rho)$ as the "marginal disutility of exerting s_1." In the second period, the "marginal disutility of exerting s_2" is instead v.

This rearrangement of terms and thinking about s as the choice variable is a useful trick. Note that in the second row the expression inside the square brackets is independent of s_1. Interpreting s_{2i} as the choice variable, we can see that we can do here as we did in the case of no persistence and write the continuation utility of the agent independently of the first period's choice for s_1:

$$\sum_{j=L,H} \left[\pi_j(s_{2i}) u_{ij} - v s_{2i} \right] = \widetilde{W}'_{1i}(\mathbf{u}, s_{2i}).$$

Hence, we obtain expressions that parallel those of the standard RMH formulation in (19). The expression in (29) can then simply be rewritten as

$$\widetilde{W}_0(\mathbf{u}, \mathbf{s}) = \sum_{i=L,H} \pi_i(s_1) \left[u_i + \beta \widetilde{W}'_{1i}(\mathbf{u}, s_{2i}) \right] - v(1 - \beta\rho) s_1.$$

Note also that the structure of the incentive constraints simplifies as it did in the case of the RMH; in the second period, the first-period choice of s drops out:

$$\sum_{j=L,H} \pi_j(s_{2i})\, u_{ij} - v s_{2i} - \rho v \widehat{s_1} \geq \sum_{j=L,H} \pi_j(\widehat{s_{2i}})\, u_{ij} - v\widehat{s_{2i}} - v\rho\widehat{s_1}, \quad \forall \widehat{s_1}, \widehat{s_{2i}}.$$

Again, all these changes of notation are simply aimed at pointing to the following fact: The problem in which effort is persistent has a similar structure to that of a standard RMH problem in which s is interpreted as effort that is not persistent, but has marginal disutility of $v(1-\beta\rho)$ at $t=1$ and of v at $t=2$. To make this explicit, using the intertemporal regrouping of s_1, the problem of the principal in SB can be written as problem SB':

$$\begin{array}{rcl} & & \max_{\mathbf{u},\mathbf{s}} V(\mathbf{u},\mathbf{s}) \\ & & s.to \\ w_0 & \leq & \widetilde{W}_0(\mathbf{u},\mathbf{s}) \\ \widetilde{W}_0(\mathbf{u},\mathbf{s}) & \geq & \widetilde{W}_0(\mathbf{u},\widehat{\mathbf{s}}) \quad \forall \widehat{\mathbf{s}} \neq \mathbf{s} \\ u_i & \in & \mathcal{U}_i \quad \forall i \\ s_1 & \in & S_1 \\ s_{2i} & \in & S_2 \quad i = L, H, \end{array}$$

with $S_1 = \left[\underline{e}, \overline{e}\right]$ and $S_2 = \left[\rho s_1 + \underline{e}, \rho s_1 + \underline{e} + \overline{e}\right]$. This rewriting leads to the following observation: If problem SB' were in fact formally equivalent to a standard RMH problem (with the modified structure of the marginal disutility), this would help us enormously to find and characterize the solution to SB, since we would know how to solve it (or at least compute it numerically). However, a close inspection of SB' points to a small but potentially important difference with a standard RMH problem: In problem SB', the domain S_2 depends on the choice of s_1, while in a standard RMH problem this domain would be exogenously given.

Using a Related RMH Problem without Persistence as an Auxiliary Problem

Following Jarque (2010), we now show that, in some instances, we can work around the difficulty that an endogenous domain S_2 poses by using a related auxiliary problem for our purposes instead of SB'. Consider a problem SB_{aux} that is equal to SB' except for the domain S_2, which is substituted by an auxiliary domain $\widetilde{S}_2 = \left[\underline{e}, \overline{e}\right]$. Note that $\widetilde{S}_2$ is exogenous so, interpreting s as effort, problem SB_{aux} is a standard RMH. We will now argue that, under some conditions, the solution to SB' coincides with the solution to SB_{aux}, and hence we can easily obtain a solution to our problem with persistence.

The solution to problems SB' and SB_{aux} coincides when two conditions are satisfied: (i) $W_0(\cdot)$ is concave in **s,** and (ii) the resulting optimal choices for effort are interior. This is a set of sufficient conditions because if the expected utility of the agent is concave in his choice of s, then the relevant effort deviations are those close to the optimal (interior) s, and not those at the limits of the domain. This implies that using an auxiliary domain that does not exactly overlap with the true domain is not changing the solution to the problem, as long as this true solution is contained in the auxiliary domain. Are each of these conditions satisfied in our framework?

(i) **Concavity of** $W_0(\cdot)$ **in s.** In our particular example, it is easy to argue that the problem of the agent is concave in s_t for all t. In fact, the argument is the same that we used earlier to argue that problems PIC_1 and PIC_2 were concave: There are only two outcomes, the probability of observing y_H is concave in s_t, and current and future utility assigned to y_H is always higher than current and future utility assigned to y_L.

(ii) **Effort is interior**. This is not satisfied trivially. Constraint (ED) implies that two restrictions need to be checked to establish that the true solution is contained in the proposed auxiliary domain:

$$s_{2i} < \rho s_1 + \underline{e} + \overline{e}, \quad i = L, H, \tag{30}$$

$$s_{2i} > \rho s_1 + \underline{e}, \quad i = L, H. \tag{31}$$

Under the probability specification in (7), equation (30) is always satisfied. Other specifications are easy to find for which the upper bound of effort in (30) is not binding. The lower bound, however, is endogenous, and equation (31) cannot be checked without having the solution for s in hand. We conclude that the interiority cannot easily be guaranteed ex ante. The strategy to go around this problem that is proposed in Jarque (2010) is the following: Solve the problem assuming that the domain constraint can be substituted—and hence the equivalence to the RMH can be used—and then, with a candidate solution for s in hand, check the constraint ex post. We follow this route in the numerical computation of an example presented next. As it turns out, it is easy to find parametrizations for which the ex post check on the nonnegativity of effort is satisfied.

The Optimal Contract for Hidden Human Capital Accumulation

What do we conclude about the properties of the optimal contract in the presence of hidden human capital accumulation? Denote as $\widetilde{\mathbf{c}}^*$ and $\widetilde{\mathbf{e}}^*$ the solution to problem SB_{aux}. Whenever the sufficient conditions discussed above are satisfied, we have that, in the optimal contract:

1. The optimal consumption sequence in problem SB, $\mathbf{c}^*(P, SB)$, is equal to $\widetilde{\mathbf{c}}^*$.

2. The optimal human capital sequence in SB, $\mathbf{s}^*(P, SB)$, is equal to $\widetilde{\mathbf{e}}^*$.

3. The optimal effort sequence in SB, $\mathbf{e}^*(P, SB)$, can be recovered from the effort solution to problem SB_{aux} using

$$\begin{aligned} e_1^*(P, SB) &= \widetilde{e}_1^* \\ e_2^*(P, SB) &= \widetilde{e}_2^* - \rho \widetilde{e}_1^*. \end{aligned}$$

Importantly, the optimal consumption sequence has the same properties as in the solution to a standard RMH problem without persistence. Also, the optimal *human capital sequence* has the same properties as the *effort sequence* in a standard RMH problem. These properties were discussed at length in Section 3. Using these properties, we can reflect on the economic meaning of the ex post check implied by equation (31).

Whenever the ex post check in (31) is satisfied, the optimal contract asks the agent to increase human capital *in every period*. That is, the remaining level of human capital from the previous period, after depreciation, ρs_1, is never sufficient to cover the requirement of human capital for the current period, s_{2i} for $i = L, H$. In light of the properties of effort in a standard RMH problem, it is easy to see that this condition may not be satisfied in some examples since a decrease in the level of human capital from one period to the next could be part of the optimal solution for the principal. In particular, we learned in Section 3 that in an interior solution we will typically have $e_{2H} < e_1$, since the smoothing of incentives that is present in the first period is not available in the second, making effort in the second period relatively more expensive. Given the results we just established for the case with persistence, this means that we will typically have $s_{2H} < s_1$ in the optimal contract with hidden human capital accumulation. How does this lead to a violation of the ex post check in equation (31)? For certain parameters, we may have that s_{2H} is *so much smaller* than s_1 that, in fact, we have $s_{2H} < \rho s_1 + \underline{e}$, violating the interiority of effort choices. That is, if it were feasible, the principal would choose to have s_2 lower than $\rho s_1 + \underline{e}$. However, in the true problem with human capital accumulation (problem SB), effort needs to stay within its domain in each period, i.e., $e_{2i} > \underline{e}$ for all i, which rules out the possibility of decreasing s_2 below $\rho s_1 + \underline{e}$. Any adjustment should be made in the first period, when the principal anticipates the added cost of future incentives. That is, the solution for s_1 should differ from the one that was just presented. Unfortunately, characterizing how exactly the solution for s_1 changes is not easy. Solving for the optimal contract in this case becomes more complicated. As we argued, the independence of second-period choices from first-period choices breaks down, both for the principal and for the agent. In practice, even the numerical computation of examples is more involved, since all feasible combinations of effort across the two periods (and choices contingent on realizations of output) need to be tested for incentive compatibility. The simple recursive structure with w_{2i} as a state variable is no longer valid, and the dimensionality of the computational problem is similar to that of the strategy proposed in Fernandes and Phelan (2000).

The next section presents an example for which the ex post check in (31) is satisfied, and hence solving for the optimal contract is simple. Using the numerical solution, we discuss the implications of persistence for consumption and effort paths by comparing the solution to that of the case without persistence ($\rho = 0$).

6. NUMERICAL EXAMPLE WITH UNOBSERVED EFFORT: A COMPARISON

For cases in which the equivalence to a RMH is valid, we can find the solution to our problem with persistence using the usual numerical methods to solve standard RMH problems without persistence.

Figures 1 and 2 illustrate the implications for effort and consumption in the solution to an example with the parameter values listed in Table 1. The example without persistence has $\rho = 0$, while the example with persistence has $\rho = 0.2$. For the numerical examples we use the functional forms $u(c) = 2\sqrt{c}$ and the probability specification in (7). We also set $\underline{e} = 0.01$ and $\overline{e} = 0.99$ in order to restrict to cases with full support.

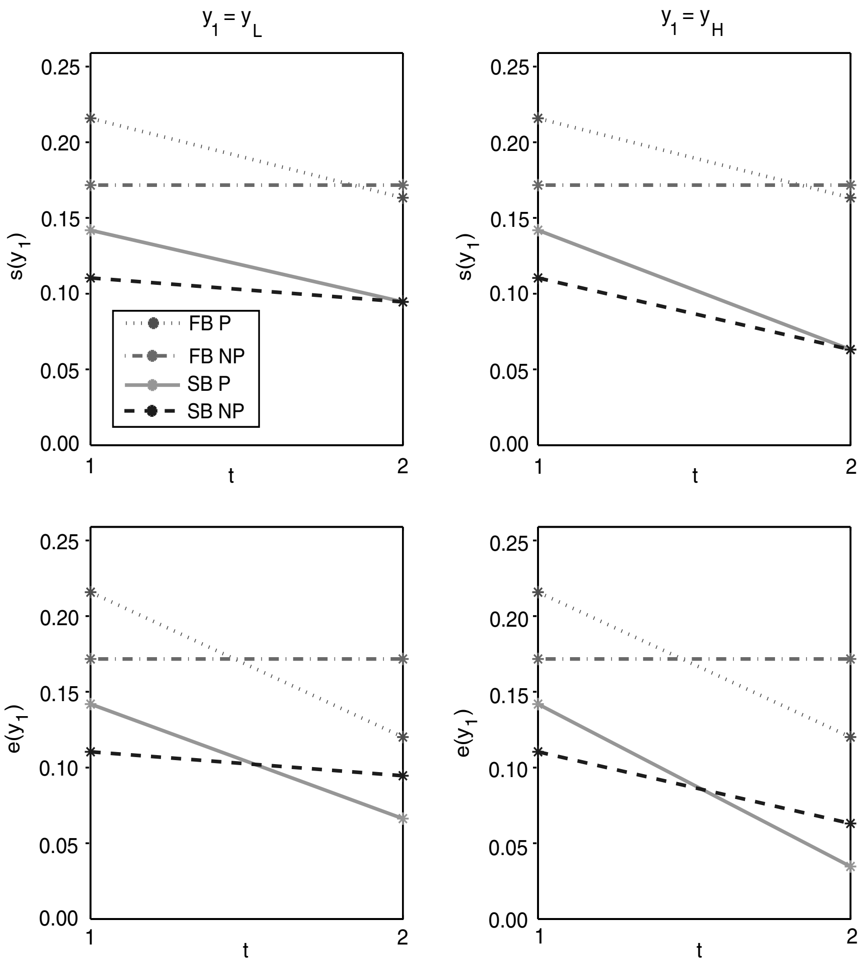
In Figure 1, the solution for s and e in the SB problem with persistence is plotted with a solid line. As we can see in the top panels, the level of s_1 in problem SB is always higher with persistence than without persistence (dashed line). Since $s_1 = e_1$, a higher level of s_1 with persistence reflects the fact that human capital is accumulated in the first period with the same cost as nonpersistent effort, but it lasts (partially) until the following period.[12]

The solutions for the paths of optimal s in the FB model are also represented in Figure 1 (dotted and dash-dotted line respectively for the persistent and nonpersistent case). The comparison clearly shows that human capital accumulation makes frontloading of s optimal. (This also translates into frontloading of effort as shown clearly in the bottom panels of Figure 1.) The main difference with the solutions to the respective SB problem is the level (higher in the FB problem). A second difference is that, even without persistence, in the second period the requirement for s may decrease in the SB problem, for incentive reasons, following both realizations (although the decrease may be more pronounced after y_H), and hence we have $s_1 > s_{2i}$ for all i.

As we can see in the bottom panels of the solution to the SB problem, both with persistence and without, effort is higher in the initial period than in

[12] The level of s_{2i} in this example coincides with and without persistence for all i. This is particular to this example and is violated if, for example, the level of w_0 is modified. Although human capital in the second period is equivalent to nonpersistent effort (because there are no further periods to exploit the persistence of human capital), the optimal choice for w_{2i} will typically be different across the two models.

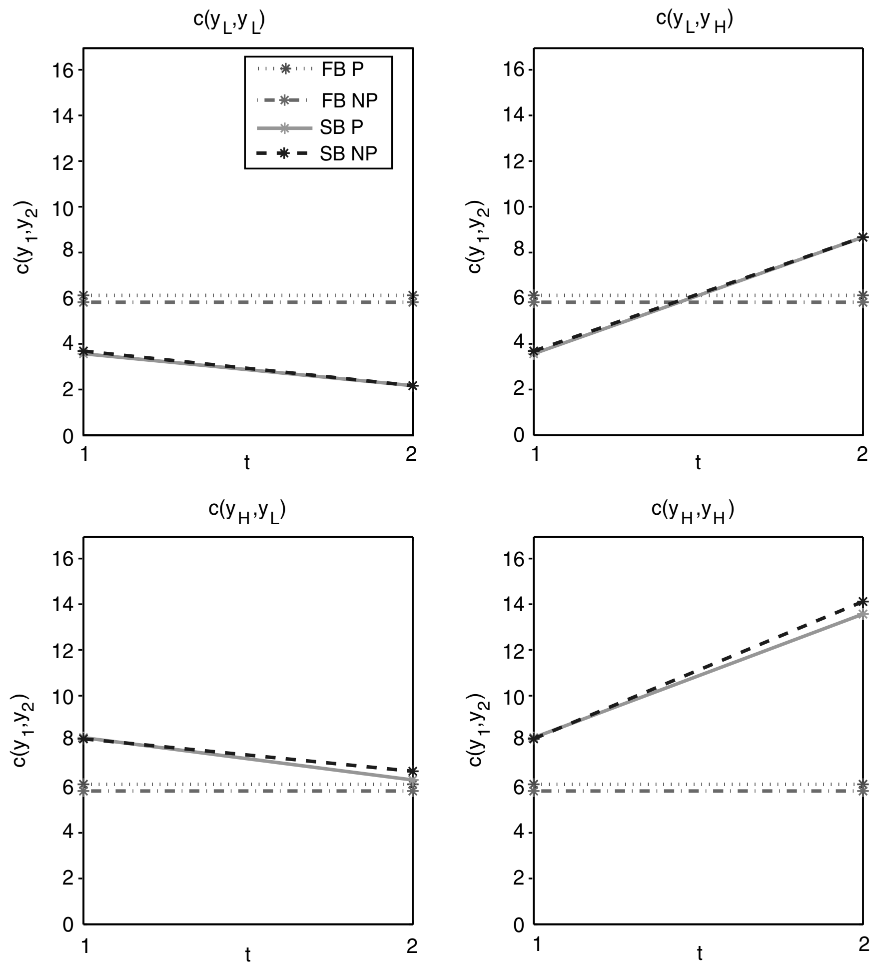
Figure 1 Contingent Paths for Human Capital and Effort in the Optimal Contract with and without Effort Persistence, both for the First-Best and the Second-Best Models



Notes: $s\,(y_1)$: human capital contingent on history y_1; $e\,(y_1)$: effort contingent on history y_1; P: see Table 3, $\rho = 0.2$; NP: see Table 3, $\rho = 0$.

the second. However, the frontloading of effort is much more pronounced with persistence. This is also true when comparing the solutions for the FB problem: While effort stays constant from one period to the next in the case without persistence, with persistence it is frontloaded, as discussed in Section 2.

Figure 2 Contingent Paths for Consumption in the Optimal Contract with and without Effort Persistence, both for the First-Best and the Second-Best Models



Notes: $c\,(y_1, y_2)$: consumption contingent on history (y_1, y_2); P: see Table 3, $\rho = 0.2$; NP: see Table 3, $\rho = 0$.

Consumption, depicted in Figure 2, is, in the SB case, virtually the same with and without persistence. It simply increases when the realization is y_H and decreases when it is y_L for the standard incentive provision reasons discussed in the earlier sections. However, we can see in the FB case that consumption is slightly lower in the case with persistence. Since the FB case

Table 3 Summary Statistics

	$\rho = 0.2$ **(FB)**		$\rho = 0.0$ **(FB)**		$\rho = 0.2$ **(SB)**		$\rho = 0.0$ **(SB)**	
	$t = 1$	$t = 2$	$t = 1$	$t = 2$	$t = 1$	$t = 2$	$t = 1$	$t = 2$
$E\left[c_t^*\right]$	6.12	6.12	5.82	5.82	5.30	5.47	5.16	5.30
$E\left[u\left(c_t^*\right)\right]$	4.95	4.95	4.83	4.83	8.26	11.32	7.74	10.90
$Var\left[c_t^*\right]$	0	0	0	0	4.96	13.69	4.34	13.96
$E\left[e_t^*\right]$	0.22	0.16	0.17	0.17	0.14	0.05	0.11	0.084
$Var\left[e_t^*\right]$	0	0	0	0	0	0.00023	0	0.00022
$E\left[s_t^*\right]$	0.22	0.16	0.17	0.17	0.14	0.0828	0.11	0.0842
$Var\left[s_t^*\right]$	0	0	0	0	0	0.00023	0	0.00022

is calculated numerically but without using a grid, we conclude that most likely consumption is also slightly lower with persistence in the true solution to the unobservable effort case.

Table 3 reports the value of some simple statistics of the comparison across the two models presented in Figures 1 and 2. The FB model statistics are included for reference, since they correspond to the solutions reported already in Sections 1 and 2. All expectations in the first period are conditional on s_1^*, and those in the second are conditional on s_{2i}^*. When comparing the statistics for the SB problem, we see that persistence implies a higher level of expected consumption, expected utility, and a slightly higher variance of consumption in the first period. When looking at these three moments across periods we see that persistence implies a steeper increase of expected consumption in time. Again, the statistics on consumption need to be interpreted with care since they are likely influenced by the use of a grid.

As for expected effort, we see that the level is higher with persistence in the initial period, but it drops below the no persistence case in the second period (a much steeper decrease than without persistence). The comparison of the expected accumulated human capital explains this: The expected level of s_1 with persistence is much higher than the level of e_1 without persistence, but the solution for s_2 with persistence is similar (in this particular example, identical) to the solution for e_2 without persistence.

7. CONCLUSION

When learning by doing is an important factor in a repeated agency relationship, solving for the optimal contract is generally very difficult. In the framework studied here, with linear disutility of effort and the productivity of the agent being a distributed lag of past efforts, we provide an example with a simple solution. This allows us to numerically establish some properties of the optimal contract. On one hand, the human capital of the agent in equilibrium and, hence, his productivity tend to be higher with learning by doing than without. Moreover, the optimal contract offered to the employee implies a lower productivity in the final years of the contract. The human capital of the agent is left to depreciate since, close to the end of the contract, the cost of incentives of requiring a higher productivity is not justified by the benefit of future productivity. This implies that, over the contractual relationship, effort is frontloaded and follows a steeper decreasing pattern than in the case without learning by doing. On the other hand, we find that the properties of wage dynamics remain unchanged with respect to those of the optimal contract without learning by doing.

REFERENCES

Ales, Laurence, and Pricila Maziero. 2009. “Accounting for Private Information.” Mimeo.

Arrow, Kenneth J. 1962. “The Economic Implications of Learning by Doing.” *The Review of Economic Studies* 29 (June): 155–73.

Fernandes, Ana, and Christopher Phelan. 2000. “A Recursive Formulation for Repeated Agency with History Dependence.” *Journal of Economic Theory* 91 (April): 223–47.

Gibbons, Robert, and Kevin J. Murphy. 1992. “Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence.” *Journal of Political Economy* 100 (June): 468–505.

Grossman, Sanford J., and Oliver D. Hart. 1983. “An Analysis of the Principal-Agent Problem.” *Econometrica* 51 (January): 7–45.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. “Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents.” *Review of Economic Dynamics* 1 (January): 1–58.

Jarque, Arantxa. 2010. “Repeated Moral Hazard with Effort Persistence.” *Journal of Economic Theory* 145 (November): 2,412–23.

Jewitt, Ian. 1988. “Justifying the First-Order Approach to Principal-Agent Problems.” *Econometrica* 56 (September): 1,177–90.

Kapička, Marek. 2008. “Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach.” Mimeo, University of California, Santa Barbara.

Kwon, Illoong. 2006. "Incentives, Wages, and Promotions: Theory and Evidence." *RAND Journal of Economics* 37 (Spring): 100–20.

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent. 2009. "Performance Pay and Wage Inequality." *Quarterly Journal of Economics* 124 (February): 1–49.

Lucas, Robert E., Jr. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (July): 3–42.

MacLeod, W. Bentley, and Daniel Parent. 1999. "Job Characteristics, Wages, and the Employment Contract." Federal Reserve Bank of St. Louis *Review* (May): 13–27.

Mukoyama, Toshihiko, and Ayşegül Şahin. 2005. "Repeated Moral Hazard with Persistence." *Economic Theory* 25: 831–54.

Phelan, Christopher. 1994. "Incentives, Insurance, and the Variability of Consumption and Leisure." *Journal of Economic Dynamics Control* 18: 581–99.

Phelan, Christopher, and Robert M. Townsend. 1991. "Computing Multi-Period, Information-Constrained Optima." *Review of Economic Studies* 58 (October): 853–81.

Prescott, Edward S. 1999. "A Primer on Moral-Hazard Models." Federal Reserve Bank of Richmond *Economic Quarterly* 85 (Winter): 47–77.

Rogerson, William P. 1985a. "Repeated Moral Hazard." *Econometrica* 53 (January): 69–76.

Rogerson, William P. 1985b. "The First-Order Approach to Principal-Agent Problems." *Econometrica* 53 (November): 1,357–67.

Spear, Stephen E., and Sanjay Srivastava. 1987. "On Repeated Moral Hazard with Discounting." *Review of Economic Studies* 54 (October): 599–617.

Wang, Cheng. 1997. "Incentives, CEO Compensation and Shareholder Wealth in a Dynamic Agency Model." *Journal of Economic Theory* 76 (September): 72–105.