

Learning in a Hiring Logic and Optimal Contracts

JOSE IRANILDO SALES¹ AND FRANCISCO DE SOUSA RAMOS²

¹Post Graduate Program in Production Engineering, Federal University of Pernambuco, Recife 50670-901, Brazil

²LabGRC, Department of Economics, Federal University of Pernambuco, Recife 50670-901, Brazil

Corresponding author: Jose Iranildo Sales (iranildo.sales2@gmail.com)

ABSTRACT This paper examines a hiring logic problem in which all players involved in this game are exposed to scenarios where they can learn from the changes and these modifications influence their preferences; consequently, their decision-making differs from the classical agency theory proposed by [1]. Therefore, how this new learning approach of the agents involved in the delegation of activities changes the methodology of the hiring logic. Concepts such as nonlinear preferences, partial understanding of performances by repetition, and economic cycles of employability are introduced into the classical model, bringing a series of significant changes in the structuring of the game according to the perception and knowledge of the agents involved in the model. As a result, the model indicates that there is a different way to understand hiring logic using the principal-agent model, in which the optimal contract is adapted for learning agents, due to the natural change of behavior by changing perception and preferences in the game.

INDEX TERMS Agency theory, hiring logic, learning, nonlinearity, repeated games.

I. INTRODUCTION

The optimal formulation of labor contracts has always been a topic of discussion regarding the incentives present in the relationship between the agents involved in the hiring process, due to the series of factors present in the scenario. Game theory can help us understand the main issues involved in this model to analyze the different scenarios present in this context [13].

In recent decades, several researchers [2]–[5] studied the behavior of the relationship between employers (principal) and employees (agents) in a hiring logic and optimal contracts of labor worldwide. As a result, many questions have been raised, such as the incentives found in each relationship, the kind of information presented in the game, the possibility of controlling the task delegation of hiring to different kinds of agents, and the behavior of the preferences. Understanding these dynamic incentives is central to decision-making.

According to [6] the most game-theoretic research relies on the availability of spectrum statistics in order to formulate the game and cope with spectrum dynamic changes, especially in stochastic [7] and repeated games [8]. Such information is not known a priori, limiting the applicability of this approach [9].

In hiring logic, an account must be taken of the degree of agents' global perception of the scenario in which they are

involved. In other words, what is needed to understand how much the players learn from basic differences in the scenario, such as changes in performance to a greater perception of the game, understanding (the gains of other players, the greater the degree of effort needed to perform a certain task, understanding changes in the labor market, etc. This corroborates the fact that there have been changes in the agent's preferences, which modifies how to draw up an optimal contract for each situation.

Therefore, models with dynamic contracts with limited commitment [10]–[12] are one way to investigate problems of optimal contracting in environments where one or more of the contracting parties are exposed to outside opportunities that tempt them to leave the ongoing relationship.

On the other hand, contract theory, which focuses on non-monetary incentives, is set out by [13], which introduces goal setting into the hiring scenario and demonstrates that monetary incentives are not a unique motivation tool. Moreover, [14] criticize the unilateralism of agency theory and seek a way to formulating contracts bilaterally and [15] considers a principal-multiple agent model in which agents are privately informed about their intrinsic motivations for collaborating or for competing.

Furthermore, hiring processes have a long tradition of explicit and implicit human biases, which may lead to different consequences [4]. There are several kinds of bias [16], [15] that may influence hiring decisions. In all

The associate editor coordinating the review of this manuscript and approving it for publication was Keli Xiao¹.

these scenarios, the relationship between the players is based on the preference structure of the agent involved in the process. Thus, an analysis must be conducted regarding the consequences of different types of changes in the model.

To analyze this context, we use the principal-agent theory, which has been a very influential theoretical apparatus for gaining insights into the design of labor incentive contracts [17]–[21]. In most principal-agent relationships, the principal has to induce the agent to engage in several tasks simultaneously.

The agent's performance can often be measured fairly accurately in some tasks, but in others, the available performance measures may be very noisy, or a verifiable performance measure that can be used to provide explicit incentives to the agent may not even exist.

For example, a production worker may have to produce a certain amount of output that is easily measurable, but he may also have to ensure that the quality of output is high and that the machinery he is working with is properly maintained, which may be more difficult to monitor.

Another example is a schoolteacher who has to teach his students basic skills, such as the three R's (reading, writing, arithmetic), which can be measured in standardized tests, but also has to stimulate their creativity and teach them social skills, which are much harder to evaluate [22].

Therefore, the starting point of incentive theory in a labor contract corresponds to the problem of delegating a task to an agent with private information. This private information can be of two types: either the agent can take an action unobserved by the principal called hidden action, or the agent has some private information about its cost or valuation that the principal is unaware of, called hidden knowledge [1].

When two parties engage in a business relationship, their interests are usually not perfectly aligned, and information asymmetry can further exacerbate the tension between them. The principal-agent model is a stylish framework for studying such a problem [23]. Furthermore, the agents performing the task incur marginal costs that are associated with the number of units of work to be carried out [1].

However, according to the theory of labor [24], the marginal cost of doing one more one unit of any activity increases by degrees i.e. the more the individual works the greater the marginal cost of performing a certain activity, which [24] called the marginal disutility of labor.

Therefore, the agent's utility is affected by the increase in the total cost of production, and utility decreases because the rate of pleasure increases, which changes the strategies the principal must use to draw up the contract.

Moreover, in a principal-agent situation, the agent chooses an action "on behalf of" the principal. The consequence of this depends on the random state of the environment as well as on the agent's action. After observing the consequence, the principal makes a payment to the agent according to a pre-announced reward function, which depends directly only on the consequence observed [25].

Another approach to increasing efficiency is the theory of repeated games. If a game with two or more players is repeated, the resulting situation can be modeled naturally as a "supergame" in which the players' actions in any one repetition are allowed to depend on the history of the previous repetitions. In the principal-agent situation, the repetition of the game gives the principal an opportunity to observe the results of the agent's actions over a number of periods and use a statistical test to infer whether or not the agent was choosing the appropriate actions.

The repetition of the game would also provide the principal with opportunities to "punish" the agent for apparent departures from the appropriate actions.

The nonlinearity and the repetition included in a labor contract, using the principal-agent model, changes the structure of the utility functions for all the agents involved in the game.

Repetition is possible because when a principal observes the agent's performance, his observation is not conclusive, and it may vary his performance in the next hiring due to the individual's fluctuations that arise from the perception of changes in the market and in their behavior. In addition, the nonlinearity is justified because of the theory set out by [24] in the marginal disutility of labor. In both cases, the players involved in the hiring logic learn from the changes imposed by the scenario, which can alter the terms of the optimal contract proposed by [1].

This study seeks to analyze the differences in the optimal contract formulation proposed by [1], when all players learn from the scenarios imposed on them. The learning deals with changes in the utility structures of each player for each scenario, whether these are the marginal disutility of labor, the partial understanding of performance by the repetition game of hiring in the repeated model, or the influence of market economic cycles on the preferences of those involved.

Section 2 describes the nonlinear model of the preferences. Section 3 examines the optimal contract under the assumption that repetition can influence the understanding of the performance of all players. Section 4 discusses the impact of structuring the utility of understanding economic employability cycles. Finally, Section 5 concludes the paper.

II. THE NONLINEAR MODEL OF PREFERENCES

According to [24], labor is the painful exertion, which we undergo to ward off pains of greater amount, or to procure pleasures, which leaves a balance in our favor. Labor can also be agreeable at the time and conducive to future good; however, it is only agreeable in a limited amount, and most people are compelled by their wants to exert themselves longer and more severely than they would otherwise do.

Each job requires a different level of effort from the agents. To define this level of effort, [24] shows that the amount of labor will be a quantity of two dimensions (intensity and time). Intensity of labor may have more than one meaning; it may mean the amount of work done, or the painfulness of the effort of doing it. Therefore, the theory of labor involves

three parameters: the amount of painful exertion, the amount of produce, and the amount of utility gained.

The individual utility of an agent can be modified by using two parameters: Besides the quantity produced, he can vary his utility by the level of effort employed due to the amount of painful exertion. Experience shows that as labor is prolonged, the effort becomes increasingly painful as a general rule. A few hours of work per day may be considered agreeable rather than otherwise; however, as soon as the overflowing energy of the body is drained off, it becomes irksome to remain at work. As exhaustion approaches, continued effort becomes less and less tolerable.

To explain how utility decreases when effort increases because of the increase in the amount of work, some aspects of these phenomena are described in Fig 1.

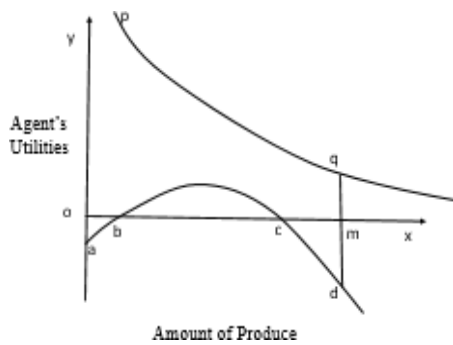


FIGURE 1. Marginal disutility of labor (Jevons, 1957).

In Fig 1, the height of the points above the line *ox* denotes pleasure and the points below it denotes displeasure. The moment of commencing labor is usually more irregular than when the mind and body are bent well to the work. Therefore, pain was measured using *the oa*. At *b*, there was neither pain nor pleasure. Between *b* and *c*, an excess of pleasure is represented as being due to exertion. However, after *c*, energy begins to be rapidly exhausted, and the resulting pain is shown by the downward tendency of line *cd*. At the same time, we may represent the degree of utility of the produce by some such curve as *pq*, the amount of produce being measured along the line *ox*.

Therefore, the utility of the agent decreases with increasing rate, whereas the marginal cost increases with decreasing rate. Thus, there is nonlinearity in labor contracts. The principal-agent model is an application of work contracts, and therefore, this model must be analyzed using the marginal disutility of labor concepts.

A. LEARNING IN THE NONLINEAR MODEL

An understanding of an activity is provided by monitoring activity and practice. Hence, players involved in a work contract learn from the daily routine, that is, with the exercise of the function. Therefore, the optimal contract formulation must consider the effort made by the agents and the

consequences of this in utility structures, as shown in the model below:

1) LEARNING WITH THE LABOR: UNOBSERVABLE COSTS

According to [1], a contract can be formulated between an individual who wishes to delegate a particular activity to another agent, but he is not aware of the actual performance of the agent. The incentives of the agents involved in the game are determined by their utility functions that describe the costs and benefits of the players. This contract is designed considering a linear relationship between the costs of carrying out the activity and the amount of labor to be performed, according the classical agency theory elaborated by [1] and as a way to simplify the approach.

However, the production costs of doing the work may not have a linear relationship with the time or quantity of work performed, because the agents learn from the activity.

With the theory of the marginal disutility of labor, it is demonstrated that the costs of performing a certain activity increases and that to produce *q* quantities costs the agent less than *3q*, not only because of the quantity realized, but also because of the cost of the growth in the marginal production of *q* quantities.

Agents can have varying levels of efficiency, with a known assumed probability distribution. It is said, then, that agents can be of different types. By way of simplification of the model the agents involved in this activity are characterized as efficient agent, those who produced larger quantities because they have lower marginal production costs and inefficient agent, which are those who produced smaller quantities because they have higher marginal costs of production.

Thus, the utility functions of the agents involved change as follows:

$$\underline{U} = \underline{t} - (\underline{\theta})^\alpha \underline{q} \tag{1}$$

$$\bar{U} = \bar{t} - (\bar{\theta})^\alpha \bar{q} \tag{2}$$

- \underline{U} is a high performance agent's utility;
- \bar{U} is a low performance agent's utility;
- \underline{q} is the quantity produced for de efficient agent;
- \bar{q} is the quantity produced for de inefficient agent;
- $\underline{\theta}$ denotes the marginal cost of the efficient agent;
- $\bar{\theta}$ denotes the marginal cost of the inefficient agent;
- \underline{t} is the amount paid for the normal work of an efficient agent;
- \bar{t} is the amount paid for the normal work of an inefficient agent;
- α is a parameter that demonstrates that the cost grows at increasing rates with the realization of an additional unit of work, because of the learning, with $\alpha > 1$.

Moreover, by analyzing the structure developed [1], it is not easily to visualize the need for changes in the quantity requested by the contracting agent for sporadic and/or floating demands of products and, in several cases, this situation can be found. To introduce the possibility of additional

requests appearing, the contract must be drawn up by considering the utility of the agents, whether additional demands do or do not occur. Thus, the utility of the agents is:

$$\underline{U} = (k) \left((\underline{t} + \underline{t}_e) - (\underline{\theta})^\alpha \underline{q} + (\underline{\gamma}^\beta) \underline{q}_e \right) + (1 - k) \left(\underline{t} - (\underline{\theta})^\alpha \underline{q} \right) \quad (3)$$

$$\bar{U} = (k) \left((\bar{t} + \bar{t}_e) - (\bar{\theta})^\alpha \bar{q} + (\bar{\gamma}^\beta) \bar{q}_e \right) + (1 - k) \left(\bar{t} - (\bar{\theta})^\alpha \bar{q} \right) \quad (4)$$

where:

- k is the probability of the need for additional demand on the part of the principal, and $(1 - k)$ is the complementary probability.
- θ is the marginal cost for each q unit stipulated in the contract for regular hours.
- γ is the marginal cost for each additional unit stipulated in the contract, in which $\gamma - \theta > 0$.
- \underline{t} is the amount paid for the normal work of an efficient agent;
- \bar{t} is the amount paid for the normal work of an inefficient agent;
- \underline{t}_e is the amount paid for the additional work of an efficient agent;
- \bar{t}_e is the amount paid for the additional work of an inefficient agent;
- \underline{q}_e is the additional amount worked by an efficient agent;
- \bar{q}_e is the additional amount worked by an inefficient agent;
- α is the factor that determines the rate of growth of the marginal cost by the increase and effort of the agent in the units defined in the contract because of learning.
- β is the factor that determines the rate of growth of the marginal cost owing to the increase and effort of the agent in the additional units. Given $\alpha > 1$ and $\beta > 1$.

The agents need a minimal benefit to accept the principal's proposal, considering that any external opportunities are equal to zero, that is, any external alternatives are less attractive (less profitable). Thus, new participation constraints are added to those that already exist. Therefore, knowing that the agent accepts the contract of the normal and additional units.

$$(k) \left((\underline{t} + \underline{t}_e) - (\underline{\theta})^\alpha \underline{q} + (\underline{\gamma}^\beta) \underline{q}_e \right) + (1 - k) \left(\underline{t} - (\underline{\theta})^\alpha \underline{q} \right) \geq 0 \quad (5)$$

$$(k) \left((\bar{t} + \bar{t}_e) - (\bar{\theta})^\alpha \bar{q} + (\bar{\gamma}^\beta) \bar{q}_e \right) + (1 - k) \left(\bar{t} - (\bar{\theta})^\alpha \bar{q} \right) \geq 0 \quad (6)$$

Hence, the agents need to be discouraged from changing productive behavior; thus, the following incentive compatibility constraints are present.

$$(k) \left((\underline{t} + \underline{t}_e) - (\underline{\theta})^\alpha \underline{q} + (\underline{\gamma}^\beta) \underline{q}_e \right) + (1 - k) \left(\underline{t} - (\underline{\theta})^\alpha \underline{q} \right) \geq (k) \left((\bar{t} + \bar{t}_e) - (\bar{\theta})^\alpha \bar{q} + (\bar{\gamma}^\beta) \bar{q}_e \right) + (1 - k) \left(\bar{t} - (\bar{\theta})^\alpha \bar{q} \right) \quad (7)$$

$$(k) \left((\bar{t} + \bar{t}_e) - (\bar{\theta})^\alpha \bar{q} + (\bar{\gamma}^\beta) \bar{q}_e \right) + (1 - k) \left(\bar{t} - (\bar{\theta})^\alpha \bar{q} \right) \geq (k) \left((\underline{t} + \underline{t}_e) - (\underline{\theta})^\alpha \underline{q} + (\underline{\gamma}^\beta) \underline{q}_e \right) + (1 - k) \left(\underline{t} - (\underline{\theta})^\alpha \underline{q} \right) \quad (8)$$

Finally, the principal needs to maximize his utility, called U_p , by formulating a contract that is the most beneficial to him, that is, knowing that there is asymmetry of information and, consequently, it is known that the utility of the principal is:

$$U_p = \text{Max } k \left(v \left(S(\underline{q}) + S(\underline{q}_e) - \left((\underline{\theta})^\alpha \underline{q} + (\underline{\gamma}^\beta) \underline{q}_e \right) \right) + (1 - v) \left(S(\bar{q}) + S(\bar{q}_e) - \left((\bar{\theta})^\alpha \bar{q} + (\bar{\gamma}^\beta) \bar{q}_e \right) \right) \right) + (1 - k) \left(v \left(S(\underline{q}) - (\underline{\theta})^\alpha \underline{q} \right) + (1 - v) \left(S(\bar{q}) - (\bar{\theta})^\alpha \bar{q} \right) \right) \quad (9)$$

where:

- S is the benefit that the principal get when the agents produces q units; where $S'(q) > 0$, $S''(q) < 0$ and $S(0) = 0$.
- v is the probability to find efficient agents.

As t refers specifically to the benefit that accrues to agents from their costs, one can rewrite the utility of the principal by considering costs and information rent, which occurs when the principal is an environment of incomplete information, that is, he wants to delegate an activity to two possible types of agents, but does not know their real productivity.

$$U_p = k \left(v \left(S(\underline{q}) + S(\underline{q}_e) - \left((\underline{\theta})^\alpha \underline{q} + (\underline{\gamma}^\beta) \underline{q}_e \right) \right) + (1 - v) \left(S(\bar{q}) + S(\bar{q}_e) - \left((\bar{\theta})^\alpha \bar{q} + (\bar{\gamma}^\beta) \bar{q}_e \right) \right) \right) + v \left(\Delta \theta^\alpha \bar{q} + \Delta \gamma^\beta \bar{q}_e \right) + (1 - k) \left(v \left(S(\underline{q}) + (\underline{\theta})^\alpha \underline{q} \right) + (1 - v) \left(S(\bar{q}) - (\bar{\theta})^\alpha \bar{q} \right) + v \left(\Delta \theta^\alpha \bar{q} \right) \right) \quad (10)$$

The utility of the principal this in the face of an expected demand, which it will seek to maximize. In addition, based on the probabilities associated with the unawareness of the agents' behavior represented by v and $(1-v)$, the benefits to the principal represented by $S(\underline{q})$ and $S(\bar{q})$, and by the marginal costs of the principal with the contracting of each profile without the possibility of additional demands $(\underline{\theta})^\alpha \underline{q}$ and $(\bar{\theta})^\alpha \bar{q}$ and as the possibility of additional demands $(\underline{\theta})^\alpha \underline{q} + (\underline{\gamma}^\beta) \underline{q}_e$ and $(\bar{\theta})^\alpha \bar{q} + (\bar{\gamma}^\beta) \bar{q}_e$. All of them represent the allocative efficiency of the utility of the principal, also possessing the informational rent with and without additional demands $v(\Delta \theta^\alpha \bar{q} + \Delta \gamma^\beta \bar{q}_e)$ and $v(\Delta \theta^\alpha \bar{q})$.

The first-order conditions for each agent must be determined because the principal needs to maximize its utility. Given that U_p is a strictly concave function, we have a global maximum.

Proposition 1: If there is no possibility of additional demands

(i.e., $k = 0$), the expected results are

$$S'(\bar{q}) = \alpha (\bar{\theta q})^{\alpha-1} \theta + \frac{v}{1-v} \alpha (\Delta \bar{\theta q})^{\alpha-1} \Delta \theta \quad (11)$$

$$S'(\underline{q}) = \alpha (\underline{\theta q})^{\alpha-1} \theta \quad (12)$$

Compared with the classic model of Laffont and Martimort (2001), the maximization result for the principal must consider more factors than only $\underline{\theta}$ and $\bar{\theta}$. Given that $\alpha (\underline{\theta q})^{\alpha-1} \theta > \underline{\theta}$ and $\alpha (\bar{\theta q})^{\alpha-1} \theta > \bar{\theta}$ with $\alpha > 1$, the principal needs to pay the agents more.

The understanding of the marginal disutility of labor is a driver of the agents' results, which leads to more principal spending.

Proposition 2: If there is the possibility of additional demands, then $k > 0$, the expected results are as follows:

$$S'(\underline{q}) = \alpha (\underline{\theta q})^{\alpha-1} \theta \quad (13)$$

$$S'(\bar{q}) = \alpha (\bar{\theta q})^{\alpha-1} \theta + \frac{v}{1-v} \alpha (\Delta \bar{\theta q})^{\alpha-1} \Delta \theta \quad (14)$$

$$S'(\underline{q}_e) = \beta (\underline{\gamma q})^{\beta-1} \gamma \quad (15)$$

$$S'(\bar{q}_e) = \beta (\bar{\gamma q})^{\beta-1} \gamma + \frac{kv}{(1-k)(1-v)} \beta (\Delta \bar{\gamma q})^{\beta-1} \Delta \gamma \quad (16)$$

Compared with the classic model of Laffont and Martimort (2001), the maximization result for the principal brought other variables that leverage the agents' marginal costs both in the normal amount worked and in the additional amount. Knowing that $\beta > \alpha$, this disutility is even greater in the additional quantities, which makes the principal reformulate his proposals for the most efficient agent with larger disbursements.

2) LEARNING WITH LABOR: EFFORT LEVEL

According to the classical principal-agent model, we consider an agent who can exert a costly effort, and this effort can take two possible values that we normalize as a zero-effort level and a positive effort of one: $[0,1]$. Exerting effort e implies a disutility for the agent that is equal to $\psi(e)$ with the normalizations $\psi(0) = 0$ and $\psi(1) = \psi$. Thus, the positive value was equal to 1.

However, when we consider the disutility of the agent as a function $\psi(e)$, we admit that e is a constant value that varies between 0 and 1. Therefore, $\psi(e) = e$.

Using the concept of learning in the marginal disutility of labor, the effort e varies with the production level, and e increases when the level of production increases. Therefore, the disutility function is $\psi(e) = e^\alpha$, where α represents the exponential growth of the production level, with $\alpha > 1$. In this sense, e is not constant.

The agent's utility changes to $U = u(t) - e^\alpha$, with u increasing and concave ($u' > 0$ $u'' < 0$).

The principal's expected utility of the agent is written as:

$$V_1 = \pi_1 (S(\bar{q}) - \bar{t}) + (1 - \pi_1) (S(\underline{q}) - \underline{t}) \quad (17)$$

If the agent makes a positive effort ($e = 1$), and

$$V_0 = \pi_0 (S(\bar{q}) - \bar{t}) + (1 - \pi_0) (S(\underline{q}) - \underline{t}) \quad (18)$$

If the agent makes a negative effort ($e = 0$).

Where:

- $S(\bar{q})$ is principal benefits of a low performance agent's work;
- $S(\underline{q})$ is principal benefits of a high performance agent's work;
- \underline{t} is the amount paid for the normal work of an efficient agent;
- \bar{t} is the amount paid for the normal work of an inefficient agent;
- π_1 is the probability associated to level of effort equal to 1;
- π_0 is the probability associated to level of effort equal to zero.

The principal wishes to induce a positive effort ($e = 1$) to maximize his utility, such that $V_1 \geq V_0$. However, when we introduce the new agent's utility function $U = u(t) - e^\alpha$, the exponential disutility is inserted in the context, thus changing the model. Knowing this, $e^\alpha = \psi(e)$ only when $\alpha = 1$.

Given that the production level q increases, the effort increases at an increasing rate with $\alpha > 1$. Therefore, the disutility of the agent is higher than that of the classical model, and α tends to increase with the growth of disutility.

Therefore, the deadline must be found that makes the principal wish to induce the agent to make a positive effort $e = 1$. Thus, the new *moral hazard incentive constraint* can be written as:

$$\pi_1 u(\bar{t}) + (1 - \pi_1) u(\underline{t}) - e^\alpha \geq \pi_0 u(\bar{t}) + (1 - \pi_0) u(\underline{t}) \quad (19)$$

This is the incentive constraint, which implies that the agent prefers to exert a positive effort. However, the utility for the positive effort should be greater than that of the classical model to compensate for this new disutility.

Therefore, the principal continues to induce a positive effort, but there is a limit that decreases the desire to induce a greater effort than that of the classical model.

Proposition 1: If the principal induces a positive effort greater than the limit for the exponential disutility, his costs could be greater than the benefits that discourage the principal from taking the action.

Proposition 2: If the disutility α is very high, the principal cannot demand much effort considering the production level q , such as increases in q increases e , which restricts the principal's ability to require the agent to make an effort.

The agent's participation constraint is now re-written as:

$$\pi_1 u(\bar{t}) + (1 - \pi_1) u(\underline{t}) - e^\alpha \geq 0 \quad (20)$$

Therefore, the expected utility of the principal when he tries to induce effort by the agent is:

$$\pi_1 (\bar{S} - \bar{t}) + (1 - \pi_1) (\underline{S} - \underline{t}) \quad (21)$$

Since the participation constraint is binding, we also obtain the value of this transfer, which is sufficient to cover the disutility of effort, namely $t^{**} = e^\alpha$

$$V_1 = \pi_1 \bar{S} + (1 - \pi_1) \underline{S} - e^\alpha \quad (22)$$

Had the principal decided to let the agent exert no effort, $e = 0$, he would make a zero payment to the agent regardless of the agent's output. Therefore, the principal obtains payoff as follows:

$$V_0 = \pi_0 \bar{S} + (1 - \pi_0) \underline{S} \quad (23)$$

Inducing effort is thus optimal from the principal's point of view when $V_1 \geq V_0$, that is, $\pi_1 \bar{S} + (1 - \pi_1) \underline{S} - h(\psi) \geq \pi_0 \bar{S} + (1 - \pi_0) \underline{S}$, or to put it differently when:

$$\Delta \pi \Delta S \geq h(\psi) \quad (24)$$

where $\Delta S = \bar{S} - \underline{S}$

However, if α is sufficiently large, $V_1 \rightarrow V_0$. Therefore, there is a value of α for each q production level that discourages the agents who want to grow because the disutility is larger than the benefits.

The principal continues to induce effort and must choose the contract, which solves the following problem.

$$\pi_1 (\bar{S} - \bar{t}) + (1 - \pi_1) (\underline{S} - \underline{t}) \quad (25)$$

$$\pi_1 u(\bar{t}) + (1 - \pi_1) u(\underline{t}) - e^\alpha \geq \pi_0 u(\bar{t}) + (1 - \pi_0) u(\underline{t}) \quad (26)$$

$$\pi_1 u(\bar{t}) + (1 - \pi_1) u(\underline{t}) - e^\alpha \geq 0 \quad (27)$$

Therefore, the amount t for both agents is larger than that of the classical model to compensate for the marginal disutility of labor by the agents involved.

Moreover, for production level q , α tends to increase, thus forcing the principal to reformulate the optimal contract considering the new disutility.

Changing the terms, we have:

$$U_p = \text{Max } \pi_1 (\bar{S} - \bar{e}^\alpha) + (1 - \pi_1) (\underline{S} - \underline{e}^\alpha) \quad (28)$$

Extracting the first-order conditions from the principal's utility function, we have to

$$\underline{S}' = \alpha \underline{e}^{\alpha-1} \quad (29)$$

$$\bar{S}' = \alpha \bar{e}^{\alpha-1} \quad (30)$$

From the information obtained in the first-order conditions, it is possible to define $\alpha e^{\alpha-1} > e$ regardless of the level of effort of the contracted agent, which means that the principal has to disburse S , which is a larger amount than the classic model by [1].

3) NUMERICAL EXAMPLE AND RESULTS

This subsection describes the numerical case that demonstrates the impact of learning with labor in the classical model for unobservable costs and levels of effort.

α: NUMERICAL EXAMPLE FOR UNOBSERVABLE COSTS

Given that in the classical principal-agent model, θ grows linearly with the quantity produced, the principal draws up a contract of execution of n tasks, given $q = n$ for both agents.

Knowing that, the efficient agent will produce more than the inefficient agent, the principal distributes the contract menu: $\{(\$ 500, x \text{ units}); (\$ 300, x \text{ units})\}$.

The principal does not observe the agents' marginal cost, but the agents know their costs for the inefficient agent $\theta = \$ 10.00$, and for the efficient agent $\theta = \$ 5.00$. Therefore, we have the following payoffs as described in the Table 1:

TABLE 1. Payoff on the classical principal agent model.

Agent	$\underline{\theta}$ - Agent	$\bar{\theta}$ - Agent
Payoff	\$ 500 + \$ 100	\$ 300

The efficient agent accepts a contract up to $x = 100$ units produced, because its marginal cost of $\theta = \$ 5.00$, increases linearly with the quantity produced. The inefficient agent accepts a contract up to $x = 30$ units produced. Since the efficient agent can be considered inefficient, he can have an additional \$ 100 cash to keep him from concealing his real performance. However, with the marginal disutility of the work proposed by [24], keeping the values of t and θ constant, one should take into account the $\alpha > 1$, which for both agents in this model grows in the same way.

TABLE 2. Production level on the learning with labor to unobservable costs.

Agent	$\underline{\theta}$ - Agent				$\bar{\theta}$ - Agent			
	$\alpha = 1$	$\alpha = 1,5$	$\alpha = 2$	$\alpha = 2,5$	$\alpha = 1$	$\alpha = 1,5$	$\alpha = 2$	$\alpha = 2,5$
Payoffs	\$ 600	\$ 600	\$ 600	\$ 600	\$ 300	\$ 300	\$ 300	\$ 300
Level of Production	100	40.25	16	6.26	30	8.53	2.40	0.66
	units	units	units	units	units	units	units	units

Thus, we have the sensitivity analysis below, with increments of 50% in the value of α every 10 units produced for the efficient agent and three units produced for the inefficient agent in the normal production.

Moreover, increments of 50% must be implemented in the value of α every five units produced for the efficient agent and three units produced for the inefficient agent in the additional production. Table 2 presents the new payoffs.

Given that the marginal cost of the efficient agent increased by 50% after the first 10 units produced, it did not produce 90 units, as in the classical model. The new denominator is $5^{1.5} = 11.18$; because of this, 450 remaining divided by 11.18, which results in 40.25 units produced, a loss of 55.27%.

TABLE 3. Production level on the learning with labor for additional demands to unobservable costs.

Agent	$\underline{\theta}$ - Agent				$\bar{\theta}$ - Agent			
	$\beta = 1.5$	$\beta = 2$	$\beta = 2.5$	$\beta = 3$	$\beta = 1.5$	$\beta = 2$	$\beta = 2.5$	$\beta = 3$
	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$
Normal Payoff	\$750	\$750	\$750	\$750	\$300	\$300	\$300	\$300
Additional Payoff	\$50	\$50	\$50	\$50	\$30	\$30	\$30	\$30
Normal Production	100 units	40.25 units	16 units	6.26 units	30 units	8.53 units	2.4 units	0.66 units
Additional Production	3.06 units	1.11 units	0.39 units	0.13 units	0.72 units	0.18 units	0.04 units	0.01 units

For the principal to have the 90 units stipulated in the contract, he will have to pay larger amounts. With $\alpha = 2$, the efficient agent produces another 16 units for the same value of the initial contract with a loss of 80%. With $\alpha = 2.5$, the agent will produce another 6.26 unit with a loss of 91.05%.

The same occurs with the inefficient agent, he will only produce another 8.53 units if $\alpha = 1.5$; 2.4 units if $\alpha = 2$ and 0.66 units if $\alpha = 2.5$, with, respectively, percentage losses of: 68.40% if $\alpha = 1.5$; 90% if $\alpha = 2$ and 96.85% if $\alpha = 2.5$.

As there is still the possibility of additional demands, the principal draws up a contract menu for that possibility. However, β , which is the increasing marginal cost of the additional activity, is greater than α . Therefore, the principal will have to pay even more to get both agents to perform the task stipulated in the contract.

The contract menu given by the principal is $\{(\$ 50 \text{ per additional unit}); (\$ 30 \text{ per additional unit})\}$. However, marginal costs are higher for both the efficient agent $\theta = \$ 6.00$, and the inefficient agent $\theta = \$ 12.00$. In addition, the efficient agent has more information rent from the emergence of additional demands, as he can be inefficient in this case as well. Therefore, the principal disburses \$ 150 to inhibit this action. All marginal costs of the model are monetary terms that facilitate the demonstration. Therefore, we have:

Therefore, because the marginal cost of the efficient agent increased by 50% after the first five units produced in the additional time, it will not produce 7,5 units, as in the classical model. The new denominator is $6^{1.5} = 14.69$, resulting in these 45 remaining units divided by 14.69, resulting in 3,06 units produced and a loss of 59.2%. For the principal to have the 7,5 units stipulated in the contract, he will have to pay out larger amounts. With $\beta = 2$, the efficient agent will produce another 1.11 units for the same value of the initial contract with a loss of 72.25%.

With $\beta = 2, 5$, the agent will produce another 0,39 units with a loss of 88.85%. With $\beta = 3$, the efficient agent will produce another 0,13 units for the same value of the initial

contract with a loss of 99.56%. The same occurs with an inefficient agent. He will only produce another 0.72 units if $\beta = 1, 5$; 0.18 units if $\beta = 2$, and 0.04 units if $\beta = 2, 5$ and 0,01 units if $\beta = 3$, with, respectively, percentage losses of: 97,60% if $\beta = 1, 5$; 99.30% if $\beta = 2$; 99.83% if $\beta = 2.5$ and 99.95% if $\beta = 3$.

b: NUMERICAL EXAMPLE FOR EFFORT LEVEL

For the classical model with moral hazard, the agents exert an effort that is not observed by the principal. This effort can be represented by e , which is a parameter of disutility embedded in the utility function of the agents involved in an asymmetric information contract. However, the parameter e can vary according to the marginal disutility of labor [24].

Therefore, the principal distributes the same contract menu of adverse selection $\{(\$ 500, x \text{ units}); (\$ 300, x \text{ units})\}$ when both agents exert positive effort ($e = 1$). The effort produced by the agents is not observed by the principal, but the agents know their real efforts for the inefficient agent $e = 5$ and for the efficient agent $e = 3$. Therefore, we have the following payoffs, as listed in Table 4.

TABLE 4. Payoff on the classical principal agent model.

Agent	\underline{e} - Agent	\bar{e} - Agent
Payoff	\$ 500 + \$ 100	\$ 300

The efficient agent accepts a contract up to $x = 200$ units produced because of the unobserved effort of $e = 3$. The inefficient agent accepts a contract up to $x = 60$ units produced because of the unobserved effort of $e = 5$. Since the efficient agent can be considered inefficient, he can have an additional \$ 100 cash to keep him from concealing his real performance.

However, with the marginal disutility of the work proposed by [24], keeping the values of t and e constant, one should take into account the $\alpha > 1$, which for both agents in this model grows in the same way.

Thus, we have the sensitivity analysis below, with increments of 50% in the value of α every 10 units produced by the efficient agent and three units produced by the inefficient agent in the normal production. Table 5 presents the new payoffs.

Given that the effort of the efficient agent increased by 50% after the first 10 units produced, he will not produce 190 units as in the classical model, $3^{1.5} = 5, 19$, that is, 570 remaining divided by 5.19, which results in 109.82 units produced, a loss of 45.09%. For the principal to have 190 units stipulated in contract, he will have to pay larger amounts. With $\alpha = 2$, the efficient agent will produce another 60 units for the same value of the initial contract with a loss of 66.67%.

With $\alpha = 2.5$, the agent will produce 32.73 more units with a loss of 81.60%. The same occurs with an inefficient

TABLE 5. Production level on the learning with labor per effort level.

Agent	\underline{e} - Agent				\bar{e} - Agent			
	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$
Payoffs	\$600	\$600	\$600	\$600	\$300	\$300	\$300	\$300
Level of Production	200 units	109.82 units	60 units	32.73 units	60 units	25.49 units	10.8 units	0.91 units

agent. He will only produce another 25.49 units if $\alpha = 1.5$; 10.8 units if $\alpha = 2$ and 0.91 units if $\alpha = 2.5$, with, respectively, percentage losses of: 55.28% if $\alpha = 1.5$; 80% if $\alpha = 2$ and 98,21% if $\alpha = 2.5$.

III. LEARNING IN THE REPEATED GAME

In labor relations, the information asymmetries related to the performance of the hired individuals diminish as these contracts are put into practice, since the principal will have a greater understanding of the performance of each employee.

Thus, an analysis needs to be made of how the optimal contract proposed by [1] behaves with the updating of the contractor’s beliefs by viewing the agent’s performance.

Therefore, the players involved learn from repetition, and so different results were found, as seen in the model below:

A. LEARNING WITH THE UNDERSTANDING BY REPETITION OF THE GAME FOR UNOBSERVABLE COSTS

The results obtained by learning in the nonlinear preferences force the principal to disburse more to compensate for the disutility of agents involved in a contract. However, when this game is played infinitely, bearing in mind that all the players like a long-run player with the discount factor, the results may be different when the nonlinear model and the classical model are compared.

In a repetition of the game, the agent’s utilities are changed considering the gains and costs for all possibilities of time. Given that time is represented by t .

During each period, the principal and agent played a one-period game with a new random environment each time. In each period, each player’s actions depend on what he has observed up to that point in time [25].

For the principal, this is the history of his own previous actions (i.e., announced reward pairs) and the history of previous successes and failures. For the agent, this is the history of his own and the principal’s previous actions, the history of previous successes and failures, and the reward pair that the principal has just announced. Neither player observes random environments, which are assumed to be independent and

identically distributed. At the end of each period, after observing the current success or failure, the principal compensates the agent according to the reward pair that he announced at the beginning of the period. A supgame strategy for a player is a sequence of decision rules that determine an action at each period as a function of their information history at that point in time.

The supgame payoff for a player is the normalized sum of the discounted expected payoff [25].

When this game is played by $t = 1$, the results are presented in section 2. However, when $t > 1$, some hypotheses must be formulated to define the model.

When $t > 1$, the hypotheses are:

- The game is played infinitely, where for each period of time, the principal formulates a different contract with the possible agents.
- The contract is updated by the set of information made available by the principal for each period when he makes a decision.
- The principal and the agents are long-run players with discount factors, and both need to maximize their pay-offs for all games.
- The discount factor determines the type of each player in this game.
- The history defines the future decisions for all players.

1) TIMELINE

- At $t = 1$, the principal does not know about the agents’ performance and makes a decision using the probabilities for each type of agent.
- In $t > 1$, the principal increases his knowledge about the agent’s behavior and starts to understand how the agent’s utility function works. Because of this, he updated the probabilities for both agents using Bayes’ rule to decrease the possibility of asymmetric information.

2) THE STRUCTURE OF THE GAME

The principal needs to delegate an activity to someone and he does not know about the agents’ real performance, which is a problem that is found in the classical model [1].

However, the game was repeated. For example, the principal needs to contract someone to carry out a service in his company and in different periods of time, which is offered to different kinds of agents.

In this situation, the principal may analyze the first agent’s performance and make suppositions in accordance with the history regarding the probability distribution of the new agent’s performance.

As mentioned above, this game is played repeatedly, and therefore, the agents’ utility functions change to incorporate the possibility of the repeated game:

$$(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} (t - \theta^{\alpha} q) \tag{31}$$

$$(1 - \bar{\delta}) \sum_{t=1}^{\infty} \bar{\delta}^{t-1} (\bar{t} - \bar{\theta}^{\alpha} \bar{q}) \tag{32}$$

where:

- δ represents the discount factor of the efficient agent;
- $\bar{\delta}$ represents the discount factor of the inefficient agent;
- $\underline{\delta} > \bar{\delta}$, the discount factor for the efficient agent is greater than the discount factor for the inefficient agent because he is more patient than the inefficient player.

The agents need a minimal benefit to accept the principal's proposal, considering that any external opportunities are equal to zero. Thus, new participation constraints are added to those that already exist. Hence, knowing that the agent accepts the contract of normal units:

$$(1-\underline{\delta}) \sum_{t=1}^{\infty} \underline{\delta}^{t-1} (t - \underline{\theta}^{\alpha} \underline{q}) \geq 0 \quad (33)$$

$$(1-\bar{\delta}) \sum_{t=1}^{\infty} \bar{\delta}^{t-1} (\bar{t} - \bar{\theta}^{\alpha} \bar{q}) \geq 0 \quad (34)$$

Thus, the agents need to be discouraged from other than usual behavior; therefore, the following incentive compatibility constraints are present.

$$(1-\underline{\delta}) \sum_{t=1}^{\infty} \underline{\delta}^{t-1} (t - \underline{\theta}^{\alpha} \underline{q}) \geq (1-\underline{\delta}) \sum_{t=1}^{\infty} \underline{\delta}^{t-1} (t - \bar{\theta}^{\alpha} \bar{q}) \quad (35)$$

$$(1-\bar{\delta}) \sum_{t=1}^{\infty} \bar{\delta}^{t-1} (\bar{t} - \bar{\theta}^{\alpha} \bar{q}) \geq (1-\bar{\delta}) \sum_{t=1}^{\infty} \bar{\delta}^{t-1} (\bar{t} - \underline{\theta}^{\alpha} \underline{q}) \quad (36)$$

This structure is applied to avoid the possibility of agents mimicking a contract today and in the future.

Finally, the principal needs to maximize his utility by formulating a contract that is the most beneficial to him in a dynamic infinity game, that is, knowing that there is asymmetry of information and, consequently, resulting information rent for the active restriction agent, it is known that the utility of the principal is

$$\begin{aligned} \text{Max Up} = & p_{h'} \sum_{t=1}^{\infty} \delta_p^{t-1} (S(\underline{q}) - \underline{\theta}^{\alpha} \underline{q}) \\ & + (1 - p_{h'}) \sum_{t=1}^{\infty} \delta_p^{t-1} (S(\bar{q}) - \bar{\theta}^{\alpha} \bar{q}) \\ & + p_{h''} \sum_{t=1}^{\infty} \underline{\delta}^{t-1} (t - \underline{\theta}^{\alpha} \underline{q}) \\ & + (1 - p_{h''}) \sum_{t=1}^{\infty} \bar{\delta}^{t-1} (\bar{t} - \bar{\theta}^{\alpha} \bar{q}) \\ & + p_{h''} \sum_{t=1}^{\infty} \bar{\delta}^{t-1} \Delta \theta q \end{aligned} \quad (37)$$

where:

- $p_{h'}$ is the conditional probability obtained from the history observed by the principal. Thus, the probabilities are updated period by period in the utility function of the principal.
- $p_{h''}$ is the conditional probability obtained from the history observed by the agents. Thus, the probabilities are updated periodically in the utility function of the agents.

All the types of the agents observe the history in the same way.

- The relationship between the discount factors is: $\delta_p > \underline{\delta} > \bar{\delta}$.
- With $0 < \bar{\delta} < \underline{\delta} < \delta_p < 1$.

Given that the principal needs to maximize utility, the first-order condition must be defined. This is shown below:

$$S'(\underline{q}) = \frac{2 - \delta_p - \underline{\delta}}{1 - \underline{\delta}} \theta \quad (38)$$

$$\begin{aligned} S'(\bar{q}) &= \bar{\theta} + \frac{1 - \delta_p}{1 - p} \left(\frac{(1 - p)\theta}{1 - \delta} + \frac{p\Delta\theta}{1 - \delta_p} \right) \\ &= \frac{1 - \delta_p}{1 - \bar{\delta}} \theta + \frac{1 - \delta_p}{1 - \delta} \frac{p}{1 - p} \Delta\theta \end{aligned} \quad (39)$$

According to the relationship between discount factors, $\delta_p > \underline{\delta} > \bar{\delta}$. $1 - \delta_p < 1 - \bar{\delta}$ and $\frac{1 - \delta_p}{1 - \bar{\delta}} < 1$. Therefore, the player in the second period earns less than the player in the first period because of the learning and the decrease in asymmetric information. As $\frac{2 - \delta_p - \underline{\delta}}{1 - \underline{\delta}} \theta > 1$.

Thus, there is an income transfer of the lower performance player to the higher performance player and, consequently, to the principal.

3) DISCUSSION

Given the above, the histories in each period of time for each player may be represented by:

- h' - The history observed by the principal when he makes a decision: $h' \in H$, and H is the set of all the histories presented.
- h'' - The history observed by the agents when he makes a decision: $h'' \in H$, and H is the set of all the histories presented.

In the one-shot game principal-agent model, the probabilities of each agent's type are built *a priori* according to the environment. However, when this game is played infinitely, it is possible that the principal will try to review his beliefs and build, from the observation, probabilities *a posteriori* using Bayes' rule. In this case, probabilities were built based on observations of past histories.

Thus, h'_1 is the history observed by the principal at $t = 2$, and h'_2 is the history observed by the principal at $t = 3$. It is possible to find ph'_1 and $(1 - ph'_1)$ at $t = 2$.

When this is known, p is the probability of contracting the efficient type for $t = 1$. The Bayes rule may be used to define ph'_1 , given p .

4) LEARNING WITH THE UNDERSTANDING BY REPETITION OF THE GAME TO EFFORT LEVEL

According to the same hypotheses, for the repeated game for adverse selection, when $t > 1$, that is,

- The game is played infinitely, where for each period of time, the principal formulates a different contract with the possible agents.
- The contract is updated by the set of information made available by the principal for each period when he makes a decision.
- The principal and the agents are long-run players with discount factors, and both need to maximize their pay-offs for all games.
- The discount factor determines the type of each player in this game.

- The history defines the future decisions for all players.

The utilities of both agents for the moral hazard repeated game is:

$$(1 - \underline{\delta}) \sum_{t=1}^{\infty} \underline{\delta}^{t-1} (\pi_1 u(\bar{t}) + (1 - \pi_1) u(t) - e^\alpha) \quad (40)$$

$$(1 - \bar{\delta}) \sum_{t=1}^{\infty} \bar{\delta}^{t-1} (\pi_1 u(\bar{t}) + (1 - \pi_1) u(t) - e^\alpha) \quad (41)$$

where:

- $\underline{\delta}$ represents the discount factor of the efficient agent;
- $\bar{\delta}$ represents the discount factor of the inefficient agent;
- $\underline{\delta} > \bar{\delta}$, the discount factor for the efficient agent is greater than the discount factor for the inefficient agent because he is more patient than the inefficient player.

The principal needs to maximize his utility by formulating a contract that is the most beneficial to him in a dynamic infinity game, that is, knowing that there is asymmetry of information and, consequently, resulting information rent for the active restriction agent, it is known that the utility of the principal is

Max Up

$$\begin{aligned} &= p_{h'} \sum_{t=1}^{\infty} \delta_p^{t-1} ((\pi_1 u(t) - e^\alpha + (1 - \pi_1) u(\bar{t}) - e^\alpha) \\ &+ (1 - p_{h'}) \sum_{t=1}^{\infty} \delta_p^{t-1} ((\pi_1 u(t) - e^\alpha + (1 - \pi_1) u(\bar{t}) \\ &- e^\alpha + p_{h''} \sum_{t=1}^{\infty} \underline{\delta}^{t-1} ((\pi_1 u(t) - e^\alpha + (1 - \pi_1) u(\bar{t}) - e^\alpha) \\ &+ (1 - p_{h''}) \sum_{t=1}^{\infty} \underline{\delta}^{t-1} ((\pi_1 u(t) - e^\alpha + (1 - \pi_1) u(\bar{t}) - e^\alpha)) \end{aligned} \quad (42)$$

where:

- $p_{h'}$ is the conditional probability obtained from the history observed by the principal. Thus, the probabilities are updated period by period in the utility function of the principal.
- $p_{h''}$ is the conditional probability obtained from the history observed by the agents. Thus, the probabilities are updated periodically in the utility function of the agents.

All the types of agents observe the history in the same way.

- The relationship between the discount factors is: $\delta_p > \underline{\delta} > \bar{\delta}$.
- With $0 < \bar{\delta} < \underline{\delta} < \delta_p < 1$

Given that the principal needs to maximize utility, the first-order conditions must be defined. These are shown below:

$$u'(\underline{q}) = \frac{2 - \delta_p - \underline{\delta}}{1 - \underline{\delta}} e \quad (43)$$

$$\begin{aligned} u'(\underline{q}) &= e + \frac{1 - \delta_p}{1 - p} \left(\frac{(1 - p) e}{1 - \delta} + \frac{p \Delta e}{1 - \delta_p} \right) \\ &= \frac{1 - \delta_p}{1 - \underline{\delta}} e + \frac{1 - \delta_p}{1 - \delta} \frac{p}{1 - p} \Delta e \end{aligned} \quad (44)$$

5) NUMERICAL EXAMPLE FOR THE LEARNING BY REPETITION

This subsection describes the numerical case that demonstrates the impact of learning with labor incorporated into the

repetition and decrease in the asymmetric information in the classical model for unobservable costs and effort levels.

a: NUMERICAL EXAMPLE FOR THE LEARNING BY THE REPETITION USING UNOBSERVABLE COSTS

Using the same example explained in subsection, we have the following answers about the quantity produced when the decrease in asymmetric information by learning in the repeated game is considered. The relationship between the discount factor of the agents is $\delta_p > \underline{\delta} > \bar{\delta}$. Table 6 describes this new scenario for $\delta = 0.8$ to the principal.

TABLE 6. Production level on the learning by repetition using unobservable costs.

Agent	$\underline{\delta}$ -Agent				$\bar{\delta}$ -Agent			
	$\bar{\delta} = 0.6$				$\bar{\delta} = 0.2$			
	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$
Payoffs	\$700	\$700	\$700	\$700	\$300	\$300	\$300	\$300
Level of Production on $t = 1$	100 units	40.25 units	16 units	6.26 units	30 units	8.53 units	2.4 units	0.66 units
Level of Production on $t > 1$	150 units	60.37 units	24 units	9.39 units	112.5 units	91.32 units	75 units	61.39 units

To describe the numerical case of learning by repetition using unobservable costs, some hypotheses need to be presented.

1. The probability of agents appearing with high and low performance (p and $1-p$) is the same.
2. The costs of doing a certain activity are assumed by the optimal contract proposed in (38) and (39).

Given that the marginal cost of the efficient agent increased by 50% after the first 10 units were produced, and there is a difference between his discount factor $\underline{\delta} = 0.6$ and $\delta_p = 0.8$, namely, 0.2, the production levels can be analyzed using the optimal contract proposed as follows: $S'(\underline{q}) = \frac{2 - \delta_p - \underline{\delta}}{1 - \underline{\delta}} \theta^\alpha$.

So, $S'(\underline{q}) = 1, 5\theta^\alpha$. Therefore, there is a gain in production when the repeated form is imposed, owing to the principal decrease in asymmetric information. Therefore, this will not produce 100 units, as in the classical model.

The new operation is the product of the $S'(\underline{q}) = 1, 5\theta^\alpha$ and the ratio between monetary gains and associated costs for each agent in determining the task. $500/5^1 = 100$ units. As a result, 150 units can be produced, thereby showing the advantages of the principal in a play to repeated games, given that there is an increment of 50% of production for the efficient agent. Thus, the principal can reduce the initial proposal to reach his goal.

However, the learning with labor that was introduced into our model is presented and generates different results; for example, if $\alpha = 1,5$, the efficient agent will produce another 60.37 units for the same value as the initial contract

when compared with the nonlinear preferences by another 40.25 units with a growth of 49.98%. Moreover, there was a decrease of 32.92% when compared with the classical model. If $\alpha = 2$, the high agent will produce another 24 units for the same value as the initial contract when compared with the nonlinear preferences by 16 more units with a growth of 50%.

In addition, there was a 70% decrease compared to the classical model. If $\alpha = 2, 5$, the agent will produce another 9.39 units for the same value as the initial contract when compared with the nonlinear preferences by another 6.26 units with a growth of 50%. In addition, there was a decrease of 86.58% compared to the classical model.

This also occurs with the inefficient agent, where the new operation uses $\frac{1-\delta_p}{1-\delta} \theta + \frac{1-\delta_p}{1-\delta} \frac{p}{1-p} \Delta \theta$, taking into account the hypothesis of the same probability of occurrence for both agents (p and $1-p$). So, $S'(\underline{q}) = 0.25.\underline{\theta} + 0.25.\Delta\underline{\theta}$.

Therefore, the new operation is the product of product $S'(\underline{q}) = 0.25.\underline{\theta} + 0.25.\Delta\underline{\theta}$ and the ratio between monetary gains and associated costs for each agent in determining the task was $300/10^1 = 30$ units. Therefore, this will produce another 112.5 units if $\alpha = 1$; 91.32 units if $\alpha = 1.5$, 75 if $\alpha = 2$ and 91.39 units if $\alpha = 2.5$ showing, respectively, a percentage growth of: 275% if $\alpha = 1$; 970.57% if $\alpha = 1.5$; 3025% if $\alpha = 2$ and 13746% if $\alpha = 2.5$. This growth is high because of the large difference between the discount factors of each player.

b: NUMERICAL EXAMPLE FOR LEARNING BY REPETITION USING EFFORT LEVEL

Using the same example explained in Subsection A.3.2, we also have the following answers about the quantity produced when decreasing asymmetric information is considered by learning in the repeated game. Table 7 describes this new scenario for $\delta = 0, 8$ for the principal.

TABLE 7. Production level on learning by repetition using effort level.

Agent	$\underline{\theta}$ -Agent				$\bar{\theta}$ -Agent			
	$\delta = 0.6$				$\delta = 0.2$			
	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$
Payoffs	\$700	\$700	\$700	\$700	\$300	\$300	\$300	\$300
Level of Production on $t = 1$	200 units	109.82 units	60 units	32.73 units	60 units	25.49 units	10.8 units	0.91 units
Level of Production on $t > 1$	300 units	164.73 units	90 units	49.09 units	75 units	53.66 units	39 units	28.48 units

To describe the numerical case of learning by repetition using the effort level, some hypotheses need to be presented.

1. The probability of agents appearing with high and low performances (p and $1-p$) is the same;
2. The costs of doing a certain activity are supposed by the optimal contract proposed in proposed in (43) and (44).

Using the same operation rules for the Unobservable Costs case and que structure for (43) and (44), it can be seen that:

Learning with labor that was introduced into our model with effort level is presented and generates different results. For example, if $\alpha = 1.5, \alpha = 2$, and $\alpha = 2.5$, the efficient agent will produce 50% more than the nonlinear model.

The inefficient agent will produce 25%, 110%, 261%, and 3029% more units, respectively, if $\alpha = 1, \alpha = 1.5, \alpha = 2$, and $\alpha = 2.5$.

IV. LEARNING IN THE ECONOMIC CYCLES OF EMPLOYABILITY

The cases shown above maintained the hypothesis that the actors involved did not suffer any influence from the economic scenario. When this hypothesis is relaxed, we have the so-called employability cycle.

Therefore, this game happens in an economic frame in which, in each period, there may be two possibilities. These economic scenarios are called employability cycles.

The first is an *economic expansion*, in which there are plenty of job offers, and it is easier for any agent to find jobs in one period and to be free to look for other possibilities in the next period. In this scenario, the bargaining power of the agents is greater than in other situations. These factors are *more critical* to the conditions and values offered by employers.

The other possibility is an *economic recession*, during which job vacancies drop to a few offers, and the probability of unemployment is higher.

In this case, bargaining power starts to decrease, and agents tend to *accept* the contracts offered by employers *faster* than normal.

The changes in the scenario were random and had unknown distributions. To describe this phenomenon in a specific period, we can use a Markov chain.

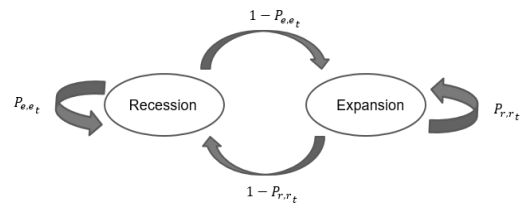


FIGURE 2. The cycles of employability.

In these two possibilities, some characteristics can be found which are described below:

- The principal evaluates the actual scenario and contracts the agents.
- The agents evaluate the scenario and accept or refuse the proposal.

To understand how scenarios change, our model uses a finite non-stationary Markov chain that is described by a sequence of transition matrices defined in a common state space (in this study, two states, recession and expansion). In period k , the system moves from state i to j with probability $\Pi(k)_{ij}$. Therefore, the probability $\Pi(t)_{ij}$ changes over time

during each period. Hence, the scenarios of this game can be represented by the transition matrix as follows:

$$\Pi_t = [P_{e,e_t} P_{e,r_t} P_{r,e_t} P_{r,r_t}] = [P_{e,e_t} 1 - P_{e,e_t} 1 - P_{r,r_t} P_{r,r_t}] \quad (45)$$

Therefore, players need to understand the scenarios and update their probabilities of accepting or refusing the formulation of the contract with these observations. To this end, we used the Dirichlet distribution in each row of the transition matrix. The use of this specific type of beta distribution is due to the possibility of changing the probabilities with each observation of the players, which leads to learning for the hiring logic.

The Dirichlet distributions for the first and second rows of the transition matrix were analyzed as follows:

$$D_{p_t} = \frac{\Gamma(a_1 + a_2)}{\Gamma(a_1) \Gamma(a_2)} P_{e,e_t}^{\alpha_1 - 1} (1 - P_{e,e_t})^{\alpha_2 - 1} \quad (46)$$

$$D_{p_t} = \frac{\Gamma(a_1 + a_2)}{\Gamma(a_1) \Gamma(a_2)} (1 - P_{r,r_t})^{\alpha_1 - 1} P_{r,r_t}^{\alpha_2 - 1} \quad (47)$$

Every period when we update the values of the probabilities of transitions, new information is generated about how likely periods of expansion and recession will be in the future. The expectation about the long-run probabilities of each state will be used to parameterize the discount factor.

From now on, the discount factor for all players is $\delta_i(\pi_e)_t$, thus reflecting the beliefs of player i in period t and their influence on the rate of impatience through successive observations.

Thus, the discount factor $\delta_i(\pi_e)_t$ is a variable that reflects the observation of the transition of scenarios of economic expansion and recession from a non-stationary Markov chain that updates the information of the players involved in the contractual relationship by using the Dirichlet distribution.

How does this expectation change the discount factor?

Two hypothesis are made:

1. Both types of agents benefit from a heated economy. Thus, if any agent deduces that by his observation in any given period that an expanding economy will frequently happen, his patience rate (discount factor) will increase, and both will be more critical.

2. The principal, on the other hand, can benefit from dealing with more brash agents as they can offer lower salaries due to their bad expectations of job vacancies in an economy in recession.

3. If an agent does not accept the contract, he or she will find a temporary job for one period.

This can be written as:

$$\frac{d\delta_p}{dp_e} < 0; \quad \frac{d\delta}{dp_e} > 0; \quad \frac{d\bar{\delta}}{dp_e} > 0 \quad (48)$$

The results above show that expansion periods are beneficial for the agents and prejudicial for the principal, given that the first-order conditions are positive for the employees (agents) and negative for the owner (principal).

In contrast, in a period of recession, the relationship between them is completely different, which is better for the owner (principal).

Proposition 1: Suppose that in a specific period τ , the discount factors are $\delta_p(\pi_e)_\tau$, $\underline{\delta}(\pi_e)_\tau$, and $\bar{\delta}(\pi_e)_\tau$, and every decision made in this period will use this discount factor to calculate its utility in the long run.

Therefore, if $t = \tau$, the highest and lowest utilities of agents can be described as follows:

$$\sum_1^T \underline{\delta}(\tau)^{t-1} (\underline{t} - \underline{\theta}\underline{q}) = \frac{1 - \underline{\delta}(\tau)^T}{1 - \underline{\delta}(\tau)} (\underline{t} - \underline{\theta}\underline{q}) \quad (49)$$

$$\sum_1^T \bar{\delta}(\tau)^{t-1} (\bar{t} - \bar{\theta}\bar{q}) = \frac{1 - \bar{\delta}(\tau)^T}{1 - \bar{\delta}(\tau)} (\bar{t} - \bar{\theta}\bar{q}) \quad (50)$$

Moreover, the principal maximizes his utility by the structure:

$$\text{Max } p \sum_1^T \delta_p(\tau)^{t-1} (S(\underline{q}) - \underline{\theta}\underline{q}) + (1 - p) \times \sum_1^T \delta_p(\tau)^{t-1} (S(\bar{q}) - \bar{\theta}\bar{q}) - p \sum_1^T \Delta\theta\bar{q} \quad (51)$$

Taking into consideration the following constraints:

- Participation Constraints

$$\sum_1^T \underline{\delta}(\tau)^{t-1} (\underline{t} - \underline{\theta}\underline{q}) = \frac{1 - \underline{\delta}(\tau)^T}{1 - \underline{\delta}(\tau)} (\underline{t} - \underline{\theta}\underline{q}) \geq 0 \quad (52)$$

$$\sum_1^T \bar{\delta}(\tau)^{t-1} (\bar{t} - \bar{\theta}\bar{q}) = \frac{1 - \bar{\delta}(\tau)^T}{1 - \bar{\delta}(\tau)} (\bar{t} - \bar{\theta}\bar{q}) \geq 0 \quad (53)$$

- Constraints on the Compatibility of Incentives

$$\sum_1^T \underline{\delta}(\tau)^{t-1} (\underline{t} - \underline{\theta}\underline{q}) \geq \sum_1^T \bar{\delta}(\tau)^{t-1} (\bar{t} - \underline{\theta}\bar{q}) \quad (54)$$

$$\sum_1^T \bar{\delta}(\tau)^{t-1} (\bar{t} - \bar{\theta}\bar{q}) \geq \sum_1^T \underline{\delta}(\tau)^{t-1} (\underline{t} - \bar{\theta}\bar{q}) \quad (55)$$

Maximizing the principal's utility, we rewrite the second best formulated by [1] as follows:

$$S'(\underline{q}) = \underline{\theta} \quad (56)$$

$$S'(\bar{q}) = \bar{\theta} + \frac{p}{1-p} \frac{(1 - \delta_p)(1 - \delta^T)}{(1 - \underline{\delta})(1 - \delta_p^T)} \Delta\theta \quad (57)$$

We can conclude that this payment to the high agent is lower than in the classical model for both scenarios, from which it can be concluded that learning with the changing scenarios of employability modifies the perception of the agents involved in the game.

$$\underline{\theta} + \frac{p}{1-p} \frac{(1 - \delta_p)(1 - \delta^T)}{(1 - \underline{\delta})(1 - \delta_p^T)} \Delta\theta < \bar{\theta} + \frac{p}{1-p} \Delta\theta \quad (58)$$

The condition to accept the contract offered by the principal is described below, namely, if for a specific period τ , the utility in the long run obtained by observing a random scenario of the high-performance agent is lower than the gains obtained in the optimal contract, the contract is accepted.

$$\sum_1^T \underline{\delta}(p_{e_t})^{t-1} (\underline{g} - \underline{\theta}\underline{q}) < \frac{1 - \underline{\delta}(\tau)^T}{1 - \underline{\delta}(\tau)} (\underline{g} - \underline{\theta}\underline{q}) \quad (59)$$

The decision-making problem for the low performance agent is analogous.

V. CONCLUSION

Understanding the context in which the player is inserted has always been paramount for behavior in decision situations. In the logic of hiring and drafting optimal job contracts, classical models offer a preference structure of the agents involved in the game and advocate the importance of analyzing behavioral changes revealed by changes in these preference structures.

However, the preference structure, that is, each player's utility function, can be changeable and can behave according to each player's perceptions. Therefore, drawing up optimal contracts without considering the participants' learning in the most diverse areas, as shown in this paper, becomes static and is not adaptable to the real and dynamic context that reflects the economic hiring scenario.

Therefore, by using the principal agent model of [1], changes in the optimality of the contract with increases and decreases in the values to be offered by the contractor can be verified.

In addition, there is the possibility of not contracting by simply understanding the current economic scenario, which shows that learning is continuous and that makes the formulation of static preference structures not adaptable to the current situation in the job market. Finally, an individual's preferences are functions of his/her continuous learning about the scenario, about the other agents, and about himself/herself regarding the object to be studied.

Thus, the main contributions of this paper are to discuss the need to verify the degree of learning of all agents involved in the changes and influences of the environment in which they are inserted in the formulation of employment contracts and demonstrate the changes in the utility structures of each agent involved in understanding the environment in which the game is inserted and, consequently, the change in the cost-benefit ratio and the different optimal offers according to each scenario and degree of learning of the players.

This is a real problem, approached theoretically, whose results were simulated, but subsequent work may seek its application.

This research suggests as possible future analyses, the investigation of different degrees of perception and learning by each player, as it is notorious that the players learn, but in different ways and speeds.

REFERENCES

- [1] J. Laffont and D. Martimort, *The Theory of Incentives: The Principal-Agent Model*. Princeton, NJ, USA: Princeton Univ. Press, 2001.
- [2] H. Bejarano, B. Corgnet, and J. Gómez-Minãmbres, *Labor Contracts, Gift-Exchange and Reference Wages: Your Gift Need Not Be Mine*. Saint Etienne, France: Hal Archives-Ouvertes, 2019.
- [3] B. H. Batalgi, Y. Deng, and X. Ma, "Network effects on labor contracts of internal migrants in China—A spatial autoregressive model," in *Center for Policy Research*, vol. 242. Syracuse, NY, USA: Syracuse Univ., 2017.
- [4] M. Vasconcelos, C. Cardonha, and B. Gonçalves, "Modeling epistemological principles for bias mitigation in AI systems: An illustration in hiring decisions," in *Proc. AAAI/ACM Conf. AI, Ethics, Soc.*, Dec. 2018, pp. 323–329.
- [5] T. Feng, S. Tai, C. Sun, and Q. Man, "Study on cooperative mechanism of prefabricated producers based on evolutionary game theory," *Math. Problems Eng.*, vol. 2017, pp. 1–6, Jul. 2017.

- [6] M. A. Shattal, A. Wisniewska, A. Al-Fuqaha, B. Khan, and K. Dombrowski, "Evolutionary game theory perspective on dynamic spectrum access etiquette," *IEEE Access*, vol. 6, pp. 13142–13157, 2017.
- [7] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic spectrum access in unknown dynamic environment: A game-theoretic stochastic learning solution," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1380–1391, Apr. 2012.
- [8] Y. Xiao, J. Park, and M. V. D. Schaar, "Repeated games with intervention: Theory and applications in communications," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 3123–3132, Oct. 2012.
- [9] Y. Xu, A. Anpalagan, Q. Wu, L. Shen, Z. Gao, and J. Wang, "Decision-theoretic distributed channel selection for opportunistic spectrum access: Strategies, challenges and solutions," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1689–1713, 4th Quart., 2013.
- [10] N. R. Kocherlakota, "Implications of efficient risk sharing without commitment," *Rev. Econ. Stud.*, vol. 63, no. 4, pp. 595–609, Oct. 1996.
- [11] C. Phelan, "Repeated moral hazard and one-sided commitment," *J. Econ. Theory*, vol. 66, no. 2, pp. 488–506, Aug. 1995.
- [12] J. Thomas and T. Worrall, "Self-enforcing wage contracts," *Rev. Econ. Stud.*, vol. 55, no. 4, pp. 541–554, 1998.
- [13] B. Corgnet, J. Gomes-Minãmbres, and R. Hernan-Gonzales, "Goal setting in the principal-agent model: Weak incentives for strong performance," *Games Econ. Behav.*, vol. 109, pp. 311–326, May 2018.
- [14] S. Pouryousefi and J. Frooman, "The problem of unilateralism in agency theory: Towards a bilateral formulation," *Bus. Ethics Quart.*, vol. 27, no. 2, pp. 163–182, Apr. 2017.
- [15] T. Daske, "Pooling hawks and doves: Interim-efficient labor contracts for other-regarding agents," ZBW—Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel, Germany, Tech. Rep., 2016.
- [16] B. Kiviat, "The art of deciding with data: Evidence from how employers translate credit reports into hiring decisions," *Socio-Econ. Rev.*, vol. 17, pp. 283–309, Aug. 2017.
- [17] C. Chapman, *Project Risk Management*. Hoboken, NJ, USA: Wiley, 2003.
- [18] J. E. Garen, "Executive compensation and principal-agent theory," *J. Political Economy*, vol. 102, no. 6, pp. 1175–1199, Dec. 1994.
- [19] J. J. Laffont and J. Tirole, *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA, USA: MIT Press, 1993.
- [20] D. E. M. Sappington, "Incentives in principal-agent relationships," *J. Econ. Perspect.*, vol. 5, no. 2, pp. 45–66, May 1991.
- [21] L. K. Stroh, J. M. Brett, J. P. Baumann, and A. H. Reilly, "Agency theory and variable pay compensation strategies," *Acad. Manage. J.*, vol. 39, no. 3, pp. 751–767, Jun. 1996.
- [22] E. Fehr and K. M. Schmidt, "Fairness and incentives in a multi-task principal agent model," *Scand. J. Econ.*, vol. 106, pp. 453–474, 2004.
- [23] H. Zhang and S. Zenios, "A dynamic principal-agent model with hidden information: Sequential optimality through truthful state revelation," *Oper. Res.*, vol. 56, no. 3, pp. 681–696, Jun. 2008.
- [24] W. S. Jevons, *The Theory of Political Economy*, 5th ed. New York, NY, USA: Cosimo Classics, 1957.
- [25] R. Radner, "Repeated principal-agent games with discounting," *Econometrica: J. Econ. Soc.*, vol. 53, pp. 1173–1198, Sep. 1985.



JOSE IRANILDO SALES received the B.Sc. degree in management from the University of Pernambuco (UPE), in 2013, and the M.Sc. degree in production engineering from the Federal University of Pernambuco of Brazil (UFPE), in 2016, where he is currently studying the Ph.D. degree in production engineering.



FRANCISCO DE SOUSA RAMOS received the D.Sc. degree in economics from the Université Catholique de Louvain, Belgium, in 1992. He was a Researcher with the National Council of Research-CNPq and LabGRC, UFPE, the Dean of the Research and Postgraduate Studies with UFPE, from 2011 to 2015, and the Research Director with UFPE, from 2008 to 2011, where he is currently a Professor of Postgraduate Program in production engineer with the Department of Economics.

...