Wage Dynamics: The Role of Learning, Human Capital, and Performance Incentives*

(Preliminary)

Braz Camargo† Fabian Lange‡ Elena Pastorino§

Abstract

Based on public individual survey data and proprietary firm personnel records, we document that the importance of performance pay declines at the end of a career. This evidence is at odds with the implication of models of implicit and explicit performance incentives that have attempted to explain the life-cycle profile of the variable component of wages. We provide a novel model that integrates uncertainty and learning about individual ability, human capital acquisition, and performance incentives to account for the life-cycle profile of wages and the declining importance of the pay-for-performance component of wages with experience. We analytically characterize the equilibrium wage contract in this environment, derive its qualitative properties, and decompose the pay-for-performance component of the equilibrium wage into four distinct terms. These terms capture the standard trade-off between risk and incentives arising in situations of moral hazard, the desire to hedge against the risk inherent in learning about ability, and the changing strength of reputational and human capital investment motives over time. We prove that the model is identified just based on a panel of wages. We derive estimators of the model's primitive parameters and use the estimated parameters to measure the relative contribution of the three sources of wage dynamics we nest—learning, human capital acquisition, and performance incentives—to the life-cycle profile of total wages and their variable component. According to our preliminary estimates, performance incentives seem to be qualitatively and quantitatively key to explaining the life-cycle profile of wages and variable pay, despite the declining importance of variable pay relative to fixed pay over time.

Keywords: Human Capital Acquisition, Learning, Performance Incentives, Wage Growth, Identification, Structural Estimation

*We thank Richard Blundell, Costas Meghir, Bob Miller, and Gabriel Ulyssea for their comments. Braz Camargo gratefully acknowledges financial support from CNPq.

†Sao Paulo School of Economics. E-mail: braz.camargo@fgv.br

‡Corresponding Author. McGill University, Department of Economics. Address: Leacock Building, 855 Sherbrooke Street West, Montreal Quebec H3A2T7, Canada. E-mail: fabolange@gmail.com.

§Stanford University, Hoover Institution, Department of Economics, and SIEPR. E-mail: epastori@stanford.edu

1 Introduction

Two large literatures in labor and personnel economics center on investment motives to explain the life-cycle profile of individual earnings. On the one hand, the *human capital* model ([1, 2, 10]) focuses on investments in productive ability as a key determinant of how individual earnings vary with experience in the labor market. On the other hand, the 'career concerns' model ([8, 5]) considers environments in which individual ability is unknown, and addresses the question of how the investment in the market's perception of individual ability or one's 'reputation' through effort on the job, which in a competitive market influences wages, affects the problem of providing explicit incentives for performance. Typically, the object of interest in this literature is the life-cycle profile of the pay–for–performance component of compensation, that is, the variable component of wages.

Since the investment motive is at the heart of both models, it is natural to expect that human capital and career concerns considerations interact in their impact on life-cycle earnings. Indeed, it has long been argued that the life-cycle profile of earnings can be ascribed to a combination of uncertainty and learning about individual ability, the acquisition of new productive skills, and dynamic incentives for performance, that is, a combination of the two models. Yet, little to date is known about the relative importance of these three sources of life-cycle wage dynamics. Models that have integrated learning and performance incentives (see [5]) have been successful at explaining the use of seemingly 'simple' wage contracts, in the face of complex incentive problem. These models are based on the intuition that explicit incentives from performance pay should become more and more important over time, as implicit incentives from concerns about one's future career weaken—ability is eventually learned and the working horizon shortens. It turns out, however, that this prediction is highly counterfactual. As we document, the sensitivity of pay to performance, which measures the strength of explicit incentives in these models, is highly variable over the life cycle and often declining at the end of it, precisely when the need to substitute explicit incentives for implicit incentives though performance pay should be strongest, according to these models.

In this paper, we analyze how career concerns and human capital investment motives interact in shaping overall wages, their fixed and variable components, and explicit incentives for performance over the life cycle. We revisit the importance of incentives for the dynamics of individual wages in a model that nests the three main sources of like-cycle earnings growth discussed above: uncertainty and learning about individual ability, human capital acquisition, and dynamic incentives for performance. We characterize the optimal wage contract in this framework and provide a recursive characterization of the pay-for-performance component of pay that allows us to decompose it into four distinct elements, which can be readily interpreted in terms of basic life-cycle considerations. Specifically, these elements capture: 1) the standard trade-off between risk and incentives that emerges in the presence of moral hazard when individuals are risk averse; 2) the dynamic hedging of the risk inherent in learning about ability, when ability is unknown; 3) concerns about the market's perception of an individual's ability; and 4) the magnitude of human capital acquisition in the form of learning-by-doing. We show that the model's primitives can be recovered from cross-sections of wages and experience only, whether from commonly used panel income surveys or firm-level personnel records, without requiring specific information on the pay-for-performance component of wages or measures of individual investments or effort over time.[1]

We estimate the model based on estimators derived from these constructive identification arguments. Our preliminary estimation results imply that learning about ability, human capital acquisition, and performance incentives are all key determinants of the profile of wages, the varying importance of performance pay, and overall wage growth over the life cycle. In particular, failing to account for the role of human capital acquisition leads to underestimating the importance of performance incentives for wages since it obscures the *declining* impact of human capital acquisition incentives on wages.

Formally, we consider a competitive market in which firms compete for individuals of unknown ability, who exert effort on the job and acquire human capital when employed depending on the effort they supply. We characterize recursively the optimal employment contract for this environment, derive analytically its key features, and decompose the pay-for-performance component of the optimal contract into four terms. Each of these terms informs on a different aspect of

[1]Of course such data would be immensely helpful when empirically implementing the model. However, as we show in the paper, it is not strictly necessary to identify the model. This demonstrates that the model as empirical content and can potentially be estimated in a wide variety of contexts.

the dynamic considerations that determine the life–cycle provision of incentives. Three of these terms reflect incentives already familiar from the literature on dynamic moral hazard and career concerns: the standard trade-off between risk and incentives already present in [7], the desire to hedge against risk inherent in learning about ability, and the changing strength of the career concerns motive over time. These first three components all lead to pay-for-performance that increases over the life cycle (see [5]).

Novel is a fourth term that reflects learning-by-doing. In particular, in a world with incomplete information, part of the marginal social returns to learning-by-doing accrues to the employer. At the margin, individuals may therefore under-invest in human capital as they do not reap the full social return from investing. The optimal contract reflects this wedge between social and private returns to the investment in human capital: this wedge tends to be largest early in an individual's life-cycle and declines as individuals age and the investment horizon shortens. A consequence of this feature of the optimal contract is that in the presence of learning-by-doing, the piece rate of the optimal wage contract, and so the responsiveness of pay to performance, will tend to fall—or increase less than in the absence of human capital acquisition—as individuals approach the end of their careers. For an intuition, observe that if learning-by-doing is present but ability is known, then piece rates decline over time. The reason for this decline is that the first–best level of effort is decreasing over time, for natural reasons, and the equilibrium piece-rates replicate this feature. Now, when an individual's life-cycle horizon is sufficiently large relative to the signal-to-noise ratio, career concerns effectively disappear and the behavior of piece-rates is governed just by the incentives provided by the human capital accumulation process, which, as just argued, implies declining piece-rates.

We next show that the model's primitives are identified from either a long enough panel on wages or repeated information on the fixed and variable component of wages, in addition to standard information on individuals' wages and experience. Intuitively, either type of data allows to recover the experience profile of the variable component of pay and so the sensitivity of pay to performance. Since the model implies an explicit known mapping between specific features of this profile and the primitive parameters of the model, up to the rate of time preference and the marginal cost of effort—as well as a scale normalization, for short panels—all primitive parameters can be recovered from the experience profile of the sensitivity of pay to performance.

Empirically, we investigate the incentive power of pay-for-performance over individuals' careers and estimate the model's primitive parameters using information from the Panel Study of Income Dynamics (PSID) and personnel data from several firms, which provide information on individuals' wages and experience over the life-cycle. These empirical patterns are instructive about how incentives evolve over the life-cycle and the sources of life-cycle wage growth we nest. (For related work on the identification and estimation of models of static and dynamic moral hazard and careers concerns, see [9, 3, 4, 6].)

The paper proceeds as follows. Section 2 presents our model. Section 3 lays the ground for our equilibrium characterization by analyzing the updating problem of firms and how incentives for effort depend on career concerns and learning-by-doing. Section 4 establishes our main theoretical results. We first show that the equilibrium is unique and provide a recursive formulation of the optimal contract and the decomposition of the optimal piece-rate. We then use these our equilibrium characterization to establish a number of results about the life-cycle profile of incentives. Section 5 discusses idenfication. Section 6 describes our data. Section 7 presents our (preliminary) estimation results. Section 8 concludes.

2 The Model

We first describe the environment and then define equilibria. We conclude with a few remarks about our model.

2.1 Environment

Time is discrete and ranges from $t = 0$ to $t = T$. The real time between two consecutive periods is $\Delta > 0$.[2] While our equilibrium characterization holds for any value of Δ, some results we establish about the life-cycle behavior of equilibrium contracts hold only for small enough Δ. There is a finite set of risk-averse workers and a finite set of risk-neutral firms that Bertrand-

[2]Thus, T is a multiple of Δ.

compete to employ the workers.[3] Firms have a constant returns to scale technology, and so can employ as many workers as they want. In what follows, we let $\mathcal{J}$ denote the set of workers, $\mathcal{S} = \{n\Delta : n \in \{0, \ldots, T/\Delta\}\}$ denote the set of points in time at which action takes place, and use $x(t)$ to denote the value of a variable x at time t.

Production. The output of worker $i \in \mathcal{J}$ at time $t \in \mathcal{S}$ is

$$y_i(t) = (e_i(t) + k_i(t) + \theta_i)\,\Delta + \varepsilon_i(t)\sqrt{\Delta}, \tag{1}$$

where $e_i(t)$ and $k_i(t)$ are, respectively, the worker's non-negative effort and stock of human capital at t and θ_i is worker i's time-invariant ability.[4] The term $\varepsilon_i(t)$ is an idiosyncratic noise, or risk, term. The worker's choice of effort is private and his ability is unknown to both him and the other agents. We assume that θ_i is normally distributed with mean m_θ and variance σ_θ^2 and that $\varepsilon_i(t)$ is normally distributed with mean zero and variance σ_ε^2. Moreover, we assume that the abilities θ_i and the noise terms $\varepsilon_i(t)$ are independent across time and workers.

We can interpret $e_i(t)$ as worker i's flow output in the time interval $[t, t+\Delta)$ due to his effort at time t. Similarly, we can interpret $k_i(t)$ and θ_i as worker i's flow output in the same time interval due to his stock of human capital at time t and ability, respectively. The interpretation of output in terms of flows is useful in what follows.

The term $\sqrt{\Delta}$ multiplying the noise term in (1) implies that the variance of output per unit of 'real' time is constant. This ensures that the speed of learning about worker ability does not depend on Δ. This fact, in turn, implies that career concerns remain an important factor shaping the provision of incentives to workers regardless of the value of Δ.

Human Capital Acquisition. The human capital of workers evolves over time according to the following learning-by-doing process. For each $i \in \mathcal{J}$ and $t \in \mathcal{S}$,

$$k_i(t + \Delta) = e^{-\eta\Delta} k_i(t) + F(e_i(t), t)\Delta, \tag{2}$$

where $\eta \geq 0$ and $F(e, t)$ is continuously differentiable and concave in e for all $t \in \mathcal{S}$. The parameter η is the (instantaneous) rate at which a worker's accumulated stock of human capital

[3] We can substitute the assumption of Bertrand competition with free entry of firms.

[4] As it turns out, in equilibrium all workers exert the same effort and acquire the same amount of human capital.

depreciates over time. The function F describes how the learning-by-doing process depends on effort and time; the partial derivatives $F_e(e,t)$ reflect a worker's 'ability to learn', that is, ability to accumulate human capital, at different points in time.[5] All workers have the same initial stock of human capital, $k(0)$.

Notice from (2) that a worker's stock of human capital depends only on real time in the sense that changing Δ while keeping the worker's choices of effort at each point in time the same does not change his stock of human capital.

Preferences. For each $t \in \mathcal{S}$, the payoff to a worker from a sequence $\{w(t+\tau\Delta)\}_{\tau\in\{0,\ldots,(T-t)/\Delta\}}$ of wage payments and a sequence $\{e(t+\tau\Delta)\}_{\tau\in\{0,\ldots,(T-t)/\Delta\}}$ of effort choices is

$$-exp\left\{-r\left[\sum_{\tau=0}^{(T-t)/\Delta} e^{-\rho\tau\Delta}\left(w(t+\tau\Delta)-\frac{\lambda}{2}e(t+\tau\Delta)^2\Delta\right)\right]\right\}, \tag{3}$$

where $r > 0$ is the workers' coefficient of absolute risk aversion and $\rho > 0$ is their discount rate. The term $\lambda e(s)^2\Delta/2$ is the (monetary) cost to a worker of exerting effort $e(s)$ at time $s \in \mathcal{S}$. This cost is proportional to both the time interval between two consecutive periods and to a parameter $\lambda > 0$. We can interpret $\lambda e(s)^2/2$ as the flow cost of effort to a worker in the time interval $[s, s+\Delta)$ when his effort at time s is $e(s)$.

Contracts. At every $t \in \mathcal{S}$, firms can offer short-term contracts to workers describing their period-t wages as a function of their period-t output.[6] We assume that firms can only offer linear contracts. So, worker i's wage at time $t \in \mathcal{S}$ is $w_i(t) = a_i(t) + b_i(t)y_i(t)$, where $a_i(t)$ and $b_i(t)$ are, respectively, the worker's base salary and piece-rate at time t. We further assume that piece-rates are restricted to be non negative. We can justify this assumption by assuming that workers can costlessly reduce their output, so that it is never optimal for firms to offer negative piece-rates.[7]

The contracting problem of the firms consists of finding the base salary and piece-rate for each worker at every time $t \in \mathcal{S}$ depending on the information firms have about the workers.

---

[5] In the Appendix, we show that our results extend to the case in which the function F also depends on k, so that a worker's ability to learn depends on his stock of human capital.

[6] Gibbons and Murphy (1992) shows that assuming that long-term contracts are not feasible is equivalent to assuming that such contracts are feasible but must be renegotiation-proof. The same holds in our environment.

[7] Alternatively, we can follow Gibbons and Murphy (1992) and allow for (costly) negative effort. Our equilibrium characterization and identification results extend to this case.

Wages. Bertrand competition among the firms implies that they earn zero expected profits at every $t \in \mathcal{S}$. Hence,

$$w_i(t) = (1 - b_i(t))\mathbb{E}\left[y_i(t)|\mathcal{I}_i(t)\right] + b_i(t)y_i(t), \tag{4}$$

where $\mathbb{E}\left[y_i(t)|\mathcal{I}_i(t)\right]$ is worker i's expected output at time t conditional on the information $\mathcal{I}_i(t)$ firms have about worker i at this point in time. This information consists of the output realizations of worker i before time t. It follows from (4) that, on average, wages equal expected productivity.

2.2 Equilibria

The action and output histories for a worker at time $t \in \mathcal{S}$ are, respectively, the sequence of the worker's effort choices and output realizations from time 0 to time $t - \Delta$. The private history for a worker at time $t \in \mathcal{S}$ is the pair consisting of his action and output history at t. A strategy for a firm specifies contract offers to the workers conditional on their respective output histories. A strategy for a worker specifies a choice of effort after each private history and contract offer. We consider pure strategies.[8]

An equilibrium specifies strategies for firms and workers such that for each worker: (*a*) after any output history for the worker the firms offer a linear contract satisfying (4) that maximizes the worker's expected lifetime payoff given the firms' and the worker's future behavior; and (*b*) the worker's choice of effort at any point in time is optimal given his private history, the contract offered to him, and his and the firms' future behavior.

Condition (*a*) follows from Bertrand competition among firms—if the firms do not choose a piece-rate that maximizes a worker's expected lifetime payoff after some output history, then any firm that does not hire the worker can profitably deviate by offering a more attractive piece-rate to the worker. Condition (*b*) is a sequential rationality constraint. While intuitive, the optimality conditions in (*a*) and (*b*) are cumbersome to write. In the Appendix, we provide a formal definition of equilibria for the interested reader. We say that an equilibrium is *non-contingent* if the equilibrium piece-rates and effort choices depend only on time.

[8]The assumption that a worker does not condition his choice of effort on past contract offers is without loss of generality when we consider pure strategies.

2.3 Remarks

Our assumptions mirror functional form assumptions typical in the literature and allow for a complete characterization of equilibria. Together with the linearity of output on ability and noise, the linearity of contracts, and the normality assumptions on risk and uncertainty about ability, our choice of worker preferences allows us to represent these preferences by certainty equivalents. This, in turn, implies that the trade-off between compensation and effort is affected neither by information about ability revealed in the market nor by the accumulation of human capital. Thus, as in Gibbons and Murphy (1992), the equilibrium is unique and non-contingent.

3 Preliminaries

In order to set the stage for our equilibrium characterization in the next section, in this section we first discuss how learning about ability takes place in equilibrium, and then discuss how career concerns and learning-by-doing affect the (private) returns of effort.

3.1 Learning in Equilibrium

Firms use a worker's performance together with his effort and human capital in equilibrium to learn about the worker's ability. Consider worker $i \in \mathcal{J}$ and suppose that his equilibrium effort and human capital at time $t \in \mathcal{S}$ are $e_i^*(t)$ and $k_i^*(t)$, respectively. In principle, $e_i^*(t)$ and $k_i^*(t)$ can depend on the worker's output history at t, that is, they can be random.[9] Now let

$$z_i(t) = y_i(t) - (e_i^*(t) + k_i^*(t))\Delta$$

be the signal about worker i's ability that the firms observe at time t; $z_i(t)$ is the part of worker i's output at time t that cannot be explained by his effort and human capital. Given that in equilibrium firms correctly anticipate worker i's effort and stock of human capital at time t, it follows that

$$z_i(t) = \theta_i\Delta + \varepsilon_i(t)\sqrt{\Delta}.$$

[9]In equilibrium, $k_i^*(t)$ is completely determined by the past effort choices of worker i. Moreover, in a pure-strategy equilibrium, $e_i^*(t)$ is completely determined by the worker's output history at t.

Since the both the ability of workers and the signals about worker ability are normally distributed, it follows from standard results that the firms' posterior belief about a worker's ability is normally distributed. Moreover, if $m_i(t) = \mathbb{E}[\theta|\mathcal{I}_i(t)]$ and $\sigma_i^2(t) = \mathrm{Var}(\theta|\mathcal{I}_i(t))$ are, respectively, the mean and variance of the posterior belief about worker i's ability at time $t \in \mathcal{S}$, with the convention that $m_i(0) = m_{\theta,i}$ and $\sigma_i^2(0) = \sigma_\theta^2$, then $m_i(t)$ and $\sigma_i^2(t)$ evolve as follows:[10]

$$m_i(t+\Delta) = \frac{\sigma_\varepsilon^2}{\sigma_i^2(t)\Delta + \sigma_\varepsilon^2} m_i(t) + \frac{\sigma_i^2(t)}{\sigma_i^2(t)\Delta + \sigma_\varepsilon^2} z_i(t); \tag{5}$$

$$\sigma_i^2(t+\Delta) = \frac{\sigma_i^2(t)\sigma_\varepsilon^2}{\sigma_i^2(t)\Delta + \sigma_\varepsilon^2} \tag{6}$$

We refer to $m_i(t)$ as worker i's *reputation* at time t.

It follows from the law of motion (6) that

$$\sigma_i^2(t) \equiv \sigma^2(t) = \frac{\sigma_\theta^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + t\sigma_\theta^2} \tag{7}$$

for all $t \in \mathcal{S}$. Thus, the variance of the posterior belief about a worker's ability does not depend on his identity and evolves deterministically over time. Moreover, this variance depends on Δ only through t; in particular, a reduction in Δ while keeping t constant does not change $\sigma^2(t)$. So, the speed of learning about worker ability depends only on real time.

It follows from the law of motion (5) that the reputation of each worker i at any time $t \in \mathcal{S}$ can be written as a weighted average of the worker's initial reputation and the average signal about his ability received so far:

$$m_i(t) = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + t\sigma_\theta^2} m_{\theta,i} + \frac{t\sigma_\theta^2}{\sigma_\varepsilon^2 + t\sigma_\theta^2} \left(\frac{1}{t} \sum_{\tau=0}^{(t/\Delta)-1} z_i(\tau\Delta) \right). \tag{8}$$

The term $\Sigma(t) = \sigma_\theta^2/(\sigma_\varepsilon^2 + t\sigma_\theta^2)$ in (8) describes the discount applied by firms to the signals they observe about a worker's ability before t when revising their beliefs about the worker's ability.

Notice that $\Sigma(t)$ is strictly decreasing in t and converges to zero as t approaches infinity: because of learning about ability, firms put progressively less weight on signals about ability as time progresses. In particular, when T is sufficiently large, there exists a point in time after which the workers' reputation stays basically the same. At this stage, career concerns stop being relevant for worker behavior and only learning-by-doing matters for shaping the provision of incentives.

[10] For completeness, we derive equations (5) and (6) in the Appendix.

3.2 Dynamic Returns to Effort

We now discuss how dynamic considerations due to career concerns and learning-by-doing affect the returns of effort to workers. We restrict attention to non-contingent equilibria. Since in a non-contingent equilibrium effort choices and piece-rates evolve deterministically over time, in this class of equilibria a worker's effort and output at a given point in time do not affect his future effort choices and piece-rates.

Consider a non-contingent equilibrium and for each $i \in \mathcal{J}$ and $t \in \mathcal{S}$, let $e_i^*(t)$ and $b_i^*(t)$ be worker i's equilibrium effort and piece-rate at time t, respectively. Moreover, let $m_i^*(t)$ be worker i's equilibrium reputation at time t. At time $t \in \mathcal{S}$, worker i chooses the effort e that maximizes the *flow* certainty equivalent

$$\frac{1}{\Delta}\sum_{\tau=0}^{(T-t)/\Delta} e^{-\rho\tau\Delta}\mathbb{E}[w_i(t+\tau\Delta)|e, m_i^*(t)] - \lambda\frac{e^2}{2},$$

where $\mathbb{E}[w_i(t+\tau\Delta)|e, m_i^*(t)]$ is worker i's expected wage at time $t+\tau\Delta$ given his choice of effort and reputation at time t. In equilibrium, the value of e that maximizes the certainty equivalent in question must be $e_i^*(t)$.

Notice that

$$\frac{1}{\Delta}\frac{\partial\mathbb{E}[w_i(t)|e, m_i^*(t)]}{\partial e} = b_i^*(t).$$

So, as in a static moral hazard problem, the marginal flow return to worker i's effort at time t depends on his piece-rate at t. However, given that worker i's effort at time t also affects his reputation and stock of human capital in the future, the derivatives

$$\frac{1}{\Delta}\frac{\partial\mathbb{E}[w_i(t+\tau\Delta)|e, m_i^*(t)]}{\partial e},$$

with $\tau \in \{1, \ldots, (T-t)/\Delta\}$, are different from zero as well. These marginal changes on worker i's future flow expected compensation reflect both the contribution of career concerns and learning-by-doing to the marginal flow return of effort e at time t to worker i. We compute these *dynamic* components of the marginal return of effort in what follows.

It follows from (1) and (8) that on the margin, effort at time t increases worker i's reputation at time $t+\tau\Delta$ by $\Sigma(\tau+t\Delta)\Delta$. This, in turn, increases worker i's flow expected compensation at the same point in time by

$$(1 - b_i^*(t + \tau\Delta))\Sigma(t + \tau\Delta)\Delta;$$

competition among firms implies that a worker's reputation affects the fixed component of his pay. Discounting and summing over his remaining life-cycle the above terms, we obtain worker i's marginal flow return of effort at time t due to career concerns:

$$R_{CC,i}^*(t) = \sum_{\tau=1}^{(T-t)/\Delta} \Delta e^{-\rho\tau\Delta}(1 - b_i^*(t + \tau\Delta))\Sigma(t + \tau\Delta).$$

Now observe from (2) that on the margin, effort e at time t increases worker i's flow output at time $t + \tau\Delta$ by $e^{-\eta(\tau-t-1)\Delta}F_e(e,t)\Delta$, the marginal increase in worker i's human capital at this point in time. This increase in flow output at time $t + \tau\Delta$ has two consequences for worker i's lifetime compensation. First, it directly affects compensation at time $t + \tau\Delta$ because of the pay-for-performance component $b_i^*(t + \tau\Delta)$. Second, the increment in output at time $t + \tau\Delta$ due to increased human capital affects worker i's reputation in all subsequent periods. The marginal increase in worker i's flow expected compensation at time $t + \tau\Delta$ due to the first effect is

$$b_i^*(t + \tau\Delta)e^{-\eta(\tau-1)\Delta}F_e(e,t)\Delta.$$

From (8), the marginal increase in worker i's flow expected compensation at time $t + \tau\Delta$ due to the second effect is

$$(1 - b_i^*(t + \tau\Delta))\Sigma(t + \tau\Delta)\Delta \sum_{s=1}^{\tau-1} e^{-\eta(s-1)\Delta}F_e(e,t)\Delta.$$

This second term reflects the cumulative impact on worker i's reputation at $t + \tau\Delta$ following the increase in his stock of human capital at all points in time between $t + \Delta$ and $t + (\tau - 1)\Delta$.

Discounting and summing over his remaining life-cycle the terms discussed above, we obtain worker i's marginal flow return of effort e at time t due learning-by-doing:

$$R_{LDB,i}^* = F_e(e,t)e^{\eta\Delta} \sum_{\tau=1}^{(T-t)/\Delta} \Delta e^{-(\rho+\eta)\tau\Delta}\Bigg[b_i^*(t + \tau\Delta) + (1 - b_i^*(t + \tau\Delta))\Sigma(t + \tau\Delta)\sum_{s=1}^{\tau-1} \Delta e^{-\eta(s-1)\Delta}\Bigg]. \quad (9)$$

Notice that since $\Sigma(t+\tau\Delta)\tau\Delta < 1$, $R^*_{LBD,i}(e,t)$ is positive if the piece-rates $b^*_i(t+\tau\Delta)$ are smaller than one. Also notice that $R^*_{LBD,i}(e,t)$ is decreasing in e by the concavity of $F(e,t)$ in e. In the Appendix, we show that

$$R^*_{LDB,i}(e,t) = F_e(e,t)e^{\eta\Delta} \sum_{\tau=1}^{(T-t)/\Delta} \Delta e^{-(\rho+\eta)\tau\Delta} \left[b^*_i(t+\tau\Delta) + R^*_{CC,i}(t+\tau\Delta)\right].$$

Since the marginal flow cost of effort e is λe, it follows from the above reasoning that worker i's equilibrium effort at time t is the unique solution to the following (necessary and sufficient) first-order condition:

$$\lambda e^*_i(t) = \max\left\{0, b^*_i(t) + R^*_{CC,i}(t) + R^*_{LBD,i}(e^*_i(t),t)\right\}.$$

The above first-order condition describes worker i's equilibrium choice of effort as a function of the static and dynamic marginal flow returns of his effort. Thus, dynamic considerations lead to an effort that is different than the effort that would be implied by static considerations only. This, in turn, implies that equilibrium piece-rates are shaped by dynamic considerations as well. In the next section, we discuss how learning about ability and learning-by-doing shape the provision of incentives and show that the latter can imply that piece-rates are eventually decreasing over time.

4 Explicit Incentives in the Presence of Career Concerns and Learning-by-Doing

In this section, we study how piece-rates are determined in equilibrium. First, we state the proposition that is central to our analysis, Proposition 1. It asserts that the equilibrium is unique and non-contingent and recursively determines the time-path of explicit incentives in our economy. Second, we consider the different forces that shape the provision of explicit incentives over the life-cycle. We discuss a decomposition of the equilibrium piece-rates into components that admit an economic interpretation. Third, and finally, we describe how piece-rates can vary over the life-cycle depending on the strength of career concerns and learning-by-doing.

One of the main substantive contributions of this paper is to illustrate that explicit incentives do not necessarily increase over the life-cycle as implied by a pure career-concerns model. Explicit incentives depend in complex ways on career concerns and learning-by-doing, and there are no global results available for how explicit incentives vary over the life-cycle. However, our analysis clarifies when we expect explicit incentives to increase over time or not. Our analysis also provides guidance on what shapes of life-cycle profiles in explicit incentives are likely to manifest themselves in the data if worker compensation is indeed determined by the interplay between career concerns, learning-by-doing, and explicit incentives.

4.1 A Recursive Formulation of the Compensation Contract

We now state Proposition 1, which is the key theoretical result of the paper.

Proposition 1. *The equilibrium is unique and non-contingent. Moreover, workers exert the same effort, $e^*(t)$, and are offered the same piece-rate, $b^*(t)$, at every $t \in \mathcal{S}$. The equilibrium efforts and piece-rates are determined as follows. For each $t \in \mathcal{S}$, let $b^0(t)$, $R^*_{CC}(t)$, $R^*_{LBD}(e,t)$, and $H^*(t)$ to be such that:*

$$b^0(t) = \frac{1}{1 + r\lambda[\Delta\sigma^2(t) + \sigma^2_\varepsilon]}; \tag{10}$$

$$R^*_{CC}(t) = \sum_{\tau=1}^{(T-t)/\Delta} \Delta e^{-\rho\tau\Delta}(1 - b^*(t+\tau\Delta))\Sigma(t+\tau\Delta); \tag{11}$$

$$R^*_{LBD}(e,t) = F_e(e,t)e^{\eta\Delta} \sum_{\tau=1}^{(T-t)/\Delta} \Delta e^{-(\rho+\eta)\tau\Delta}\left[b^*(t+\tau\Delta) + R^*_{CC}(t+\tau\Delta)\right]; \tag{12}$$

$$H^*(t) = r\lambda\sigma^2_\varepsilon\Sigma(t) \sum_{\tau=1}^{(T-t)/\Delta} \Delta e^{-\rho\tau\Delta}. \tag{13}$$

Moreover, let $\gamma^*(t) = F_e(e^*(t), t)$. Then, $e^*(t)$ and $b^*(t)$ are given by

$$\lambda e^*(t) = \max\left\{0, b^*(t) + R^*_{CC}(t) + R^*_{LBD}(e^*(t), t)\right\} \tag{14}$$

and

$$b^*(t) = \max\left\{0, b^0(t)\left[1 + \gamma^*(t)e^{\eta\Delta}\sum_{\tau=1}^{(T-t)/\Delta}\Delta e^{-(\rho+\eta)\tau\Delta} - R^*_{LBD}(e,t) - R^*_{CC}(t)\right] - H^*(t)\right\}. \tag{15}$$

The first-order condition for the equilibrium choice of effort, equation (14), follows from the discussion in the previous section. In the Appendix we show that for all $\Delta > 0$ the equilibrium is unique and non-contingent, that effort choices and piece-rates at any point in time are the same for all workers, and establish the expression for the equilibrium piece-rates.

Equation (15) has a clear economic interpretation. The term

$$1 + \gamma^*(t)e^{\eta\Delta} \sum_{\tau=1}^{(T-t)/\Delta} \Delta e^{-(\rho+\eta)\tau\Delta}$$

is the marginal (flow) social return of effort $e^*(t)$ at time t, that is, the marginal increase in a worker's present discounted expected flow output from time t on when he exerts effort $e^*(t)$ at time t. On the other hand, the term $R^*_{LBD}(e^*(t), t) + R^*_{CC}(t)$ is the marginal (flow) benefit to a worker of the same choice of effort at time t. Equation (15) shows that the equilibrium piece-rate at $t \in \mathcal{S}$ is proportional to the wedge between the marginal social benefit of the equilibrium effort at time t and the marginal private benefit of the same choice of effort at time t, minus a term $H^*(t)$ that we explain below. The constant of proportionality in (15) reflects the risk-incentives trade-off. In the first best, firms would like to align the incentives of workers with the social optimum. Because of moral hazard, this alignment is only partial.

The term $H^*(t)$ in (15) represents a dynamic hedge against the lifetime risk in compensation that workers face because of the uncertainty about their ability. Indeed, because of this uncertainty, variation in output at time t represents not only variation in compensation at time t, but also variation in future compensation. Indeed, a worker's output at time T affects his reputation in all subsequent points in time, which in turn affects his future compensation. By introducing a negative correlation between a worker's current compensation and his future compensation, the term $H^*(t)$ provides the workers with insurance against this risk in lifetime compensation.

While the dynamic hedge term is always positive, the marginal returns of effort due to career concerns and learning-by-doing can be negative. This is only possible if piece-rates are greater than one, though. The next result shows that if the workers' ability to learn is small enough, then piece-rates are smaller than one regardless of Δ. The proof of Proposition 2 is in the Appendix.

Proposition 2. *There exists $\overline{\gamma} > 0$ independent of Δ such that if $F(0, t) < \overline{\gamma}$ for all $t \in \mathcal{S}$, then $b^*(t) < 1$ for all $t \in \mathcal{S}$.*

4.2 Decomposition of Equilibrium Piece-Rates

It follows from (15) that if $b^*(t) > 0$, then

$$b^*(t) = b^0(t) - b^0(t)R^*_{CC}(t) - H^*(t) + b^0(t)\left[\gamma^*(t)e^{\eta\Delta}\sum_{\tau=1}^{(T-t)/\Delta}\Delta e^{-(\rho+\eta)\tau\Delta} - R^*_{LBD}(e,t)\right]. \quad (16)$$

This equation is useful because it decomposes the equilibrium piece-rates into four components that admit an economic interpretation. We discuss this decomposition next.

The first term in (16) is the solution to the static problem of providing incentives for effort when preferences are CARA and effort costs are quadratic. In the absence of dynamic considerations, the firms would offer the piece-rate $b^0(t)$ to the workers at $t \in \mathcal{S}$. It then follows from (16) that we can represent the optimal piece-rate at any point in time as a deviation from the static moral hazard piece-rate. As we now discuss, this deviation reflects the impact of dynamic considerations due to career concerns and learning-by-doing.

The second and third terms in (16) are familiar to researchers in personnel economics. The second term adjusts explicit incentives to account for the career-concerns incentive for effort. The intuition is familiar from Gibbons and Murphy (1992) and Holmström (1999): the career concerns motive provides workers with an incentive to exert effort even in the absence of an explicit link between pay and performance. Since workers are risk-averse, and so incentive provision is costly for them, the firms find it optimal to set piece-rates smaller than they would be in the absence of learning about ability.[11] The third term in (16) is the dynamic hedge term already discussed in the previous subsection. Notice that $H^*(t)$ is strictly decreasing in t. This is intuitive: since uncertainty about worker ability decreases over time, so does the need to hedge workers against the lifetime variation in compensation due to learning about ability.

The fourth and last term in (16) represents the contribution of learning-by-doing to the equilibrium piece-rates. In order to understand this term, notice that

$$MSR^*_{LBD}(t) = \gamma^*(t)e^{\eta\Delta}\sum_{\tau=1}^{(T-t)/\Delta}\Delta e^{-\rho\tau\Delta}$$

[11] Assuming that the terms $R^*_{CC}(t)$ are positive.

is the marginal flow social benefit of the equilibrium effort at time t due to learning-by-doing. It follows from (9) that if piece-rates are smaller than one, then $R^*_{LBD}(e^*(t), t) < MSR^*_{LBD}(t)$ for all $t < T$.[12] This is intuitive. When future piece-rates are smaller than one, a worker does not capture completely the returns to his investment in human capital. So, this investment is smaller than the socially optimal investment. The term in consideration partially offsets this undersupply of effort; it does not do so completely because of the risk-incentives trade-off.

To summarize, equation (16) describes the equilibrium piece-rate at a point in time as the static moral hazard piece-rate plus three terms. The first term describes the adjustment to the equilibrium piece-rate due to career concerns, and is negative if piece-rates are smaller than one. The second term describes the dynamic hedge against the lifetime risk in compensation due to uncertainty about ability, and is also negative. The third term describes the adjustment to the equilibrium piece-rate due to learning-by-doing and is positive if piece-rates are smaller than one. Both the static moral hazard piece-rate and the dynamic hedge term are decreasing over time and so are a force in favor of increasing piece-rates. In the next part we discuss how the remaining terms shape the provision of incentives over the life-cycle.

4.3 Explicit Incentives over the Life-Cycle

We now turn to discuss how the strength of career concerns and learning-by-doing affects the evolution of piece-rates over the life-cycle.

To begin, we mute learning-by-doing by setting $F_e(e, t) \equiv 0$ and show that piece-rates are nondecreasing over time in this case. In the absence of learning-by-doing, (15) reduces to

$$b^*(t) = \max\left\{0, b^0(t)[1 - R^*_{CC}(t)] - H^*(t)\right\}. \tag{17}$$

In the Appendix, we show that $R^*_{CC}(t)$ is strictly decreasing over time in the pure career concerns case. The desired result follows from the fact that for all $t \leq T - \Delta$, $b^*(t) > 0$ only if $R^*_{CC}(t) < 1$.

Lemma 1. *Suppose that $F_e(e, t) \equiv 0$ and $\sigma^2_\theta > 0$. Then $b^*(t)$ is nondecreasing over time and is strictly increasing over time whenever it is greater than zero.*

[12] Indeed, $b^*(t + \tau\Delta) < 1$ implies that $b^*(t + \tau\Delta) + (1 - b^*(t + \tau\Delta))\Sigma(t + \tau\Delta) \sum_{s=1}^{\tau-1} \Delta e^{-\eta(s-1)\Delta}$ is bounded above by $b^*(t + \tau\Delta) + (1 - b^*(t + \tau\Delta))\Sigma(t + \tau\Delta)\tau\Delta$, which in turn is smaller than one.

Lemma 1 extends the result in Gibbons and Murphy (1992) that piece-rates increase over time to the case in which the time interval between two consecutive periods is arbitrary, and its intuition is the same. Our direct characterization of the equilibrium piece-rates allows for a simpler proof that piece-rates are nondecreasing over time than in Gibbons and Murphy (1992).[13]

As the next step, we mute the effects arising from career concerns by setting $\sigma_\theta^2 = 0$ and consider the pure learning-by-doing case. In this case, equation (15) reduces to

$$b^*(t) = b^0 \left(1 + \gamma^*(t) e^{\eta\Delta} \sum_{\tau=1}^{(T-t)/\Delta} \Delta e^{-(\rho+\eta)\tau\Delta} [1 - b^*(t+\tau\Delta)] \right), \tag{18}$$

where $b^0 = 1/(1 + r\lambda\sigma_\varepsilon^2)$ is the static moral hazard piece-rate in the absence of uncertainty about ability. In particular, unlike in the pure career concerns case, equilibrium piece-rates can be greater than the static moral hazard piece-rates when only learning-by-doing is present.

We know that the contribution of learning-by-doing to the equilibrium piece-rates depends positively on the marginal (flow) social benefit of effort due to learning-by-doing and negatively on the marginal private benefit of effort due to learning-by-doing. Given that both these marginal benefits tend to decrease over time, as the horizon over which effort affects a worker's output decreases as the worker gets older, there is no reason a priori to expect that equilibrium piece-rates behave monotonically over time in the pure learning-by-doing case. The next result shows that, nevertheless, equilibrium piece-rates in the pure learning-by-doing case are eventually strictly decreasing over time as long as Δ is sufficiently small. The proof of Lemma 2 is in the Appendix.

Lemma 2. *Suppose that $F_e(0,t) > 0$ for all $t \in \mathcal{S}$ and $\sigma_\theta^2 = 0$. There exists $\alpha \in (0,1)$ and $\overline{\Delta} > 0$ such that if $\Delta < \overline{\Delta}$, then $b^*(t)$ is strictly decreasing in t as long as $T/t \geq 1 - \alpha$.*

The above lemma shows that at least for workers with long enough experience in the labor market, learning-by-doing can lead to the opposite prediction of career concerns in terms of how piece-rates behave over time. The next result shows that there are specific cases in which equilibrium piece-rates are strictly decreasing over time in the pure learning-by-doing case. The proof of Lemma 3 is also in the Appendix.

[13] Gibbons and Murphy (1992) characterizes the equilibrium in terms of total incentives, $B^*(t) = b^*(t) + R_{CC}^*(t)$.

Lemma 3. *Suppose that* $F_e(e,t) \equiv \gamma$ *and* $\sigma_\theta^2 = 0$*. There exists* $\overline{\gamma} > 0$ *such that if* $\gamma < \overline{\gamma}$*, then* $b^*(t)$ *is strictly decreasing over time. Moreover, for all* $\gamma > 0$*, there exists* $\overline{\Delta} > 0$ *such that* $b^*(t)$ *is strictly decreasing over time if* $\Delta < \overline{\Delta}$*.*

Lemmas 1 to 3 show that in isolation career concerns and learning-by-doing can lead to different predictions about the life-cycle behavior of equilibrium piece-rates. Intuition suggests that when both forces are present, the stronger one should shape the evolution of piece-rates. Since learning about ability takes place in the long-run, one should expect that if workers are sufficiently long-lived, then learning-by-doing eventually prevails; after a while uncertainty about ability becomes residual and career concerns stopping affecting incentives. On the other hand, one expects that if σ_θ^2 is large enough, then career concerns should prevail when workers are young, making equilibrium piece-rates initially nondecreasing. The next result confirms this intuition. A corollary of Proposition 3 is that if both σ_θ^2 and T are sufficiently large, then piece-rates are maximized at some intermediate point in time.

Proposition 3. *Suppose that* $\sigma_\theta^2 > 0$ *and* $F_e(0,t) > 0$ *for all* $t \in \mathcal{S}$*. The following two facts hold.*

1. There exists $\overline{T} > 0$ *such that for all* $T > \overline{T}$*, there exist* $\alpha \in (0,1)$ *and* $\overline{\Delta} > 0$ *such that if* $\Delta < \overline{\Delta}$*, then* $b^*(t)$ *is strictly decreasing in* t *if* $T/t \geq 1-\alpha$*. The cutoffs* $\overline{T}$*,* $\overline{\Delta}$*, and* α *can be taken to be independent of* σ_θ^2*.*

2. There exists $\overline{\sigma}_\theta^2$ *such that if* $\sigma_\theta^2 > \overline{\sigma}_\theta^2$*, then there exists* $\alpha \in (0,1)$ *such that* $b^*(t)$ *is nondecreasing in* t *it* $t/T \leq \alpha$*. The cutoffs* $\overline{\sigma}_\theta^2$ *and* α *can be taken to be independent of* T *and* Δ*.*

In general, the interplay between career concerns and learning-by-doing generates complex incentives which can lead to a wide variety of patterns in pay-for-performance over the life-cycle. For example, it is possible that piece-rates are *minimized* at some intermediate point in time. This can be the case, for instance, when learning-by-doing is important only at the beginning of a worker's career. As an extreme example, suppose that $\underline{\gamma}_0 = \lim_{e\to+\infty} F_e(e,0) > 0$, but that $F_e(0,t) = 0$ for all $t > 0$. In this case, $b^*(t)$ is smaller than $b^0(t)$ and nondecreasing in t for all $t > 0$. It is easy to see that if $\underline{\gamma}_0$ is sufficiently large, then $b^*(0) > b^0(0)$; see the Appendix for details. While our model has ambiguous predictions about how explicit incentives should vary over the life-cycle, in the next section we show that it nevertheless has empirical content.

5 Identification

In this section we discuss identification of the crucial parameters of our model. These parameters are the variances σ_θ^2 and σ_ε^2, the marginal products $\gamma^*(t) = F_e(e^*(t), t)$ for $t \in \{0, \Delta, \ldots, T-\Delta\}$, the depreciation rate η, and the product $r\lambda$. The pair $(\sigma_\theta^2, \sigma_\varepsilon^2)$ governs career concerns, the list $(\eta, (\gamma^*(t))_{t=0}^{T-\Delta})$ governs learning-by-doing, and the product $r\lambda$ governs the risk-incentives trade-off. We assume that the data is generated for a fixed value of Δ. To economize on notation, let $\Delta = 1$, so that $\mathcal{S} = \{0, \ldots, T\}$, and set $\mu = e^{-\eta}$ and $\delta = e^{-\rho}$. We assume that δ is known.

Our equilibrium characterization in the previous section shows that the evolution of piece-rates over the life-cycle reflects the relative strength of career concerns and learning-by-doing. It is not surprising, then, that the evolution of piece-rates is central to our identification argument. The identification argument presented below emphasizes the centrality of piece-rates for identification by describing how we can first identify the piece-rates $(b^*(t))_{t=0}^{T}$ together with the pair $(\sigma_\theta^2, \sigma_\varepsilon^2)$, and then showing how we can identify the remaining parameters—the list $(\mu, (\gamma^*(t))_{t=0}^{T-1})$ and the product $r\lambda$—from this information and the known value of δ.

While knowing the parameter vector $(\sigma_\theta^2, \sigma_\varepsilon^2, \eta, (\gamma^*(t))_{t=0}^{T-1}, r\lambda)$ allows us to decompose the observed piece-rates into the components discussed in the previous section, thus allowing us to uncover how the contribution of career concerns and learning-by-doing to explicit incentives changes over the life cycle, this information is not enough to conduct a counterfactual analysis. For the latter, one needs to identify the human capital production function F. In this section we also establish that F is non-parametrically non-identified, but that we can identify the efforts $(e^*(t))_{t=0}^{T}$ and the list of parameters $(F(e^*(t), t))_{t=0}^{T-1}$ up to a vertical translation of F and a choice of the initial stock of human capital. We then discuss how we can use this information to identify F under some flexible parametric assumptions.

5.1 Identifying Piece-Rates and Learning Parameters with a Wage Panel

We begin by showing how we can identify the pair $(\sigma_\theta^2, \sigma_\varepsilon^2)$ and the piece-rates $(b^*(t))_{t=0}^{T}$ from a panel of wages as long as there are at least three observation periods.

The model implies that the observed wages of a worker i at any time t are given by

$$w_i(t) = h_i(t) + (1 - b^*(t))\, \mathbb{E}[\theta_i | \mathcal{I}_i(t)] + b^*(t)\theta_i + b^*(t)\varepsilon_i(t),$$

where $h_i(t) \equiv h(t)$ is a constant term that depends on the worker's stock of human capital, θ_i is normally distributed with mean 0 and variance σ_θ^2, and $\varepsilon_i(t)$ is normally distributed with mean 0 and variance σ_ε^2.[14] Moreover, the abilities θ_i and the noise terms $\varepsilon_i(t)$ are independent across workers and time.

As a first step, for each worker i let

$$r_i(t) = (1 - b^*(t))\, \mathbb{E}[\theta_i | \mathcal{I}_i(t)] + b^*(t)\theta_i + b^*(t)\varepsilon_i(t)$$

be the random part of the worker's wage at time t; notice that $\mathbb{E}[r_i(t)] = 0$. We work with the residuals $r_i(t)$ in our identification argument. We obtain the empirical counterparts of these residuals by regressing the wages of workers against a set of dummies in experience. The next result provides a complete characterization of the variance-covariance matrix of the residuals $r_i(t)$ implied by our theory. The proof of Lemma 4 is in the Appendix. Recall that $\sigma^2(t) = \sigma_\theta^2 \sigma_\varepsilon^2 / (\sigma_\theta^2 + \sigma_\varepsilon^2)$ is the variance of the posterior belief about a worker's ability at time t

Lemma 4. *The following two facts hold.*

1. *For all* $0 \leq s < t \leq T$, $\mathrm{Cov}[r_i(s), r_i(t)] = b^*(s)\sigma^2(s) + \sigma_\theta^2 - \sigma^2(s)$.
2. *For all* $0 \leq t \leq T$, $\mathrm{Var}[r_i(t)] = b^*(t)^2(\sigma^2(t) + \sigma_\varepsilon^2) + \sigma_\theta^2 - \sigma^2(t)$.

We can now identify the pair $(\sigma_\theta^2, \sigma_\varepsilon^2)$ as follows. Let

$$\alpha = \frac{(\mathrm{Cov}[r_i(0), r_i(1)])^2}{\mathrm{Var}[r_i(0)]};$$

we can construct α from the data. By Lemma 4, it follows that

$$\alpha = \sigma_\theta^2 \cdot \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2} = \sigma^2(1) \cdot \frac{\sigma_\theta^2}{\sigma_\varepsilon^2},$$

where the second equality follows from the fact that $\sigma^2(1) = \sigma_\theta^2 \sigma_\varepsilon^2 / (\sigma_\theta^2 + \sigma_\varepsilon^2)$. Now let

$$\beta = \frac{(\mathrm{Cov}[r_i(1), r_i(2)] - \alpha)^2}{\mathrm{Var}[r_i(1)] - \alpha};$$

[14] The mean talent $m_{\theta,i}$ of worker is absorbed into the intercept of $h_i(t)$.

we can also construct β from the data. Since $\alpha = \sigma_\theta^2 - \sigma^2(1)$, it follows from Lemma 4 that

$$\beta = \sigma^2(1) \cdot \frac{\sigma^2(1)}{\sigma^2(1) + \sigma_\varepsilon^2}.$$

On the other hand, given that

$$\frac{\sigma^2(1)}{\sigma^2(1) + \sigma_\varepsilon^2} = \frac{\sigma_\theta^2}{\sigma_\varepsilon^2} \cdot \frac{1}{2(\sigma_\theta^2/\sigma_\varepsilon^2) + 1},$$

we then have that

$$\beta = \sigma^2(1) \cdot \frac{\sigma_\theta^2}{\sigma_\varepsilon^2} \cdot \frac{1}{2(\sigma_\theta^2/\sigma_\varepsilon^2) + 1} = \alpha \cdot \frac{1}{2(\sigma_\theta^2/\sigma_\varepsilon^2) + 1}.$$

So, we can identify the signal-to-noise ratio $\xi = \sigma_\theta^2/\sigma_\varepsilon^2$ from the data. To conclude, notice that as

$$\alpha = \sigma_\theta^2 \cdot \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2} = \sigma_\theta^2 \cdot \frac{\xi}{1 + \xi},$$

we can then identify σ_θ^2, and thus σ_ε^2, from the data.

To finish this part of the identification argument, notice that once we know the pair $(\sigma_\theta^2, \sigma_\varepsilon^2)$, we can identify the piece-rates $b^*(t)$, with $t \in \{0, \dots, T\}$, from the variances $\mathrm{Var}[r_i(t)]^2$. To summarize, we can identify the list $(\sigma_\theta^2, \sigma_\varepsilon^2, (b^*(t))_{t=0}^T)$ from a panel of wage payments as long as we have observations for at least three periods.

5.2 Identifying the Remaining Parameters

We now show how we can identify the vector of parameters $(\mu, (\gamma^*(t))_{t=0}^{T-1}, r\lambda)$ from the career concerns parameters $(\sigma_\theta^2, \sigma_\varepsilon^2)$ and the piece-rates $(b^*(t))_{t=0}^T$ for a known value of δ.

We begin by noting that since $b^*(T) = 1/[1 + r\lambda(\sigma^2(T) + \sigma_\varepsilon^2)]$, we can identify $r\lambda$ from $b^*(T)$. Now recall that the equilibrium piece-rate and effort at time $t \in \{0, \dots, T\}$ are given by

$$b^*(t) = b^0(t) \left(1 + \gamma^*(t) \sum_{\tau=1}^{T-t} \delta^\tau \mu^{\tau-1} - R_{CC}^*(t) - R_{LBD}^*(e^*(t), t)\right) - H^*(t)$$

and

$$\lambda e^*(t) = b^*(t) + R_{CC}^*(t) + R_{LBD}(e^*(t), t),$$

respectively. As $(b^0(t))_{t=0}^T$, $(R_{CC}^*(t))_{t=0}^T$, and $(H^*(t))_{t=0}^T$ depend only on the pair $(\sigma_\theta^2, \sigma_\varepsilon^2)$ and the piece-rates $(b^*(t))_{t=0}^T$, we can then express

$$\tilde{b}(t) = b^0(t)\left(\gamma^*(t)\sum_{\tau=1}^{T-t}\delta^\tau\mu^{\tau-1} - R_{LBD}^*(e^*(t),t)\right) \tag{19}$$

$$= b^0(t)\gamma^*(t)\sum_{\tau=1}^{T-t}\delta^\tau\mu^{\tau-1}(1-b^*(t+\tau)-R_{CC}^*(t+\tau)), \tag{20}$$

the contribution of learning-by-doing to the equilibrium piece-rate at time t, as a function of known parameters. We use the terms $(\widetilde{b}(t))_{t=0}^T$ to identify μ and $(\gamma^*(t))_{t=0}^{T-1}$.

First notice from (20) that for all $t \in \{0,\ldots,T-1\}$,

$$\widetilde{b}(t) = b^0(t)\gamma^*(t)\delta(1-b^*(t+1)-R_{CC}^*(t+1)) + \frac{b^0(t)}{b^0(t+1)}\frac{\gamma^*(t)}{\mu^{T-t-2}}\frac{\mu^{T-t-1}}{\gamma^*(t+1)}\delta\widetilde{b}(t+1). \tag{21}$$

Given that

$$\widetilde{b}(T-1) = b^0(T-1)\gamma^*(T-1)\delta(1-b^*(T)),$$

we can back out $\gamma^*(T-1)$ from $\widetilde{b}(T-1)$. Using recursion (21) for $t=T-2$, we can then back out $\gamma^*(T-2)\mu$ from $\widetilde{b}(T-2)$ and $\widetilde{b}(T-1)$ (together with $\gamma^*(T-1)$). Continuing in this fashion, we can back out $\gamma^*(t)\mu^{T-t-1}$ for all $t \in \{0,\ldots,T-1\}$ from $(\widetilde{b}(t))_{t=0}^T$.

Now observe from the definition of $R_{LBD}^*(e,t)$ that for all $t \in \{0,\ldots,T-1\}$.

$$R_{LBD}^*(e^*(t),t) = \gamma^*(t)\delta(1-b^*(t+1)-R_{CC}^*(t)) + \frac{\gamma^*(t)\mu}{\gamma^*(t+1)}\delta R_{LBD}^*(e^*(t+1),t+1). \tag{22}$$

Since $R_{LBD}^*(e^*(T-1),T-1) = \gamma^*(T-1)\delta(1-b^*(T))$ and we know

$$\frac{\gamma^*(t)\mu^{T-t-1}}{\gamma^*(t+1)\mu^{T-t-2}} = \frac{\gamma^*(t)\mu}{\gamma^*(t+1)}$$

for all $t \in \{0,\ldots,T-2\}$, it follows from recursion (22) that we know $(R_{LBD}^*(e^*(t),t))_{t=0}^{T-1}$.

To conclude our identification argument in this part, we can now use (19) to determine the marginal social return of effort due to learning-by-doing at time t,

$$MSR_{LBD}^*(t) = \gamma^*(t)\sum_{\tau=1}^{T-t}\delta^\tau\mu^{\tau-1} = \gamma^*(t)\frac{\delta\left[1-(\delta\mu)^{T-t}\right]}{1-\delta\mu},$$

for all $t \in \{0,\ldots,T-1\}$. Since we know $(\gamma^*(t)\mu^{T-t-1})_{t=0}^{T-1}$, we can then use $MSR_{LBD}^*(T-2)$ to back out μ. Once we determine μ, we obtain the list $(\gamma^*(t))_{t=0}^{T-1}$.

5.3 Identifying the Production Function

We now discuss identification of the production function F. We begin by showing that the human capital production function F is non-parametrically non-identified.

First notice that by adding a constant to F we do not change the marginal products $(\gamma^*(t))_{t=0}^{T-1}$, and thus do not change the equilibrium efforts and piece-rates. Now notice that for each worker i,

$$\mathbb{E}[w_i(t)] = e^*(t) + m_\theta + k^*(t).$$

Iterating on the law of motion (2) for the stock of human capital, we obtain that

$$\mathbb{E}[w_i(t)] = e^*(t) + m_\theta + \sum_{s=1}^{t} \mu^{s-1} F(e^*(t-s), t-s) + \mu^t k(0). \tag{23}$$

In particular, if $G(e,t) = F(e,t) + c$, where c is a constant, then

$$\mathbb{E}[w_i(t)] = e^*(t) + \left(m_\theta + \frac{1}{1-\mu}c\right) + \sum_{s=1}^{t} \mu^{s-1} G(e^*(t-s), t-s) + \mu^t \left(k(0) - \frac{1}{1-\mu}c\right).$$

So, the production function G generates the same piece-rates and average wages as the production function F if we adjust the mean ability of workers and their initial stock of human capital accordingly. This establishes our claim.

We now show that we can identify $(e^*(t))_{t=0}^{T}$ and $(F(e^*(t),t))_{t=0}^{T-1}$ up to a vertical translation of F and a choice of $k(0)$. Since $(b^*(t))_{t=0}^{T}$, $(R^*_{CC}(t))_{t=0}^{T}$, and $(R^*_{LBD}(t))_{t=0}^{T}$ are known, we can use the first-order conditions for the equilibrium choices of efforts to determine $\lambda e^*(t)$ for all $t \in \{0,\dots,T\}$. In particular, we know the ratio $e^*(1)/e^*(0)$. On the other hand, given that

$$\mathbb{E}[w_i(1)] - \mathbb{E}[w_i(0)] = e^*(1) - e^*(0) - (1-\mu)k(0) + F(e^*(0), 0).$$

we also know the difference $e^*(1) - e^*(0)$ up to a normalization of $F(e^*(0),0)$ and a choice of $k(0)$. Thus, can determine $e^*(0)$ and $e^*(1)$, and from this determine λ and the remaining equilibrium efforts. To finish the identification argument, we can use (23) to compute $(F^*(e^*(t),t))_{t=1}^{T-1}$ recursively from the differences $\mathbb{E}[w_i(t)] - \mathbb{E}[w_i(t-1)]$, with $t \in \{2,\dots,T\}$.

To conclude this part, notice that if $F(e,t) = e^{\alpha(t)}\beta(t)$, then

$$e^*(t)\frac{\gamma^*(t)}{F(e^*(t),t)} = \alpha(t),$$

from which it follows that we can identify $(\alpha(t))_{t=0}^{T-1}$ and $(\beta(t))_{t=0}^{T-1}$ from $(\gamma^*(t))_{t=0}^{T-1}$, $(e^*(t))_{t=0}^{T-1}$, and $(F(e^*(t),t))_{t=0}^{T}$.

5.4 Remarks

We conclude our discussion of identification with a few remarks. First notice that we have identified the parameters σ_θ^2 and σ_ε^2 of the learning process and the product $r\lambda$ separately from the parameters $(\gamma^*(t))_{t=0}^{T-1}$ and μ of the learning-by-doing process. Since the parameters σ_θ^2 and σ_ε^2 and the product $r\lambda$ are sufficient to solve for the optimal piece-rates when learning-by-doing is absent, this means, loosely speaking, that the parameters $(\gamma^*(t))_{t=0}^{T-1}$ and μ are identified by matching $(b^*(t))_{t=0}^{T}$ to the piece-rates that would result in a world without learning-by-doing.

Also notice that we have much more information available than what is required for our identification argument, even when T is small. Indeed, for $T > 3$, the variance-covariance matrix of wages has substantially more moments than necessary to identify the piece-rates and the parameters of the learning process.

Other approaches in the literature rely on observations of productivity (or correlates of productivity) in addition to wages to identify learning models. See, for instance, Altonji and Pierret (2001), Lange (2007), Kahn and Lange (2014), Margiotta and Miller (2000), Gayle and Miller (2009, 2015), Gayle, Miller, and Golan (2015), and Pastorino (2016). We show how information on piece-rates can be used instead of information on productivity to identify learning models. The advantage of our approach is that this information can, in principle, be obtained from the second moments of the wage distribution. Of course, we do rely on strong functional form assumptions, but such assumptions are not uncommon in the literature.

Finally, models of human capital accumulation are also traditionally estimated using the information from the conditional mean of wages. We can see two consequences arising from this. First, one could test our current model using earnings profiles over the life-cycle. Second, one could use a structure such as ours to separately identify the contribution of on-the-job training from learning-by-doing, thus helping settle the long-standing question of whether returns to experience reflect on-the-job training or learning-by-doing.

6 Data

In this section we describe our data.

6.1 Sample

We focus on a PSID sample consisting of male heads of households, excluding individuals part of the poverty, latino, and immigrant sub-samples, with the following characteristics: are between 16 and 70 years old, worked more than 45 weeks in a year, have valid education information—more than 0 and up to 17 years of education, the largest value for the years of interest—have non-missing positive total labor income in the year, and are not self employed. We interpret an individual's labor income as the individual's *wage*. These selection criteria lead to more than 17,000 year-individual observations for the years between 1992 to 2012, since the last wave in our sample is for 2013.

We construct a measure of *performance or variable* pay given by the sum of earnings due to tips, bonus, and commissions, available from the 1993 survey onward. We drop observations for which performance pay is bigger than total labor income, for a total of 2 observations across all sample years. Accordingly, individuals who in a year do not report any tip, bonus, or commission have performance pay equal to zero. We deflate all income and wage variables using the BLS's consumer expenditure price index (PCE) so as to express them in 2009 US dollars.

Finally, we measure *experience* in the labor market, t, as potential experience, which we calculate as the difference between an individual's age and years of education according to the following equation: $t = \text{age} - \text{education} - 6 + 1$.[15]

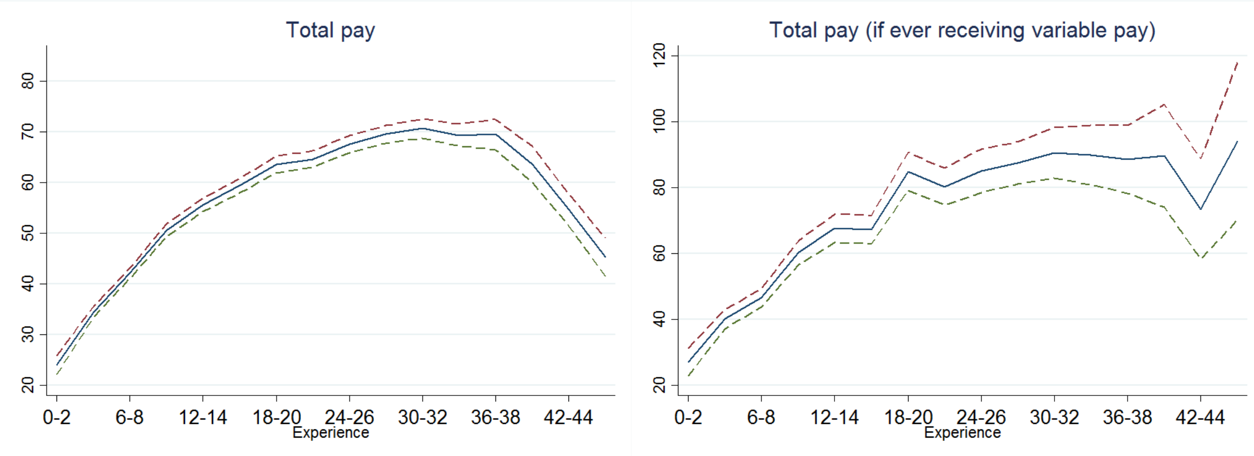
6.2 Descriptive Statistics

The mean age of individuals in our sample is of 40.40 years, with a standard deviation of 10.48 years. The mean educational attainment is 13.61 years (high school degree and some college), with a standard deviation of 2.19 years. To minimize the influence of possibly spurious extreme observations, we censor the distribution of total labor income by experience at the 2^{nd} and 98^{th} percentiles of the distribution. The resulting mean wage is 60.17 (in thousands of 2009 US dollars),

[15] In the few cases in which this calculation resulted in negative experience, we replaced the negative value by 1.

with a standard deviation of 38.82.

Figure 1: PSID Evidence on Life-Cycle Wages



The mean variable pay or wage, expressed in thousands of 2009 U.S. dollars, is 1.71, with a standard deviation of 16.44. For individuals ever receiving variable pay, the mean variable wage is 13.02, with a standard deviation of 43.72. As for (potential) experience, the sample mean is 21.80 years, with a standard deviation of 10.63 years—its distribution across years has a fairly wide range: the 25^{th}, 50^{th}, and 75^{th} percentiles of the experience distribution are, respectively, 13, 22, and 30 years. In terms of employment, only a few sectors of the economy absorb 11% or more of individuals: manufacturing, transportation, communication and utilities, whole and retail trade, and services. Most (at least 11%) of the individuals ever receiving variable pay are employed in these sectors or in finance, insurance, and real estate (FIRE).

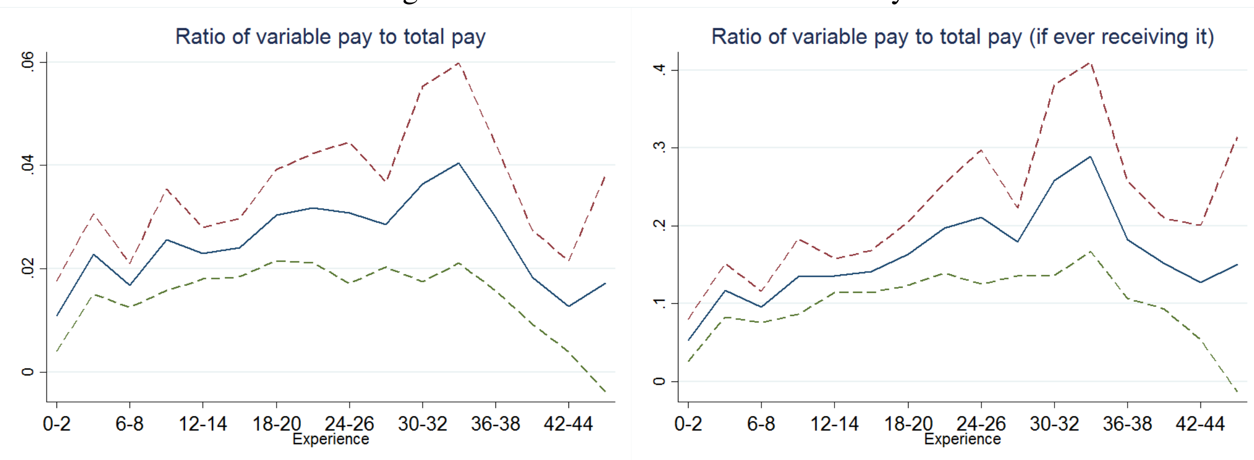
Following the identification argument of the previou section, we measure the sensitivity of pay to performance at experience level t, which corresponds to the variable b_t in our model, as the ratio of average real performance pay to average real labor income, based on the following equation:

$$w_t \equiv \underbrace{\text{fixed pay at } t}_{f_t} + \underbrace{\text{variable pay at } t}_{v_t}.$$

We graph the experience profile of wages, $\{w_t\}$, and performance pay, $\{b_t\}$, expressed in thousands of 2009 dollars, in Figures and 1 and 2 together with their associated 5% confidence intervals.[16] To this purpose, we group individuals in three-year experience groups.

[16] We compute these confidence intervals as $CI_t^x = x_\tau \pm 1.96 \times \sigma_{x_\tau}$, where $x \in \{w, b\}$ and $1 \leq t \leq 43$

Figure 2: PSID Evidence on Variable Pay



As expected, wages increase with experience at a decreasing rate: the wage profile is increasing until $t = 35$, and declining with experience thereafter; this pattern also holds for the wages of individuals who ever receive variable pay, even if the wages of these individuals tend to be higher. Interestingly, a similar pattern emerges for variable pay, whether we consider all observation-years—the left panel of Figure 2—or only observations on individuals ever receiving variable pay—the right panel of Figure 2. This evidence is at odds with the implications of standard models of career concerns, such as Gibbons and Murphy (1992), which predict that the sensitivity of pay to performance should increase with experience at *all* experience levels.

7 Estimation

In this section we detail how we estimate the model based on its discrete time version and present preliminary estimation results.

7.1 Details

In our estimation, we set $\delta = 0.95$ and $T = 43$ to allow for a long enough time window of observation on individuals and yet obtain a large enough sample for estimation to be feasible and estimates to be sufficiently precise. We estimate the remaining primitive parameters of our model, namely, σ_θ^2, σ_ε^2, r, γ, and μ, by deriving constructive estimators from the identification argument of Section 5. We proceed sequentially as follows.

First, we estimate the sequence $\{b_t\}_{t=1}^T$ of piece-rates as follows:

$$\widehat{b_t} = \frac{\widehat{\mathbb{E}_t(v_t)}}{\widehat{\mathbb{E}_t(w_t)}},$$

where $\widehat{\mathbb{E}_t(v_t)}$ and $\widehat{\mathbb{E}_t(w_t)}$ are, respectively, the sample means of variable pay and total pay at experience t.[17] The estimators of the remaining parameters are obtained from known transformations of $\widehat{\mathbb{E}_t(v_t)}$ and $\widehat{\mathbb{E}_t(w_t)}$, with $t \in \{1, T-2, T-1, T\}$, the empirical covariance of wages at $t = 1$ and $t = 2$, $\widehat{\mathrm{Cov}(w_1, w_2)}$, and the empirical variance of wages at $t = 1$, $\widehat{\mathrm{Var}[w_1]}$.

Specifically, we begin by estimating σ_θ^2 and σ_ε^2 as

$$\widehat{\sigma_\theta^2} = \frac{\widehat{\mathbb{E}_1(w_1)} \cdot \widehat{\mathrm{Cov}(w_1, w_2)}}{\widehat{\mathbb{E}_1(v_1)}},$$

where $\mathrm{Var}[m_1] = 0$, and

$$\widehat{\sigma_\varepsilon^2} = \frac{\widehat{\mathrm{Var}[w_1]}}{\widehat{b_1}^2} - \widehat{\sigma_\theta^2} = \frac{\widehat{\mathbb{E}_1(w_1)}}{\widehat{\mathbb{E}_1(v_1)}} \left[\frac{\widehat{\mathrm{Var}[w_1]} \cdot \widehat{\mathbb{E}_1(w_1)}}{\widehat{\mathbb{E}_1(v_1)}} - \widehat{\mathrm{Cov}(w_1, w_2)} \right],$$

respectively.

Following that, we estimate r as

$$\widehat{r} = \frac{1}{\widehat{\sigma_\varepsilon^2} + \widehat{\sigma_T^2}} \left(\frac{1}{\widehat{b_T}} - 1 \right),$$

where σ_T^2 is estimated as

$$\widehat{\sigma_T^2} = \frac{\widehat{\sigma_\varepsilon^2}\widehat{\sigma_\theta^2}}{\widehat{\sigma_\varepsilon^2} + (T-1)\widehat{\sigma_\theta^2}}.$$

We next estimate γ as

$$\widehat{\gamma} = \frac{\left[\widehat{b_{T-1}} + \frac{\delta(1 - \widehat{b_T})\widehat{\sigma_\theta^2}}{\widehat{\sigma_\varepsilon^2} + (T-1)\widehat{\sigma_\theta^2}} \right] \left[1 + \widehat{r}(\widehat{\sigma_{T-1}^2} + \widehat{\sigma_\varepsilon^2}) \right] - 1 + \delta \widehat{r} \widehat{\sigma_{T-1}^2} \widehat{b_T}}{\delta(1 - \widehat{b_T})},$$

where σ_{T-1}^2 is estimated as

$$\widehat{\sigma_{T-1}^2} = \frac{\widehat{\sigma_\varepsilon^2}\widehat{\sigma_\theta^2}}{\widehat{\sigma_\varepsilon^2} + (T-2)\widehat{\sigma_\theta^2}}.$$

[17]Note that the estimator of each $\widehat{b_t}$ inherits the randomness of $\widehat{\mathbb{E}_t(v_t)}$ and $\widehat{\mathbb{E}_t(w_t)}$.

Lastly, we estimate μ as

$$\widehat{\mu} = \frac{\widehat{B_{T-2}}[1 + \widehat{r}(\widehat{\sigma_\varepsilon^2} + \widehat{\sigma_{T-2}^2})] + \delta\widehat{r}\widehat{\sigma_{T-2}^2}\left(\widehat{B_{T-1}} + \delta\widehat{b_T}\right) - 1 - \delta\widehat{\gamma}(1 - \widehat{B_{T-1}})}{\delta^2\widehat{\gamma}(1 - \widehat{b_T})},$$

where σ_{T-2}^2 is estimated as

$$\widehat{\sigma_{T-2}^2} = \frac{\widehat{\sigma_\varepsilon^2}\widehat{\sigma_\theta^2}}{\widehat{\sigma_\varepsilon^2} + (T-3)\widehat{\sigma_\theta^2}},$$

and $\widehat{B_{T-1}}$ and $\widehat{B_{T-2}}$ are auxiliary estimated parameters given by

$$\widehat{B_{T-1}} \equiv \widehat{b_{T-1}} + \frac{\delta(1 - \widehat{b_T})\widehat{\sigma_\theta^2}}{\widehat{\sigma_\varepsilon^2} + (T-1)\widehat{\sigma_\theta^2}}$$

and

$$\widehat{B_{T-2}} \equiv \widehat{b_{T-2}} + \frac{\delta(1 - \widehat{b_{T-1}})\widehat{\sigma_\theta^2}}{\widehat{\sigma_\varepsilon^2} + (T-2)\widehat{\sigma_\theta^2}} + \frac{\delta^2(1 - \widehat{b_T})\widehat{\sigma_1^2}}{\widehat{\sigma_\varepsilon^2} + (T-1)\widehat{\sigma_\theta^2}}.$$

7.2 Estimation Results

Our preliminary estimation results are as follows. Our estimates of the variance of the distribution of ability and output noise are, respectively, $\sigma_\theta^2 = 4,438.598$ and $\sigma_\varepsilon^2 = 400,894.100$. Our estimate of the product of the coefficient of absolute risk aversion and the parameter λ is $r\lambda = 0.00033$. Our estimate of the rate of human capital acquisition is $\gamma = 2.998$. Finally, our estimate of the rate of human capital depreciation is $\mu = 0.145$. Notice that all estimates satisfy the positivity restrictions implied by our model, even though we did not impose any such restriction in our estimation procedure.

To interpret these results in terms of the learning parameters observe that the implied signal-to-noise ratio ($\sigma_\theta^2/\sigma_\varepsilon^2$) is 0.011. This implies both a high degree of uncertainty about ability and a slow degree of learning. Hence, uncertainty about ability is large and persistent over time. If we calibrate a degree of relative risk aversion of 1, the implied value of λ is 0.018. This, in turn, implies, at the estimated values of the other parameters, a wage growth of 207.63% according to our model. This wage growth is remarkably close to the one we measure in the data, 198.58%. We obtain similar results for other plausible values of the risk aversion parameter, say, between 1 and 5, but the closest fit to the data in terms of life-cycle wage growth is indeed for a value of 1.

We also evaluate model fit in terms of the predicted profile of piece-rates. The model matches this profile qualitatively. It also matches quantitatively well its decline towards the end of the working life cycle. Specifically, both in the model and in the data the profile $\{b_t\}$ of piece-rates is hump-shaped with experience: it peaks at experience $t = 35$ in the data and at experience $t = 39$ in the model. The peak to end-of-life trough change in $\{b_t\}$ is -3% in data and -2% in model. The presence of human capital acquisition is critical for the shape of this profile: without it, the $\{b_t\}$ profile would be increasing throughout the working life cycle.

8 Conclusion

We propose a model that nests employer learning, learning-by-doing, and incentives for performance to account for the life-cycle profile of individual wages. We characterize the optimal wage contract in this framework and provide a recursive characterization of the pay-for-performance component of the optimal wage contract that allows us to decompose this component into four distinct elements, which can be readily interpreted in terms of basic life-cycle considerations: the standard trade-off between risk and incentives, the desire to hedge against the risk inherent in learning about ability, and the changing strength of the career concerns and human capital investment motives over time. We show that the model's primitives can be recovered from cross-sections of wages and experience only, whether from standard surveys or from firm-level records, without requiring specific information on the pay-for-performance component of wages or measures of individual investments and effort over the life-cycle. Our preliminary estimation results imply that learning about ability, human capital acquisition, and incentives are all key determinants of the life-cycle profile of wages as well as individuals' overall wage growth.

References

[1] Gary S. Becker. Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5, Part 2):9–49, 1962.

[2] Yoram Ben-Porath. The production of human capital and the life cycle of earnings. *Journal of Political Economy*, 75(4, Part 1):352–365, 1967.

[3] George-Levi Gayle and Robert A. Miller. Has moral hazard become a more important factor in managerial compensation? *American Economic Review*, 99(5):1740–1769, 2009.

[4] George-Levi Gayle and Robert A. Miller. Identifying and testing models of managerial compensation. *The Review of Economic Studies*, 82(3):1074–1118, 2015.

[5] Robert Gibbons and Kevin J. Murphy. Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy*, 100(3):468–505, 1992.

[6] Limor Golan, George-Levi Gayle, and Robert A. Miller. Promotion, turnover and compensation in the executive labor market. *Econometrica*, 83(6):2293–2369, 2015.

[7] Bengt Holmstrom. Moral hazard and observability. *The Bell Journal of Economics*, pages 74–91, 1979.

[8] Bengt Holmstrom. Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169–182, 1999.

[9] Mary Margiotta and Robert A. Miller. Managerial compensation and the cost of moral hazard. *International Economic Review*, 41(3):669–719, 2000.

[10] Jacob A. Mincer et al. Schooling, experience, and earnings. *NBER Books*, 1974.