

Incentives to Work or Incentives to Quit?

Raicho Bojilov*

15 February 2015

Abstract

Pay incentives affect not only effort choice but also turnover and thereby the quality mix of the workforce. This paper investigates how considerations about the quality mix shape pay policy and profits within a structural model of effort choice, symmetric learning about match quality, and turnover. Using unique data from a call centre in South Carolina, I estimate the model in two steps adapting estimation methods for dynamic structural models to the analysis of employment dynamics. Then, I consider three classes of contracts: (1) compensation depends only on current output; (2) compensation depends also on past output; (3) compensation depends on all available information and may vary with tenure. The results indicate that experimentation to improve the quality mix is a primary concern that affects the optimal contracts in the three classes. Experimentation requires high turnover which is also associated with the destruction of accumulated specific human capital. The trade-off between experimentation and the accumulation of specific capital determines the characteristics of the optimal contracts.

KEYWORDS: Piece Rates, Learning, Turnover

JEL CODE: J22, J33, D82, D83, M51, M52

*Raicho Bojilov, Department of Economics; Ecole Polytechnique, Route de Saclay; 91128 Palaiseau, France. E-mail: raicho.bojilov@polytechnique.edu. I would like to thank V. J. Hotz, B. MacLeod, L. Munasinghe, B. O’Flaherty, J.-M. Robin, F. P. Vinay, participants at the Paris Search and Matching conference, and seminar participants at CREST-ENSAE for their helpful comments and suggestions. I am greatly indebted to P.A. Chiappori and B. Salanié for their advice and support in my work on this project.

1 Introduction

Employees and firms often learn about the quality of their match over time and this learning influences separation decisions. As long as at least some of the worker's specific match quality or ability is non-exportable to other firms, the employer may capture at least some of the surplus from the relation. Matters become even more interesting when one considers the interplay between pay incentives and turnover: in some cases, the firm may be tempted to encourage a worker of average or slightly better than average match quality to quit, so that his or her position is taken by someone whose match quality is unknown and, therefore, potentially higher. A firm in such a context faces two related questions: What contract should it offer to its employees? and How should the firm balance between the competing concerns of experimentation with new workers and retaining existing employees of high match quality? These issues suggest that pay incentives affect profits not only through their impact on effort choice but also through their effect on the quality mix of the workforce.

In this paper, I investigate the relative importance of the two channels to maximizing profits in a structural model of employment dynamics that includes effort choice, learning about match quality, and labor turnover. My empirical work focuses on the optimal contract in such an environment. However, I also study contracts that are linear in output: compensation is equal to the sum of a base pay and a bonus proportional to hourly output (performance from now on). I do so for two reasons. First, the firm whose personnel records I use itself implemented such linear contracts and one of the objectives of this paper is to characterize the profitability of the firm's compensation policies. Second, firms often apply simple compensation policies based on such linear contracts and the problem of finding and characterizing the optimal linear contract is of interest on its own.¹

The firm's data are ideally suited for the empirical analysis of pay incentives; they come

¹Within a dynamic setting as the one in this paper, the firm's flexibility in designing the contract is reduced. This general idea is first explored by Holmstrom and Milgrom (1987) who show that in some settings the optimal compensation is to provide workers with incentives that are linear in output. Besides, the cost of implementing a complicated nonlinear contract is high: see, for example, Holmstrom and Milgrom (1990) as well as Ferrall and Shearer (1999).

from a call center in South Carolina and contain an objective measure of individual performance (defined as output per hour), a known compensation policy based on linear contracts in performance, and a variation in the pay policies that does not depend on what the firm learns about its employees. My estimates show that steeper incentives are associated with higher performance, that persistent differences in individual performance are driven by differences in the quality of the employer-employee match, and that employees learn about the quality of the match on the job. Their posterior beliefs are largely responsible for their decision to stay or quit and the interaction between incentives and turnover appears to be crucial to evaluating the impact of pay incentives on profits. Thus, I conclude that finding the optimal pay policy requires the explicit modelling of all three: effort choice, learning about match quality, and separation decisions.

Unobserved effort, labor turnover and learning about match quality are the subject of intensive study in the structural literature but usually separately from one another. For example, Shearer and Paarsch (2009) analyze the effect of incentives on effort and conduct a related policy analysis, but the experimental design of their study does not allow them to analyze the effects of incentives on the pool of entering employees and turnover. However, Lazear (2000) points out that in the context of his study about one-third of the improvement in performance after the introduction of pay incentives can be traced back to the improvement in the quality mix of entering employees. When the quality of the match between a potential employee and the firm becomes known to the worker in the hiring process, the firm may use its pay policy not only to induce effort but also to shape the quality mix of the newly hired employees, as discussed in Lazear (1998); the firm also faces a trade-off between the extra revenue generated by workers who stay and the associated increase in pay that is necessary to make them stay. The interaction of all these considerations determine the firm's pay policy and the associated turnover is not necessarily low or nonexistent but depends on the characteristics of the technology, the workforce and the alternative jobs. When the firm and workers learn about match quality over time, an additional consideration arises. Depending on how much starting employees know about their match, turnover becomes the primary

channel through which pay incentives affect the quality mix. In the special case of a common prior, there is no selection at entry, and turnover is the only source of changes in the quality mix at the workplace.

Similarly, the models in the structural literature on learning about match quality and turnover do not incorporate effort choice. The dynamic programs in such models are complicated enough even as they are, since they involve heterogeneity across employees and a sequence of posterior beliefs. To my knowledge, Miller (1984) is the first paper to estimate a model with learning about match quality. Pastorino (2009) considers a variation of this model that incorporates correlation between ability at one job and ability at others. Furthermore, Nagypal (2007) applies the method of indirect inference to distinguish between learning about match quality and learning-by-doing within a structural setting. More recently, Camargo and Pastorino (2010) estimate a structural model of career concerns with learning-by-doing. While these models consider selection and job mobility, they assume away potential problems of moral hazard. Furthermore, the computational complexity associated with estimating the models usually requires some strong assumptions about the production technology and the heterogeneity among workers.

This paper is closest to Pugatch (2012) who studies the decision to re-enter school in a model of educational and career choice in South Africa. To my knowledge, he is the first to introduce Bayesian learning in a dynamic search model of schooling and career choice. His estimation method follows in the tradition of Hotz and Miller (1993) and relies heavily on the generalization of this approach to cases of unobserved heterogeneity discussed in Arcidiacono and Miller (2011). His estimation strategy is, however, suited to models with discrete binary signals about the unobserved parameter of interest.

Here, I propose a very simple two-step procedure to estimate a structural model of learning about match quality, effort choice, and turnover when the observed signal and match quality are normally distributed. In the first step, I estimate a semi-structural attrition model and recover the stochastic technology up to a constant, as well as a scaled version of the value of continued employment. Bojilov (2011) shows that the first-step provides consistent estimates

of the effect of changes in pay incentives on effort, while popular alternatives may overestimate the same effect by a factor of two. I use these estimates in the second step to recover the remaining parameters using the method of moments. That is, I use an indirect approach to estimate the value of continued employment without directly solving for the value function. The estimates of the model are then used as a basis for counterfactual policy analysis.

The results suggest that firms choose their pay policy for reasons that go beyond effort choice. Most of the increase in profits from switching to the optimal policy from hourly wage can be traced back to the effect of incentives on the quality mix. The optimal linear contract induces low quality employees to quit and in this way it helps the firm build a workforce of high match quality over time. This effect more than offsets the loss associated with replacing an experienced worker with a newly hired one of no experience and unknown ability. Furthermore, the employer exploits the firm-specific nature of match quality to capture most of the surplus generated by the employment relation. To achieve that, the firm offers pay incentives that induces little effort, so high level of effort and low turnover are not necessarily attributes of the profit-maximizing pay policy. Finally, the optimal policy in linear contracts that I find generates only 4.8% higher profits than one of the actually implemented pay regimes. An exercise in comparative statics shows that as turnover costs grow, the firm increases compensation to induce lower turnover by offering much steeper incentives. The result is a decline in the quality mix of the workforce and an increase in the importance of effort to the firm's profits. Another counterfactual experiment shows that the firm's profits would have been 27% higher if match quality was known to the employees at the time of hiring. The primary reason is that workers of high match quality self-select into the firm which leads to low turnover and high level of experience.

In addition, I also consider history dependent contracts. In particular, the structure of the learning process implies that the average of the past signals about match quality is a sufficient statistic for the mean of the posterior beliefs. For this reason, I focus on nonlinear contracts that incorporate the average of the past signals. Since the informativeness of the average increases with tenure, it is natural to allow also for variation in the compensation

schedule with tenure. Thus, I conclude the counterfactual experiments by studying a tenure- and history-dependent nonlinear contracts. The results show that switching from the optimal linear contract to the optimal tenure-dependent nonlinear contract in past signals leads to an increase of 56 % in profits. This increase can be traced back to the greater flexibility of the tenure and history dependent contract which allows the firm to provide the maximal incentives to the worker to exert effort while also capturing a greater share of the surplus associated with match quality.

The results indicate that incentives matter and in this way they are consistent with the literature on investigating incentive effects represented by Paarsch and Shearer (2000), Lazear (2000) and Shearer (2004). More specifically, they show that workers are very responsive to changes in the slope of incentives. This is consistent with previous results obtained in Paarsch and Shearer (1999, 2009), as well as Haley (2003). The novelty here is that optimal pay incentives are allowed to affect not only effort choice but also the composition of the workforce at different tenure horizons through turnover. Thus, the paper extends the work in Lazear (1998, 2000) on the effect of incentives on the quality mix by studying how turnover shapes the properties of the optimal pay policy. In particular, the results show that turnover may be the primary channel through which pay incentives affect profits when workers learn about match quality on the job. They also indicate that the considerable contribution of improved match quality to worker's compensation that is estimated in some structural papers of job mobility, such as Hoffmann (2010), may depend on the strong assumptions of low or non-existent turnover costs.

The rest of the paper is organized as follows. Section 2 presents the model and section 3 the data. Section 4 introduces the estimation of the model. Section 5 discusses the estimates of the structural parameters and presents the policy analysis. Section 6 concludes.

2 Model

The model is a variation on the standard model of search by experience, first introduced in Jovanovic (1979): workers choose not only whether to stay or quit, but also how much effort to exert. The crucial element in the model is match quality θ_i that is unknown at the time of hiring and both the employee and the employer learn symmetrically about its value over time through a sequence of noisy performance signals y_{it} . In contrast to other search models, the compensation in this setting is not pinned down by competition among firms for exportable match quality: the compensation schedule is endogenously determined by the firm to maximize its profits.

2.1 Worker's Problem

At the beginning of the employment relation, the worker's prior belief are represented by random variable θ_{i1} . At the beginning of each period t , she receives an outside offer ξ_{it} , and decides to stay if the value of continued employment is greater than the outside offer. If the worker decides to stay, she chooses a level of effort l_{it} , that is not observable or verifiable by the firm. Then she receives an output signal y_{it} governed by the following stochastic technology:

$$y_{it} = \theta_i + g(t) + l_{it} + \varepsilon_{it}$$

where ε_{it} is iid over time and across individuals, θ_i is independent of the error process, and $g(t)$ represents the accumulation of firm-specific knowledge.² The strong separability of the technology in effort and ability is a crucial element of the model. However, Bojilov (2011) shows that this functional form restriction is consistent with the observed patterns in the performance data used in this study. More details on the choice of functional form for the technology can be found in Appendix A.

²Jovanovic and Nyarko (1994) provide an alternative specification for the accumulation of firm-specific knowledge, which is sometimes referred to as learning-by-doing. However, the data do not support the prediction of their model that the variance of individual performance declines over time. Note that the model also assumes that the accumulation of knowledge does not depend on past or present effort.

After observing the output signal y_{it} the worker is paid $w_{it} = \alpha_{it} + \beta_{it}y_{it}$, according to the piece rate $R_{it} = (\alpha_{it}, \beta_{it})'$. Regime R_{it} is said to be more generous than regime R'_{it} , $R_{it} > R'_{it}$, if both $\alpha_{it} > \alpha'_{it}$ and $\beta_{it} > \beta'_{it}$. Workers are assumed to be risk-neutral with a separable utility function:

$$u(R_{it}, l_{it}, y_{it}) = \alpha_{it} + \beta_{it}y_{it} - \psi(l_{it}).^3$$

Employees do not expect the piece rate to change in the future.⁴ The cdf of the posterior belief θ_{it} is a function of the cdf of the initial prior θ_{i1} and the noisy signals about θ_i up to period t , $\{y_{ik} - g(k) + l_{ik}\}_{k=1}^t$. Let μ_{it} denote the mean of the posterior belief. Since θ_i and l_{it} enter additively in the utility function, optimal effort choice does not depend on θ_{it} and is function of R_{it} only, $l(R_{it})$. Then, the worker's problem can be represented by the following optimal stopping problem

$$V(\theta_{it}, R_{it}, t) = \int \max \left[\xi_{it}, E_{\theta_{it}}(u(R_{it}, l_{it}, y_{it})) + \delta \int_{\Theta} V(\theta_{it+1}, R_{it}, t+1) dF(\theta_{it+1}|\theta_{it}) \right] dF_{\xi},$$

where $E_{\theta_{it}}$ indicates that expectation is taken with respect to the cdf of θ_{it} . $F(\theta_{it+1}|\theta_{it})$ is the transitional kernel over the set of possible beliefs at the beginning of period $t+1$ based on the information at t and F_{ξ} is the cdf of outside offers. Let

$$G(\theta_{it}, R_{it}, t, X_{it}) = E_{\theta_{it}}(u(R_{it}, l_{it}, y_{it})) + \delta \int_{\Theta} \int_{\Xi} \max[\xi_{it}, G(\theta_{it+1}, R_{it}, t+1)] dF(\theta_{it+1}|\theta_{it}) dF_{\xi}$$

After the realization of the outside offer, i decides to stay if the value of continued employment is higher than the value of the outside offer:

$$G(\theta_{it}, R_{it}, t, X_{it}) - \xi_{it} > 0$$

The assumptions of this general model, the existence of a solution to the worker's problem

⁴While at first sight this assumption may appear restrictive, Kanemoto and MacLeod (1992) show that when firms learn about individual ability the existence of an outside option for workers disciplines firms to keep piece rates fixed even after beliefs are updated.

and its characterization are presented in Appendix B. Under assumptions AA1 to AA4 in the appendix, if R_{it} is more generous than R'_{it} optimal effort $l(R_{it}) > l(R'_{it})$ and $G(\theta_{it}, R_{it}, t) > G(\theta_{it}, R'_{it}, t)$. Furthermore, the value of continued employment increases in θ_{it} in the sense of the likelihood ratio property.

For the purposes of counterfactual policy analysis, I impose the following parametric form on utility:

$$u(R_{it}, l_{it}, y_{it}) = \alpha_{it} + \beta_{it}y_{it} - \frac{\gamma}{1 + \frac{1}{\psi}}l^{1 + \frac{1}{\psi}}.$$

This specification for the disutility of labor is popular in the related literature; for example it is used in Shearer (2004) and Paarsch and Shearer (2009). Here ψ is the elasticity of effort to the slope of incentives β . Optimal effort is then

$$l_{it} = \left(\frac{\beta_{it}}{\gamma} \right)^\psi$$

Intuitively, this assumption about the functional form of the utility implies that conditional on one's ability, output is proportionate to β^ψ .

Furthermore, for the purposes of forming tractable sequence of posterior beliefs, I maintain the following assumption:

Assumption A (i) $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$ and $\xi_{it} \sim N(\mu_\xi, \sigma_\xi^2)$ are iid across tenure horizons and individuals, independent from the rest of the covariates. (ii) $\theta_i \sim N(\theta, \sigma_\theta^2)$ is iid across individuals and is independent from the rest of the covariates.

Under the assumption above, the posterior belief θ_{it} is also normally distributed for all t : $\theta_{it} \sim N(\mu_{it}, \sigma_t^2)$, where for $t > 1$

$$\mu_{it} = (1 - K_t)\mu_{it-1} + K_t(y_{it-1} - l(R_{it}) - g(t-1))$$

$$\sigma_t^2 = \frac{\sigma_\varepsilon^2 \sigma_\theta^2}{\sigma_\theta^2(t-1) + \sigma_\varepsilon^2}$$

$$K_t = \frac{\sigma_\theta^2}{\sigma_\theta^2(t-1) + \sigma_\varepsilon^2}$$

Thus, Bayesian updating becomes quite tractable. In particular, the precision of beliefs depends only on t , so the average of the demeaned past signals is a sufficient statistic to characterize posterior beliefs. The equation for the posterior mean can be solved recursively to arrive at:

$$\mu_{it} = (t-1)K_t \cdot \left(\frac{1}{t-1} \sum_k^{t-1} (y_{ik} - l(R_{it}) - g(k)) \right) + (1 - (t-1)K_t) \mu_{i1}.$$

K_t represents the precision that the worker attaches to the average past signals in forming her beliefs about her match quality. Since $(t-1)K_t$ increases over time, she attaches greater and greater weight to the average of the demeaned past performance and less to the mean of the initial belief.

Then, the expected utility from working in period t is

$$U(\mu_{it}, R, t, \gamma, \psi) = \alpha_{it} + \beta_{it} \left(\mu_{it} + \left(\frac{\beta_{it}}{\gamma} \right)^\psi + g(t) \right) - \frac{\gamma}{1 + \frac{1}{\psi}} \left(\frac{\beta_{it}}{\gamma} \right)^{\psi+1}$$

By these observations, the independence of θ_i from ε_{it} and ξ_{it}^* , and by the additivity of ε_{it} in the stochastic technology, the optimal problem of the worker can be formulated as functional equation (P)

$$v(\mu_{it}, R_{it}, t) = \int [\max(\xi_{it}, U(\mu_{it}, R, t, \gamma, \psi)) + \delta \int v(\mu_{it+1}, R_{it}, t+1) f(\mu_{it+1}|\mu_{it}, t) d\mu_{it+1}] f_\xi(\xi_{it}) d\xi_{it},$$

where $f(\mu_{it+1}|\mu_{it}, t)$ is the conditional density of μ_{it+1} , given μ_{it} and t .

Proposition 1. *Given the specification of the model above*

i. *The functional equation (P) has a unique continuous solution $V(\mu_{it}, R_{it}, t)$ and the optimal policy*

$$A(\mu_{it}, R_{it}, t) = \{l_t \in L \mid (P) \text{ holds.}\}$$

is a continuous function.

ii. Optimal effort $l(R_{it}) > l(R'_{it})$ if $R_{it} > R'_{it}$.

iii. $V(\mu_{it}, R_{it}, t) > V(\mu_{it}, R'_{it}, t)$ if $R_{it} > R'_{it}$, and $V(\mu_{it}, R_{it}, t)$ increases μ_{it} .

The proof of Proposition 1 is presented in Appendix B. Let

$$H(\mu_{it}, R_{it}, t) = U(\mu_{it}, R_{it}, t, \gamma, \psi) + \delta \int \int \max\{\xi_{it+1}^*, H(\mu_{it+1}, R_{it}, t+1)\} \varphi(\mu_{it+1}|\mu_{it}, t) f_{\xi}(\xi_{it+1}) d\mu_{it+1} d\xi_{it+1}$$

After the realization of the outside offer, i decides to stay if the value of continued employment is higher than the value of the outside offer

$$H(\mu_{it}, R_{it}, t) - \xi_{it} > 0.$$

2.2 Firm's Problem

Based on the firm's records, I take the revenue from a successfully processed call to be $r = \$8.5$. This approximation is based on the firm records for average outbound and inbound calls, the reward that the firm receives from processing each type of calls, and the relation between the number of processed calls and accounts serviced by the company.⁵ I assume that inbound and outbound calls, as well as the number of processed calls per account is independent from the implemented contract. Furthermore, I assume that the firm's monthly discount factor is $\tilde{\delta} = 0.99$, implying an annual discount factor of just below 0.9. Quitting disrupts the production process and necessitates spending money to advertise the available job position, and train the replacement. In what follows, I incorporate turnover costs, which according to some estimates of the firm itself amount to approximately \$750. Furthermore, I also allow the firm to hire a replacement immediately after a worker quits. The firm is assumed to face

⁵The actual contract between the call center and the cable TV company was stated in more complicated terms. The cable TV company transferred accounts to the call center after the latter had successfully processed previously transferred calls. Thus, the call center made its profits from successfully processing accounts; the cable TV company expected more than 95% rate of collection. Yet, the contract recognized that not all attempts to contact a cable TV subscriber are successful, so it conditioned pay per account on the inbound and outbound calls that operators make to the client. Nevertheless, the underlying factor that drives profits is the successful collection of debt because that leads to the transfer of more accounts. The approximation establishes a relation between the processed calls and serviced accounts.

constant returns to scale. Finally, the profit function that I consider below abstracts away from fixed costs.

Given these assumptions, the expected profits per employee in period t , conditional on staying, match quality, t , and the pay policy, are defined as

$$\pi_{it}(\theta_i, R, t) = (r - \beta) \cdot (\theta_i + l(R) + g(t)) - \alpha,$$

where r is the revenue per call. Let the probability of staying at least until period t be $p_{it}(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}^*\}_{k=1}^{t-1})$. The expected profits of the firm are:

$$\begin{aligned} \pi(R) = E_\theta \{ & \sum_{t=1}^{\infty} \tilde{\delta}^{t-1} [p_{it}(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}^*\}_{k=1}^{t-1}) \pi_{it}(\theta_i, R, t) + \\ & + (1 - p_{it}(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}^*\}_{k=1}^{t-1})) (\pi(R) - c)] \}. \end{aligned}$$

The expectation operator indicates that the expectation is taken with respect to the initial prior belief, which coincides with the distribution of match quality in the population. Each period, the employee either stays with probability $p_{it}(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}^*\}_{k=1}^{t-1})$ and generates profits $\pi_{it}(\theta_i, R, t)$ or quits and the firm hires a new employee who at entry is expected to generate exactly the same profits as the original employee, $\pi(R)$. This equation can be solved for $\pi(R)$

$$\begin{aligned} & \pi(R) \\ = E_\theta \left\{ & \sum_{t=1}^{\infty} \tilde{\delta}^{t-1} \frac{p_{it}(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}^*\}_{k=1}^{t-1}) \pi_{it}(\theta_i, R, t) - (1 - p_{it}(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}^*\}_{k=1}^{t-1})) c}{1 - \sum_{t=1}^{\infty} \tilde{\delta}^{t-1} (1 - p_{it}(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}^*\}_{k=1}^{t-1}))} \right\} \end{aligned}$$

The firm chooses (α, β) to maximize $\pi(R)$ subject to

$$l(R) = \left(\frac{\beta}{\gamma} \right)^\psi.$$

The probability of staying $p_{it} \left(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}\}_{k=1}^{t-1} \right)$ and the optimal effort conditions connect the firm's problem to that of the worker presented above.

The structure of the problem implies that when the firm considers linear contracts it faces a trade-off between providing more incentives to exert effort and shaping the quality mix through turnover. This trade-off is a particular feature of linear contracts and is, therefore, an important consideration in practice given the prevalence of linear contracts in the real world. In the context of history- and tenure-dependent contracts, it disappears. In fact, the policy analysis is simplified by the risk-neutrality of both the firm and the worker, since then it is optimal for the firm to sell the contract. Given the linear technology, this implies that all workers will be given 100 per cent of the output that is due to effort and the problem boils down to finding how the 'price' offered by the firm varies with match quality given the trade off between experimentation and keeping high quality workers.

3 Data

The data set contains a clean performance measure and three known compensation regimes that were implemented in a way that allows to identify each one's effect on performance. It comes from a call center in North Carolina owned and operated by a multinational company, which center collects outstanding debt and fees on behalf of cable TV companies, which ensures a stable demand for its services. An automated switchboard operator allocates inbound and outbound calls, so that the longest weighting customer is matched with the longest weighting operator. Employees rotate their work stations on a daily basis.

As part of a reorganization plan, the central management implemented a linear contract at the beginning of January 2005: a linear function of the performance metric, , the number of calls per hour that end with collection of the outstanding debt. Before January 2005, compensation was based on an hourly wage of \$9.5. The central management was concerned that the company was paying "too much," so it implemented a new regime for the newly-hired employees in June 2005 (regime 2). Relative to regime 1, regime 2 offered a lower

base pay, decreased the slope of the piece rate for those with performance less than 3.8, and increased the slope of the piece rate for those with performance greater than 3.8 (regime 2). All previously hired employees continued to be paid according to regime 1. Since the central management was worried about possible negative effects of the piece rate on the quality of service, it changed the pay regime yet again in November 2005. The new regime 3 had two components: all employees were paid according to the pay schedule of regime 2, but in addition employees had to meet certain minimum quality standards of service to qualify for the piece rate. Twenty per cent of one's calls were randomly monitored and the quality of service was rated on a scale from 0 to 100. An employee who did not meet the minimum quality standard was relegated to an hourly wage equal to the base pay of the piece rate. Since 99% of performance lies between 1.05 and 3.8, regimes 2 and 3 effectively lowered incentives relative to regime 1. Diagram 1 shows a time line for the implementation of the three regimes.

Bojilov (2011) provides a detailed descriptive analysis. What follows summarizes only the most relevant pieces of this analysis. The call center experienced high turnover rates under all pay regimes: more than 50% of all employees under regime 1 quit within the first six months of employment, while under regimes 2 and 3 the turnover for the first six months approached 67%. There also appears to be a noisy downward trend in the separation rates as tenure increases. This noisiness is probably due to the small sample size, but it also suggests that separation decisions depend to a large extent on individual-specific factors. Table 1 reports the average performance for the first six months of employment across regimes. Again, as one may expect, the average performance under regime 1 is higher than its counterparts for regimes 2 and 3. Furthermore, the average performance on the subset of stayers is higher than the simple average, suggesting that poor performers quit. This evidence suggests that steep pay incentives lead to high performance; that attrition appears to be non-random, since workers with higher performance are more likely to stay; that individual-specific effects are present; and finally that workers accumulate experience or knowledge in the course of their first six months of employment. Summary statistics and some functional form nonparametric tests that are relevant to this paper are presented in Appendix A.

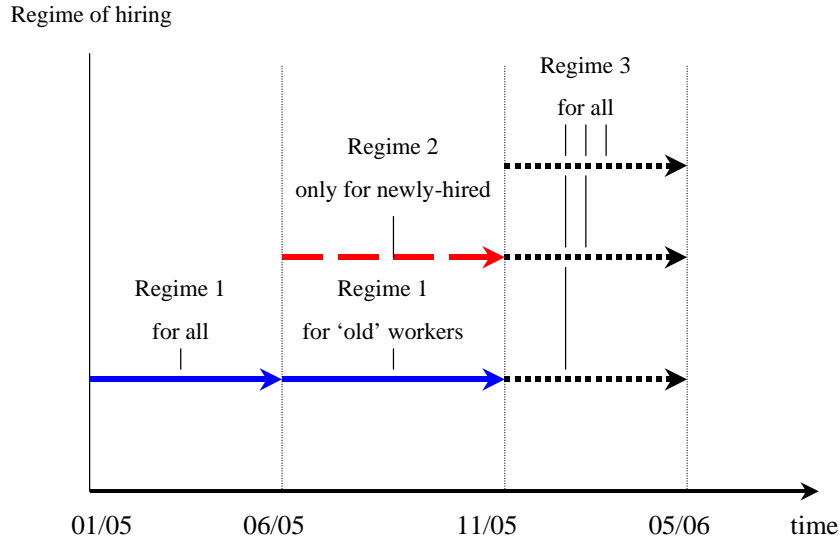


Diagram 1. Timeline of pay regimes.

4 Estimation

Moral hazard, Bayesian learning and labor turnover have been studied intensively but usually separately in the literature on structural estimation. For example, Shearer (2004) and Shearer and Paarsch (2009) analyze the effect of incentives on performance and conduct a related policy analysis but the context of their study allows them to assume away issues related to turnover. Still, moral hazard and labor turnover are defining features of the analytical environment at most workplaces; their interaction shapes employment outcomes and through them profits and individual welfare. The estimation of a structural model including moral hazard, Bayesian learning, and labor turnover is, however, a complicated exercise. To my knowledge, Nagypal (2007) is the only recent paper that estimates a structural model of labor turnover and Bayesian learning in order to distinguish between learning-by-doing and learning about match quality. The dynamic programs in such models are quite complicated, since they involve posterior beliefs about an unobserved individual-specific parameter. As a result, estimation methods that rely on solving for the value function at each step of the optimization algorithm are computationally intensive. Nagypal uses indirect inference to estimate her model, but even this method is not without its problems. In particular, discreteness of some

of the dependent variables still posits a challenge. A small change in the structural parameters leads to a discrete jump in the value of the discrete choice variables and in turn a discrete jump in the objective function. Nagypal addresses these issues by applying the simplex algorithm in the optimization stage, but such a solution is viable only when the number of structural parameters is small. While a number of authors have suggested solutions based on "smoothing" the discrete variables, there is by no means consensus on how to address these issues.⁶

Here, I propose a simple two-step procedure to estimate the structural model incorporating effort choice, learning about match quality, and separation decisions. In principle, the structural parameters can be recovered by estimating the following model

$$y_{it} = \theta_i + l(R_{it}) + g(t) + \varepsilon_{it}, \text{ if for all } k = 1, \dots, t - 1$$

$$s_{ik} = 1 [G(\theta_{ik}, R_{it}, k) - \xi_{ik}^* > 0]$$

Doing so, however, involves solving for the value function of each individual for each belief at each step of the optimization algorithm, which is computationally intensive. Therefore, in practice I estimate the model in two steps. In the first step, I estimate a semi-structural attrition model and recover the stochastic technology up to a constant, as well as a scaled version of the net expected utility of continued employment. I use these estimates in the second step to estimate the remaining structural parameters using the method of minimum distance estimation.

Step 1 In the first step, I estimate the following semi-structural model:

$$y_{it} = \theta_i + l(R_{it}) + g(t) + \varepsilon_{it}, \text{ if for all } k = 1, \dots, t - 1$$

$$s_{ik} = 1 [H(\theta_{ik}, R_{it}, k) - \xi_{ik} > 0],$$

⁶Smith (2003) provides a summary of the econometric challenges and outlines a smoothing approach that addresses them.

where $H(\theta_{ik}, R_{ik}, k)$ is a scaled, demeaned, and flexible approximation of $\frac{\mu_\xi}{\sigma_\xi} + \sigma_\xi G(\theta_{ik}, R_{it}, k)$, $\xi_{it} \sim N(0, 1)$, and

$$\theta_{it} = \frac{\frac{\sigma_\theta^2}{\sigma_\xi^2} \cdot (t-1)}{\frac{\sigma_\theta^2}{\sigma_\xi^2} (t-1) + 1} \cdot \left(\frac{1}{t-1} \sum_k^{t-1} (y_{ik} - l(R_{it}) - g(k)) \right).$$

Since scale and location parameters cannot be identified in a discrete choice environment similar The value function is approximated using a linear combination of orthogonal polynomials of the explanatory variables.⁷ I assume that

$$E \left[H(\theta_{it}, R_{it}, t) - \frac{1}{\sigma_\xi} G(\theta_i, R_{it}, t) - \frac{\mu_\xi}{\sigma_\xi} \right] = 0$$

This model incorporates the restrictions on the stochastic technology, but imposes no structure on the utility function. As a result, it does not impose a link between effort in the performance equation and disutility of effort in the attrition equation and the distribution of outside offers is normalized to be $N(0, 1)$. By estimating the semi-structural model, I recover the stochastic technology up to a constant and $G(\theta_{it}, R_{it}, t)$ up to a scaling parameter and an additive constant, $\hat{H}(\theta_{ik}, R_{it}, k)$. I estimate the model using maximum likelihood as discussed in Appendix C. These estimates are used to obtain the remaining structural parameters: the marginal disutility of one unit of effort γ , the curvature of the disutility of effort ψ , and the mean and variance of ξ , μ_ξ and σ_ξ^2 , as well as the discount factor δ .

The main advantage of this two-step estimator is its computational simplicity, but it has also one important limitation. The initial flexible estimator can be imprecise in small samples and this can generate a finite sample biases in the two-step estimator of structural parameters. One way to investigate the magnitude of the potential problem and limit its effect is to apply a K-step procedure as presented in Aguirregabiria and Mira (2007). This issues is left for future research.

⁷Bellman, Kaleba, and Kotkin (1963) first propose the use of such a linear approximation to the value function. The approximation method remains popular in both economics and machine learning, where it is still the workhorse for approximating dynamic programs as discussed in Kveton and Hauskrecht (2004).

Step 2 Let the difference in exerted effort under regimes 1 and 2 be Δl . Then from the performance equation,

$$\Delta l = \left(\frac{1}{\gamma}\right)^\psi \left(\beta_1^\psi - \beta_2^\psi\right) \quad (1)$$

The first step of the estimation provides the empirical counterpart of Δl , $\widehat{\Delta l}$. Thus, one can solve

$$\widehat{\Delta l} = \left(\frac{1}{\gamma}\right)^\psi \left(\beta_1^\psi - \beta_2^\psi\right)$$

for γ in terms of ψ and $\widehat{\Delta l}$; let the solution be $\gamma(\psi, \widehat{\Delta l})$.

To save on notation, define

$$\begin{aligned} & \lambda(H(\theta_{it}, R_{it}, t)) \\ &= E_\xi \max\{\xi_{it}, H(\theta_{it}, R_{it}, t)\} \\ &= H(\theta_{it}, R_{it}, t) \cdot \Phi(H(\theta_{it}, R_{it}, t)) + \varphi(H(\theta_{it}, R_{it}, t)) \end{aligned}$$

Note that $V(\theta_{it}, R_{it}, t)$ can be expressed in terms of $H(\theta_{it}, R_{it}, t)$ as follows

$$V(\theta_{it}, R_{it}, t) = E_\xi \max\{\xi_{it}, G(\theta_{it}, R_{it}, t)\} = \lambda(H(\theta_{it}, R_{it}, t))$$

Thus, (2) implies that

$$H(\theta_{it}, R_{it}, t) = U(\theta_{it}, R_{it}, t, \gamma, \psi) + \delta [E_{\theta_{it+1}|\theta_{it}}(\lambda(H(\theta_{it+1}, R_{it}, t+1)))] ,$$

where $E_{\theta_{it+1}|\theta_{it}}$ indicates that expectation is taken with respect to the distribution of the posterior beliefs in $t+1$ given the information available at t . After adjusting for the normalization in the first step, (2) gives rise to the following conditions for $t = 1, \dots, \tau_i$

$$E\left(\widehat{H}(\theta_{it}, R_{it}, t) - M_{it}(\theta_{it}, R_{it}, t, \Theta_2)\right) = 0,$$

where

$$M_{it}(\theta_{it}, R_{it}, t, \Theta_2) = \frac{1}{\sigma_\xi} \left\{ U(\theta_{it}, R_{it}, t, \gamma, \psi) + \delta \left[\mu_\xi + \sigma_\xi E_{\theta_{it+1}|\theta_{it}} \left(\lambda \left(\widehat{H}(\theta_{it+1}, R_{it}, t+1) \right) \right) \right] - \mu_\xi \right\},$$

and Θ_2 is the vector of structural parameters recovered at the second stage. Θ_2 does not contain γ as it can be recovered from $\gamma(\psi, \Delta \widehat{l})$.

Identification Theorem *Assume that the distribution of outside offers remain the same at all tenure horizons and does not depend on individual observed or unobserved characteristics. Then $\delta, \psi, \gamma, \mu_\xi, \sigma_\xi$ are identified.*

(a). The discount factor is identified from the variation in

$$E_{\theta_{it+1}|\theta_{it}} \left(\lambda \left(\widehat{H}(\theta_{it+1}, R_{it}, t+1) \right) \right)$$

originating from variation in beliefs, other things equal. By $t = 12$ the accumulation of experience has come to an end and the the difference in expected current utility $U(\theta_{it}, R_{it}, t, \gamma, \psi)$ from one period to the next is zero when the same pay regime prevails. Thus, any variation in $H(\theta_{it}, R_{it}, t)$ across periods originates from variation in beliefs, and the discount factor is identified from variation in the first-difference of $H(\theta_{it}, R_{it}, t)$ and the first-difference in $E_{\theta_{it+1}|\theta_{it}} \left(\lambda \left(\widehat{H}(\theta_{it}, R_{it}, t+1) \right) \right)$.

(b). The structural parameters γ and ψ are identified from condition (1) and from the first difference in $U(\theta_{it}, R_{it}, t, \gamma, \psi)$ when the pay regimes in the two periods differ. ■

As with most binary choice models, it is impossible to identify from the data the mean and the variance of the outside offer and the estimation proceeds under the assumption that $\mu_\xi = 0$ and $\sigma_\xi = 1$.

Define

$$M_t(\Theta_2) = \sum_i M_{it}(\theta_{it}, R_{it}, t, \Theta_2) \text{ and } \widehat{H}_t = \sum_i \widehat{H}(\theta_{it}, R_{it}, t)$$

where the summation is over the individuals who make the decision to stay or quit in period

t. Furthermore, let

$$M(\Theta_2) = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ M_t(\Theta_2) \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \quad \text{and} \quad \widehat{H} = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \widehat{H}_t \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

To find the remaining structural parameters Θ_2 , I use minimum distance estimation. The remaining parameters Θ_2 are the solution to

$$\min_{\Theta_2} \left(\widehat{H} - M(\Theta_2) \right)' \Omega^{-1} \left(\widehat{H} - M(\Theta_2) \right),$$

where Ω is the optimal weighing matrix. Let the covariance matrix of the structural parameters Θ_1 that enter in H and estimated in the first step be Σ . By the Delta method

$$\Omega = W' \Sigma W$$

where $W = \frac{\partial H}{\partial \Theta_1}$. Finally, the covariance matrix Σ is obtained from the first-step MLE estimates. Then, following Hansen (1982), the asymptotic covariance matrix for Θ_2 is $(J' \Omega^{-1} J)^{-1}$, where $J = \frac{\partial M}{\partial \Theta_2}$.

Note that the estimation of the structural parameters Θ_2 is in its essence a consistency check for the estimates of the first-step attrition model: the second step can be interpreted as a search for structural parameters that generate a data process that is consistent with the findings in the first step. The criterion function evaluated at the optimum has χ^2 distribution with 10 degrees of freedom under the null hypothesis that the theoretical model are valid.

4.1 Firm's Problem

For any regime, the probability of staying at tenure t , $p_{it} \left(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}\}_{k=1}^{t-1} \right)$, cannot be estimated analytically, so I resort to simulations to evaluate the profitability of pay regimes. For the set of employees who enter the firm, I draw paths of ε_t and ξ_t^* and θ to generate 1000 data sets. Using the point estimates from steps one and two, I generate the sequence of noisy performance signals and posterior beliefs. The generation of the separation indicators, s_{ik} , requires some care. Given a regime R , I solve for the value function of each individual for each posterior belief. I assume that conditional on staying for 2 years employees know the true value of their match quality and the accumulation of experience has stopped. Then, the worker's problem becomes

$$V(\theta_i, R) = \int \max[\xi_{it}^*, U(\theta_i, R) + \delta V(\theta_i, R)] dF_{\xi^*},$$

where $U(\theta_i, R)$ stands for the expected utility after the individual knows her match quality θ_i , there is no more experience to be gained, and $V(\theta_i, R)$ is the value of continued employment. This problem can be solved as a standard fixed-point problem using value function iteration. The starting value for the iterations is the discounted sum of expected utility, i.e.

$$V^0(\theta_i, R) = \frac{1}{1 - \delta} U(\theta_i, R)$$

Then, I solve backwards for the utility of continued employment $V(\mu_{it}, R, t)$, using the appropriate posterior. I use the Gauss-Hermite method with 8 nodes of integration. This approach to solving for the value function is similar to the one employed in Nagypal (2007). Comparing the drawn outside offers and the values of continued employment from above generates the sequence of separation indicators. It should be noted that all workers eventually quit. For each of the simulated data sets, I find the expected profits per entering employee by averaging the discounted sum of individual profits for the duration of stay. To find the expected profits per workstation, I take into account that all quits are replaced by new workers who have exactly the same expected profits at entry as the original cohort. This simulation method is used to

estimate the profits of the firm under the actually implemented regimes and to evaluate the candidates for the optimal linear contract at each step of the optimization. Given the low dimension of the optimization problem, I use a version of the simplex algorithm to find the optimal linear contract.

5 Results

This section presents the results from estimating the structural model and then shows how they can be used to find the profit-maximizing pay policy under various assumptions about the employment environment. The policy analysis indicates that turnover is a major channel through which pay incentives affect both performance and profits.

5.1 Estimates of Structural Parameters

In this subsection, I present the results from estimating the structural model and characterize the employment environment. The attrition model of the first step is estimated using MLE. The results, their econometric implications, and possible alternative specifications are discussed in greater detail in Bojilov (2011). The explanatory variables for the performance equations include second degree orthogonal polynomials of tenure and calendar time, dummies for regimes of operation and regimes of hiring, unobserved match quality, and controls. Specifically, regime 2 enters additively as implied by the theoretical model. Since regime 3 has the same pay schedule as regime 2 but conditions pay on the quality of service, the performance equation incorporates interaction terms between the tenure polynomials and regime 3. The attrition equations include orthogonal polynomials interacted with regimes and, depending on the specification, θ_i or μ_{it} , controls, calendar time, and regime of hiring. As a preliminary step, I conduct a specification search for the degrees of the orthogonal polynomials in the performance and attrition equations. I find that orthogonal polynomials of degree 2 for the performance equation and orthogonal polynomials of degree 3 for the attrition equations fit the data best. The estimates can be found in Tables 1-3. They are very similar to the ones

Table 1: Estimates for the performance equation in the attrition model

Dependent Variable: Performance	Attrition Model	
Explanatory Variable	Coef.	S.E.
t , orthog. pol. 1	-0.58	0.15
t , orthog. pol. 2	-0.61	0.09
regime 2	-0.21	0.07
regime 3	0.13	0.08
t , (regime 3), orthog. pol. 1	0.78	0.13
t , (regime 3), orthog. pol. 2	0.43	0.08
% outbound calls	-0.16	0.05
Constant	3.43	0.11
Log-likelihood	-3745.73	

Notes: The specification also includes calendar time orthogonal polynomials of degree 2 and individual controls: gender, age, marriage status, distance from home, and race.

Table 2: Estimates for the separation equation in the attrition model

Dependent Variable: Decision to stay	Attrition Model	
Explanatory Variable:	Coef.	S.E.
t , orthog. pol. 1	-0.51	0.17
t , orthog. pol. 2	-0.22	0.08
t , orthog. pol. 3	0.23	0.13
t , (regime 2), orthog. pol. 1	0.41	0.14
t , (regime 2), orthog. pol. 2	0.42	0.11
t , (regime 2), orthog. pol. 3	-0.78	1.59
t , (regime 3), orthog. pol. 1	0.40	0.16
t , (regime 3), orthog. pol. 2	0.23	0.08
t , (regime 3), orthog. pol. 3	-0.03	0.13
avg. % outbound calls in past	-0.31	0.09
Constant	0.58	0.10
Log-likelihood	-3745.73	

Notes: The specification also includes calendar time orthogonal polynomials of degree 2 and individual controls: gender, age, marriage status, distance from home, and race.

reported in Bojilov (2011) for the basic attrition model. The main difference is that some not significant variables have been omitted, along with the dummies for regime of hiring. In what follows, I make a brief summary of those results that are directly related to the second step of estimation and profits. Furthermore, I measure the contribution of effort, match quality, and experience to performance in terms of successful calls per hour (just calls per hour for short from now on).

Table 3: Estimates of parameters related to ability and learning in the attrition model

Parameter or explanatory variable	Attrition Model	
	Coef.	S.E.
σ_{ε}^2	0.17	0.01
σ_{ξ}^2	1	
$\rho(\varepsilon, \xi)$	0.01	0.04
σ_{θ}^2	0.48	0.02
μ_{it}	0.18	0.05
$t.\mu_{it}$, orthog. pol. 1	0.06	0.03
$t.\mu_{it}$, orthog. pol. 2	0.02	0.04
$t.\mu_{it}$, orthog. pol. 3	0.03	0.03
Log-likelihood	-3745.73	

Notes: The specification also includes calendar time orthogonal polynomials of degree 2 and individual controls: gender, age, marriage status, distance from home, and race.

In the first step, I estimate the distribution of match quality at entry and characterize the associated dynamics of learning and turnover. The variance of match quality is 0.48 and accounts for the greater part of the variance in performance at entry under regimes 1 and 2. Moreover, it has an important effect on attrition. Figure 1 presents, among other things, the distribution of match quality at entry and how it changes by the sixth months of employment. The value of θ is on the horizontal axis, while the vertical axis represents the proportion of agents of a certain match quality who are present in the firm at a given tenure horizon. The figure shows that the conditional distribution of θ shifts to the right as tenure increases under both regimes 1 and 2 and only workers with very high match quality remain employed after six months of work. As expected, the switch from regime 1 to regime 2 generates an increase in turnover at any tenure horizon. The variance of the disturbance term in the performance equation is estimated at 0.17 which implies that the signal-to-noise ratio, defined as the ratio of the variance of match quality over the variance of noise, is approximately 2.6. Consequently, within 6 months the variance of the posterior beliefs declines to approximately 0.05 and the weight on the initial belief declines to almost zero.

Furthermore, I recover the technology up to an additive constant. The estimated parameters for the performance equation are broadly consistent with the theoretical predictions. In economic terms, the switch from regime 1 to regime 2 leads to a decline in worker's effort and in turn performance by about 0.2 calls per hour which translates into a decline in hourly

pay by approximately \$2. The estimates imply a significant improvement in performance over time due to the accumulation of experience: in the first 6 months of employment performance increases by approximately one successful call per hour, or 35% growth in the first six months under regime 1. Under regime 1, this growth translates in an increase in hourly pay by approximately \$3.3. Finally, I also estimate flexibly the normalized and scaled value of continued employment $G(\mu_{it}, R, t, X_{it})$ which provides the basis for the second step estimation.

Table 4 presents the estimates from the second step. Before discussing the results from the second step, I first check the validity of the overidentifying restrictions. The χ -square test with 10 degrees of freedom for the overidentifying restrictions fails to reject the null hypothesis that the restrictions are valid, since the test statistic is 6.14. Thus, I conclude that the data is consistent with the restrictions imposed by the model.⁸ The second step results allow for the characterization of optimal effort choice under regimes 1 and 2. The elasticity of effort to pay incentives ψ is estimated at 3.24 with a standard error of 0.2, implying that workers' supply of effort is highly sensitive to changes in pay incentives. The relative benefit of effort to its subjective cost represented by γ is estimated at 3.92 with a standard error of 0.23, so that effort amounting to one call per hour costs to the individual \$3.9. Given the estimates of γ and ψ , the level of effort under regime 1 translates into an increase in performance by approximately 0.59 calls per hour and under regime 2 by 0.38 calls per hour. Compared to the variation in match quality, the contribution of effort to performance is relatively small: less than one standard deviation under regime 1 and even less under regime 2. In contrast, the mean of match quality in the population of entering workers is approximately 2, i.e. independent of pay incentives an employee of average quality successfully completes two calls per hour when starting work. Furthermore, the results indicate that in the absence of selection at entry and exit the optimal piece rate involves $\beta = 6.55$ ⁹. The fact that the implemented

⁸**Note that the estimation of the structural parameters Θ_2 is a test for consistency of the first-step estimates with the specified utility in the model.**

⁹Solving the profit-maximization problem of the firm subject to a participation constraint yield

$$\beta = \frac{p\psi}{\psi + 1},$$

where p is the revenue generated from the employment relation and in the present context is \$8.5 per call. Substituting the estimated ψ gives $\beta = 6.55$

Table 4: Estimates of the structural parameters of the model.

Parameter:	Coefficient.	Std. Err.
ψ	3.24	0.20
γ	3.92	0.23
δ	0.76	0.10
Δl	-0.21	0.06
Δ disutility experience by $t = 6$	-0.65	0.07
σ_θ^2	1.02	0.09
μ_θ	0.48	0.02
χ_{12}^2 test stat.	2.02	0.11
	6.14	

pay regimes have β much lower suggests that turnover has a nontrivial effect on profits.

The monthly discount factor is estimated at 0.76 with standard error of 0.1, indicating a strong preference for present to future consumption: when making decisions workers assign a weight 0.001 to consumption after two years.

5.2 Profits and Policy Analysis

The estimates of the structural model provide the basis for counterfactual policy analysis. In this subsection, I start by discussing the profitability of the implemented regimes 1 and 2, as well as the contribution of effort, experience, and match quality to profits under these regimes. Then, I consider the problem of finding the optimal linear contract. I compare the results to those previously presented for regimes 1 and 2. Finally, I consider some counterfactual changes in the firm environment and their effect on profits. In particular, I consider a higher level of turnover costs than the one reported by the firm. I also evaluate the implications for profits and the optimal compensation policy when workers learn about their match quality before deciding whether to enter the firm.

The decomposition of profits is difficult because profits depend on both performance and the probability of staying, while the latter is a highly nonlinear function of effort, beliefs about match quality and experience. The additive structure of the technology allows for isolating

the contribution of effort, ability, and experience to profits. I define

$$\pi_{\theta}(R) = E_{\theta} \left\{ \sum_{t=1}^{\infty} \tilde{\delta}^{t-1} P_{it} \left(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}^*\}_{k=1}^{t-1} \right) (r - \beta) \theta_i \right\}$$

where

$$P_{it} \left(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}\}_{k=1}^{t-1} \right) = \frac{p_{it} \left(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}\}_{k=1}^{t-1} \right)}{1 - \sum_{t=1}^{\infty} \delta^{t-1} \left(1 - p_{it} \left(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}\}_{k=1}^{t-1} \right) \right)}$$

to be the profits associated with match quality. In a similar way,

$$\pi_l(R) = E_{\theta} \left\{ \sum_{t=1}^{\infty} \tilde{\delta}^{t-1} P_{it} \left(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}\}_{k=1}^{t-1} \right) (r - \beta) l(R) \right\}$$

$$\pi_t(R) = E_{\theta} \left\{ \sum_{t=1}^{\infty} \tilde{\delta}^{t-1} P_{it} \left(R, \theta_i, t, \{\varepsilon_{ik}, \xi_{ik}\}_{k=1}^{t-1} \right) (r - \beta) g(t) \right\}$$

$\pi_l(R)$ and $\pi_t(R)$ stand for the contribution to profits by effort and with experience, respectively. I take hourly wage as a benchmark regime with respect to which I evaluate how total profits and the contributions of effort, experience and match quality defined above change as the pay regime changes. An alternative approach that I also apply to the study of the effect of effort on profits is to compare profits under the same pay regime when effort affects performance and stay and when it is restricted to have no effect on them: the difference in the profits provides a conservative estimate for the contribution of effort to profits.

5.2.1 Implemented Regimes

Table 5 presents the linear pay policies that I analyze, along with profits,¹⁰ effort, average match quality and average tenure per workstation under each of them. Table 6 considers the channels through which pay incentives affect profits. It considers the effects on the contributions of effort, match quality, and tenure to profits when switching from the initial hourly

¹⁰Recall that from the definition of profits, total profits stands for the discounted infinite sum of hourly profits starting from the month of hiring the worker.

wage to some alternative pay regimes. Under an hourly wage, employees do not exert effort. Moreover, equal hourly pay implies that workers of different match quality are equally likely to quit at any tenure horizon. I fix the hourly wage to \$9.5 which was actually implemented by the firm prior to January 2005. This hourly wage is clearly quite low relative to the mean of the outside offer and leads to a very high turnover: more than 93% of the employees last at most six months in the firm. This high turnover leads to a low level of experience in the workforce as indicated by an average tenure of 3.23. Furthermore, the failure of the hourly wage to distinguish between workers of high and low match quality leads to an average match quality of 1.99 calls per hour. Taken together, these effects of the hourly wage lead to total profits of \$19.4.¹¹

Switching from the hourly wage to regime 1 induces all workers to exert effort of 0.59 calls per hour but also rewards workers of high match quality more than workers of low match quality. As discussed in Bojilov (2011), the result is that workers of high match quality stay longer in the firm than workers of low match quality. These differences lead to an increase in average match quality to 2.88 calls per hour. The net effect of the change in the compensation policy on the separation decisions is a decline in turnover illustrated with an increase in average tenure to 11.3 months. The retention of employees of high match quality, along with the decline in their probability of quitting at any tenure horizon leads to an impressive increase in π_θ by \$90. The lower turnover also leads to an increase in the profits associated with experience by approximately \$49. Finally, the introduction of the bonus rate of \$3.3 per successful call induces effort that generates profits associated with effort in the amount of \$31.42.¹² Total profits jump to \$167. These numbers indicate that the increase in π_θ , followed by the increase in π_t , rather than the increase in π_l makes the greatest contribution to the increase in profits when switching from hourly wage to regime 1. The results suggest that the firm benefits considerably from the accumulation of workers of high match quality

¹¹Under the given specification of the stochastic technology and the utility, the profits associated with match quality π_θ are \$28.89 and the profits associated with experience π_t are \$10.4.

¹²Recall that the firm incurs a flat hourly pay of α and turnover costs which must be subtracted to obtain the total profits.

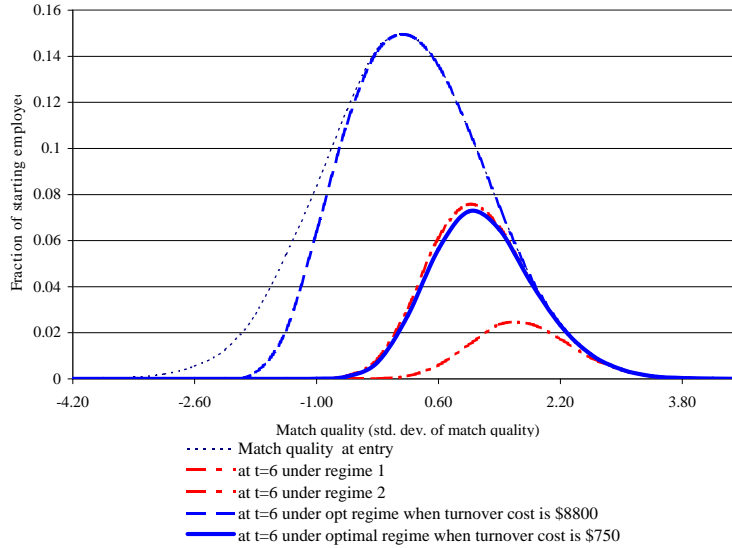
through turnover.

Next, I consider the effect of regime 2 on profits. Recall that this regime stipulates both lower base pay and lower piece rate. This less generous compensation policy leads to a sharp increase in the probability of quitting during the first six months which approaches the levels under the hourly wage. The result is average tenure of 6.2 months, a decrease by more than 40% relative to regime 1, which implies also lower levels of accumulated experience. While the probability of quitting increases at each tenure horizon, the firm still retains workers of very high match quality, as discussed in paper 3. However, average match quality under regime 2 is not higher but slightly lower than average match quality under regime 1: 2.83 calls per hour. This result indicates that the negative effect of high turnover more than offsets the effect of retaining only the workers of highest match quality. At the same time, effort declines to 0.34 calls per hour. The combined effect of these factors implies that regime 2 yields much lower profits than regime 1. Despite the fact that the piece rate declines by \$0.8 calls per hour, the profits associated with match quality are still \$70 higher compared to their level under the hourly wage. However, as a result of the high rate of destruction of accumulated experience, π_t is quite close to its level under the hourly wage: it is only \$16 higher. The profits associated with effort π_l are approximately \$10, and total profits amount to about \$110. Thus, under regime 2 match quality continues to be a crucial determinant of profits and the decline relative to regime 1 is smallest in the case of π_θ .

5.2.2 Optimal Regime

The solution of the profit-maximization problem is the optimal pay regime R^w defined by $\alpha_w = 3.65$ and $\beta_w = 3.24$. Several factors affect its properties. While steep incentives induce more effort and increase the probability of staying, they also surrender a larger proportion of the revenues to the employees. Since quitting of an employee comes with the possibility of hiring a better one in the future, the firm chooses a pay schedule that among other things, balances the benefit from continued employment of a worker and the benefit from finding one of higher quality. The results here depend crucially on the firm-specific nature of the

Figure 1: Ability under regime 1, 2, and the optimal regimes when the turnover cost is \$750 and when it is the industry average of \$8800.

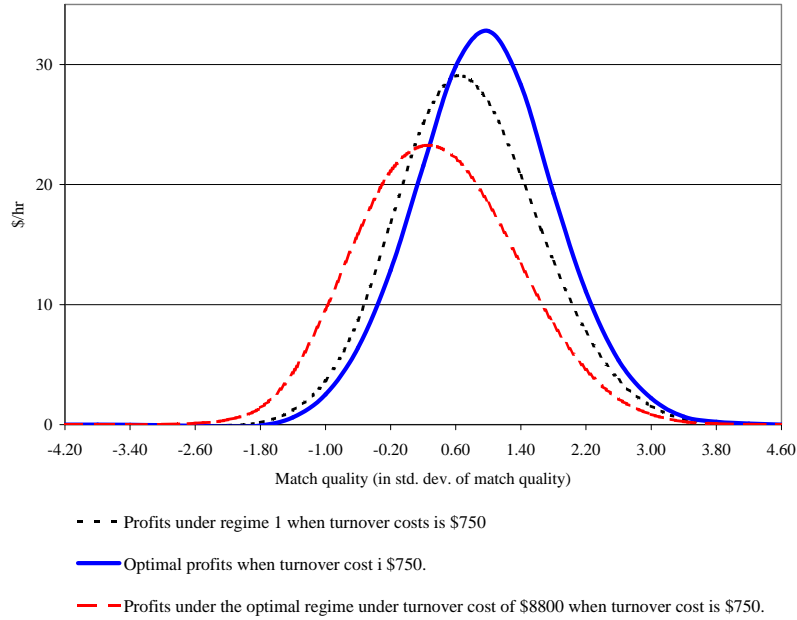


match quality parameter: in particular, the ability of the firm to extract much of the surplus from the employment relation will be limited if workers can export their match quality θ to alternative jobs. The findings also depend to some extent on the simple nature of the compensation policy: for example, the properties of the optimal pay regime will change if the firm can condition base pay α on posterior beliefs about match quality.

Table 5: Profits, effort, ability and tenure per workstation under different regimes.

Pay policy:	α	β	l	$E(t)$	$E(\theta)$	π
hourly wage, \bar{w}	9.5	0	0	3.22	2.03	19.45
regime 1	3.8	3.3	0.59	11.3	2.88	167.81
regime 2	3.5	2.8	0.38	6.23	2.83	110.17
regime R^w	3.65	3.24	0.55	9.85	2.83	174.24
regime R^w , known ability	3.65	3.24	0.55	12.30	3.12	215.92
regime R^n , known ability	3.74	3.09	0.47	11.61	3.17	221.74
regime R^h , high costs	1.82	5.44	2.91	21.56	2.17	162.30

Figure 2: Comparison between profits under regime 1, the optimal regime when turnover cost is \$750 and when it is \$8800. Turnover cost is \$750.



The slope and base pay of the optimal pay regime R^w are very close to those implemented under regime 1. The optimal pay regime R^w induces considerable turnover: only about 55% of the employees stay more than six months in the firm. Furthermore, it not only induces a similar rate of turnover but also leads to a similar quality mix at different tenure horizons as regime 1. Figure 1 shows that the conditional distributions of match quality after six months under regime 1 and regime R^w are almost identical. It also indicates that only workers of high match quality (match quality greater than $\theta + \sigma_\theta$) experience little or no turnover. The

Table 6: Effects of different pay regimes relative to hourly wage.

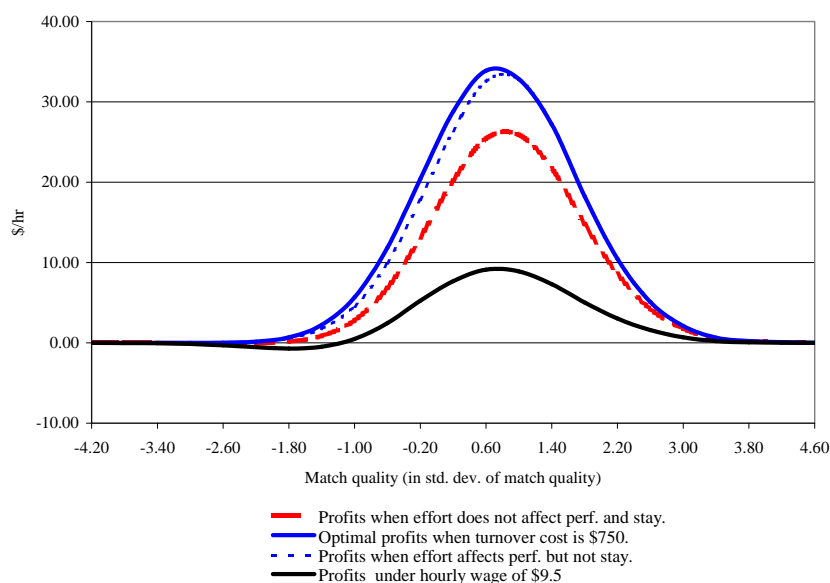
Pay policy:	α	β	$\Delta\pi_l$	$\Delta\pi_t$	$\Delta\pi_\theta$
regime 1	3.8	3.3	35.06	49.14	89.85
regime 2	3.5	2.8	9.74	15.68	70.33
regime R^w	3.65	3.24	31.42	44.32	99.32
regime R^w , known ability	3.65	3.24	33.31	49.93	131.18
regime R^n , known ability	3.74	3.09	29.18	48.37	140.57
regime R^h , high costs	1.82	5.44	191.04	56.20	23.42

Note: $\Delta\pi_x = \pi_x(R) - \pi_x(\bar{w})$, $x = l, t, \theta$.

small slope of incentives induces effort that translates into only 0.55 calls per hour, less than one standard deviation of match quality among starting employees. The distributions of expected profits under the optimal pay regime R^w and regime 1 are again very similar but some differences are also present, as evident from Figure 2. Regime 1 generates more income on employees of average quality while regime R^w generates more profits on the top performers. This pattern is explained by the fact that under regime R^w the firm captures more of the surplus from the top performers while the slightly higher slope of incentives and base pay under regime 1 induce more workers of average ability to stay and work. The net effect is that optimal pay regime R^w generates approximately 4.8% higher profits than regime 1.

Relative to regime 1, the optimal regime is less generous which leads to a small increase in the probability of quitting across posterior beliefs and tenure horizons. The result is average tenure of about 10 months, slightly lower than the 11.3 months under regime 1. This finding implies that the level of accumulated experience under the optimal regime R^w is just below that for regime 1. The probability of quitting increases at each tenure horizon relative to regime 1, but the firm still retains workers of very high match quality, as discussed in the previous paragraph in the context of Figure 2. The net effect is that the average match quality under R^w is 2.91 calls per hour, slightly higher than its counterpart for regime 1 which leads also to higher profits. Furthermore, shaving off 6 cents from the bonus rate reduces effort only to 0.55 calls per hour relative to the 0.58 call per hour under regime 1. Due to the improved quality mix, the small negative effect on turnover, and the small reduction in the variable costs, the contribution of match quality to profits, π_θ , increases by \$99 relative to its level under hourly wage. The contribution of tenure to profits, π_t , increases by \$44 when switching from hourly wage to regime R^w . The contribution of effort, π_l , declines slightly relative to regime 1 to \$31. Combining all of these with the costs of turnover and base pay yields total profits of \$174. The results show that switching from the benchmark hourly wage to the optimal regime R^w leads to an increase in profits by more than eight times, or in absolute terms by \$154. The results show that much of this change is due to the effect of incentives on the quality mix of the workforce rather than the effect of incentives on effort.

Figure 3: Gains from switching to the optimal regime from hourly wage. Turnover cost is \$750.



An alternative approach to evaluate the effect of incentives on profits through effort is to compare profits under the optimal pay regime when effort affects performance and stay and when it is restricted to have no effect on them. My analysis starts with the quality mix, assuming that effort is not a channel through which incentives affect performance or separation decisions. Figure 3 shows that under the hourly wage of \$9.5, implemented until January 2005, the firm makes losses on some workers of below average ability and its expected profits amount only to approximately \$20, due to a high quitting rate and the associated costs. The figure also indicates that the introduction of the optimal regime induces high quality employees to stay while low quality employees to quit. The firm captures 75 % of the additional surplus and profits increase by more than a factor of three.

Next, I relax the restriction that incentives do not affect effort, but still maintain that effort choice has not effect on separation decisions. Figure 3 shows that the exerted effort leads to an additional increase in profits by 114%. Finally, I also allow effort choice to affect separation decisions, but Figure 3 indicates that only but a few separation decisions remain

unchanged: the combined effect of effort choice and match quality for those who switch from quitting to staying accounts for a 19% increase in profits. Thus, the total effect of switching from hourly wage to the benchmark rate results in a dramatic increase in profits, but two-thirds of the increase would have materialized even if pay incentives did not affect effort choice or the separation decisions.

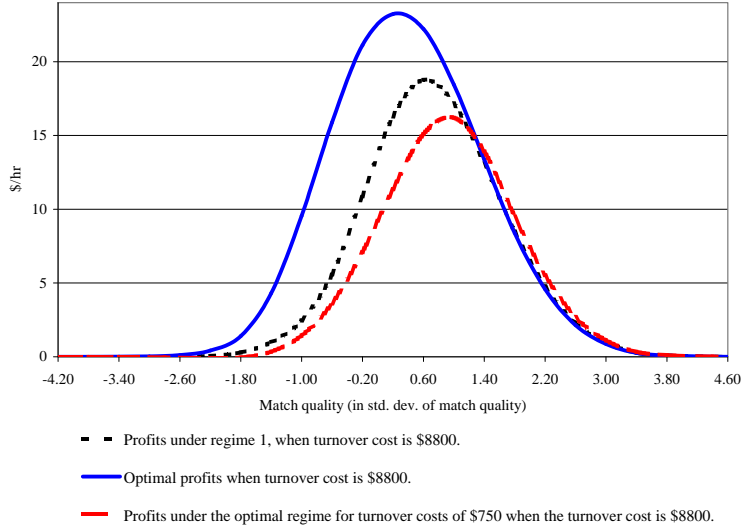
To summarize, these results show that most of the increase in profits from switching to the optimal pay regime can be traced back to the effect of incentives on the quality mix. Pay incentives not only induce high quality employees to stay but also act as a selection mechanism that helps the firm build a workforce of high match quality over time. In the present context, the firm exploits the firm-specific nature of the relation to capture most of the surplus generated by the employment relation.

5.2.3 Counterfactual Experiments

The turnover costs of \$750 reported by the firm appear very low relative to industry averages published in Superb Staff Services (2011) which vary between \$4,100 and \$25,000. In this subsection, I explore the effect of high turnover costs on profits under the regime R^w , optimal under a turnover cost of \$750, and search for the optimal regime under turnover costs equal to the industry average of \$8,800. Furthermore, I study the effect on profits when the worker knows her match quality before deciding to start working but the employer does not.

Turnover Costs The optimal pay regime R^h when turnover costs are \$8,800 is defined by $\alpha_h = 1.82$ and $\beta_h = 5.44$. These slope and base pay are much different from the one's implemented by the firm. The high-powered incentives induce little turnover, mainly in the first two months of employment, and a high level of effort resulting in 2.91 calls per hour. Figure 4 presents profits under regimes 1, R^w and R^h when turnover costs are \$8,800. A comparison of profits under pay regime R^h and regime R^w reveals that the two have a very similar expected profits from the top performers, while regime R^h accumulates much higher profits on the employees of low and average quality. Thus, the top performers capture much

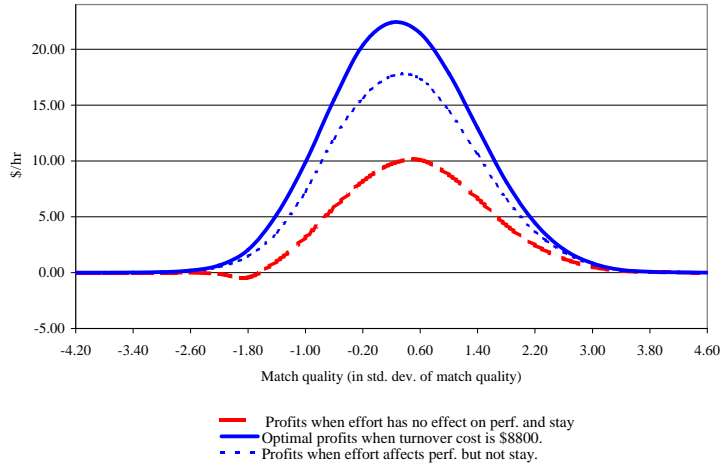
Figure 4: Comparison between profits under regime 1, the optimal regime when turnover cost is \$750 and the optimal regime when it is \$8800. Turnover cost of \$8800.



of the revenue under regime R^h , while the firm increases its profits from the higher effort exerted by employees who would not have stayed under regimes 1 or R^w . Table 5 shows that the low turnover rate leads not only to a high average tenure of about 20 months but also to a low average match quality of 2.17 calls per hour. In contrast to the case of the optimal contract when turnover cost is only \$750, the profits under optimal regime R^h come mainly from high levels of effort: π_l for this regime is \$191. Still, total profits are only \$162 because of the high turnover costs.

Next, I analyze the composition of profits. I start with the quality mix, assuming that implementing regime R^h does not affect performance or separation decisions through effort choice. Figure 5 shows that the introduction of regime R^h , even in the absence of any effect through effort choice, allows many employees to remain in the firm and in turn generate revenue of which more than 67% go to the workers. Then, I relax the restriction that incentives do not affect effort but still maintain that effort is not a channel through which incentives affect separation decisions. Figure 5 shows that, under these new restrictions, pay incentives induce

Figure 5: Gains from switching to the optimal regime from hourly wage. Turnover cost is \$8800.

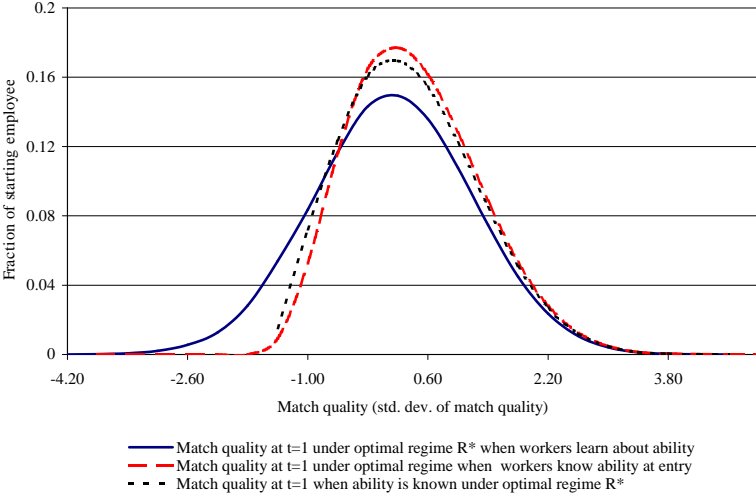


effort that increases revenues considerably in contrast to the case of the optimal regime when turnover costs are \$750. Finally, I allow effort choice to affect separation decisions. Given the assumptions of the model about the utility function, the regime induces higher effort and higher utility. Thus, the introduction of effort choice changes some but not all separation decisions: some workers who would have otherwise left now decide to stay. The employees who now stay contribute to profits with their match quality and effort: as evident from Figure 5, the effect is not negligible.

Summing up, the total effect of switching from hourly wage to regime R^e results in an impressive increase in profits, but only 27% of this growth would have materialized if pay incentives did not affect performance and separation decisions through effort choice. Thus, this counterfactual experiment indicates the sensitivity of the solution to the profit maximization problem to turnover costs.

Workers Know the Match Quality Table 5 and 6 also report average match quality, average tenure, profits, and their decomposition when workers know their match quality before

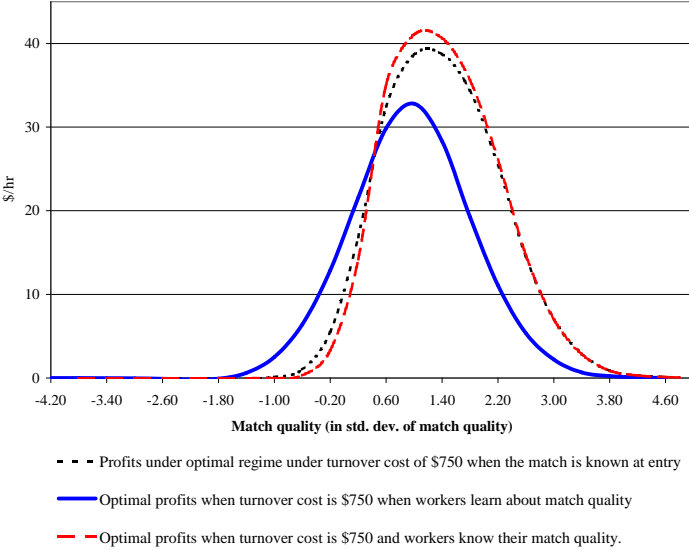
Figure 6: Ability at entry under the optimal regimes when workers know match quality, when they learn about it, and when the latter is applied to an environment in which workers know their match quality. Turnover cost is \$750.



deciding to join the firm. When regime R^w , optimal when workers learn their match quality, is implemented in this environment, profits increase to \$216. Much of this increase can be traced back to self-selection at entry: some workers know that their match with the firm is of low quality and decide to opt out for an alternative. Figure 6 shows the mean of the distribution of match quality at $t = 1$ under R^w when workers know their match quality is 2.35 calls per hour compared to 2.01 calls per hour when they learn about it. As a result, the firm accumulates workers of high match quality faster than when workers learn about match quality and the average match quality increases to 3.12 calls per hour, while the average tenure increases to 12.3 months. These effects lead to a considerable increase of \$131 in π_θ relative to its level under hourly wages, which is largely responsible for the increase of total profits to \$216.

The next step is to find the optimal pay regime when workers know their match quality before deciding to enter the firm. This problem is a special case of the more general model presented above: the prior belief is a degenerate distribution centered at the true value of

Figure 7: Comparison between profits under the optimal regimes when workers know match quality at entry, when they learn about it, and when the latter is applied to an environment in which workers know it. Turnover cost of \$8800.



match quality. The solution of the profit-maximization problem is the optimal pay regime R^n defined by $\alpha_n = 3.74$ and $\beta_n = 3.09$. The results for this regime are reported in tables 5 and 6. The firm offers lower incentives to exert effort in order to capture a greater share of the profits associated with match quality which is partially offset by a modest increase in the base pay. Thus, the growth in income and the variance of the distribution of income in this environment are smaller than their counterparts when workers learn about match quality. Still, one cannot generalize too much from this result because the properties of regime R^n depend considerably on the restriction to search for the optimal regime within the family of linear contracts only. The average match quality under R^n is 3.17 calls per hour and the distribution of match quality among the entering employees is not much different from that under R^w , as shown on Figure 6. Average tenure is 11.6 months, compared to 12.3 months under R^w , while effort amounts to only 0.47 calls per hour. Total profits are \$221. Figure 7 presents profits under regime R^n when workers know their match quality, under regime R^w when workers learn about their match quality, and when they know it. It shows that the

profits under R^n and R^w when workers know their match quality are similar. This is driven by the fact that more or less the same type of people enter the firm under both regimes and all differences arise from the fact that R^n shaves off more of the revenue from top performers by decreasing the bonus rate at the expense of a slightly higher turnover. Consequently, the results indicate that employers can benefit considerably if they can introduce a technology that helps workers find out their match quality before they decide to enter the firm.

5.2.4 More General Contracts

In the final set of counterfactual experiments, I relax the restriction that the firm considers only linear contracts in current output and allow for more general contracts that depend on both current and past performance. In such an environment Holmstrom (1999), among others, points out that the initial symmetry of learning is broken and the worker could try to manipulate the beliefs of the employer to influence her compensation. While this issue is a very important, it is outside the scope of this paper and is, therefore, left for future research. Thus, I proceed with the analysis under the additional assumption that effort is observable, but not verifiable. Such an assumption has some justification because call center agents work in teams of 10 to 15 in the almost constant presence of their supervision.

Given the additive structure of the technology and that both the performance signal and the unobserved match quality are normally distributed, the tenure and the average of the demeaned past performance are sufficient statistics that summarize posterior beliefs. Based on these observations, I focus on contracts that depend only on these two sufficient statistics and on current output. Furthermore, since both the firm and the individual are risk-neutral, the economic intuition suggests that the optimal contract will have the interpretation of sale of the firm's claims to future output to the worker. For this reason, finding the most profitable pay schedule can be interpreted as a problem of evaluating the stream of surplus generated by an employee. Based on these considerations, I study contracts that consist of the sum of two terms: the first one depends on current output and provides the maximal incentives to the worker to exert effort, so that conditional on working and beliefs the first best is achieved;

Table 7: Effects of different pay regimes relative to hourly wage.

Pay policy:	$E(t)$	$E(\theta)$	π
regime, time invariant	8.53	3.23	256.84
regime, varying with t	9.12	3.46	272.65
regime, optimal linear	9.85	2.93	174.24

and the second term conditions the 'sale' price of output on posterior beliefs. This second term captures the tradeoff that a firm faces between extracting the surplus from an existing relation, the benefits from exploring the possibility that a new employee could turn out to be better than the existing, and the threat that even if the firm wants the employee to stay an attractive outside offer may entice her to quit. In technical terms, it depends on the option value of experimentation for the firm, the option value of staying for the employee, and the distribution of outside offers.

The analysis proceeds in two steps. I start with contracts that depend on current performance and the average of the past signal (the posterior mean), but do not vary with tenure. Then, I allow for contracts whose schedule also depends on tenure, i.e. on the precision of the posterior mean. In this way, the results illustrate the effect of improvement in the precision of posterior beliefs on the firm compensation policy. Throughout, I maintain that the most that the firm can, at most, take all of the surplus from the employment relation, i.e. it cannot impose arbitrarily large punishments to ensure that the worker quits. This constraint is imposed for both technical reasons and for the benefit of greater realism in the set-up.

The compensation schedule for the optimal tenure-invariant contract is represented on Figure 8. It is increasing and hits the lower bound of zero total pay at about match quality of 1.4 (-1 standard deviation from the mean). Furthermore, for match quality above 3.14 (1.6 standard deviations), the compensation schedule grows very slowly and eventually becomes virtually flat. The schedule is largely concave and pay starts declining very rapidly as match quality declines from 1.75 (-0.5 standard deviations). Finally, the results indicate that the firm manages to capture the greater part of the employment surplus.

Table 7 shows that the optimal history dependent but tenure-invariant contract generates

Figure 8: Compensation schedule of the optimal history-dependent tenure-invariant contract as a function of deviations from μ_{it} .

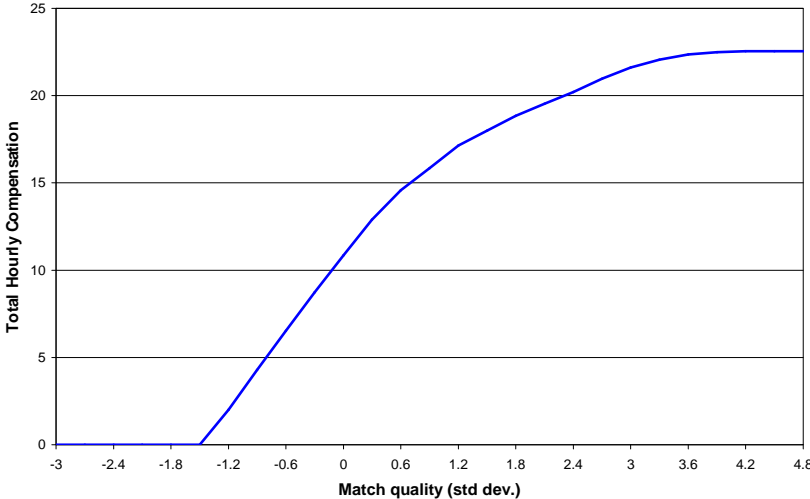
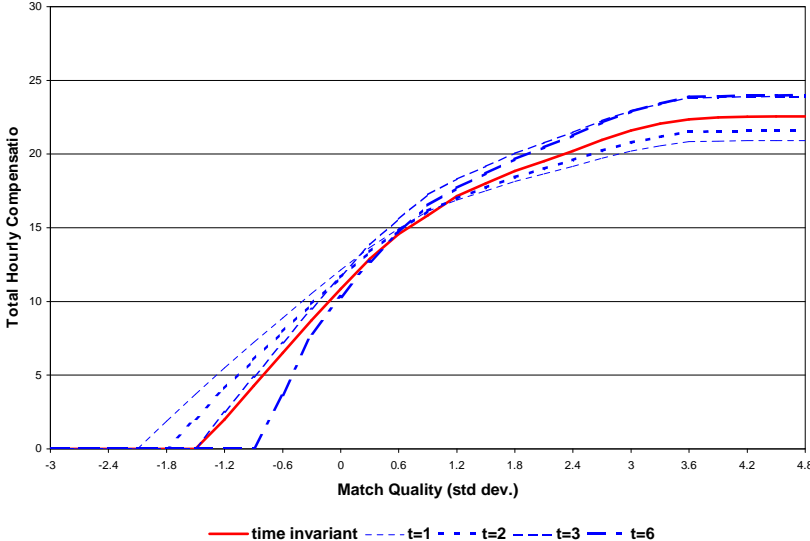


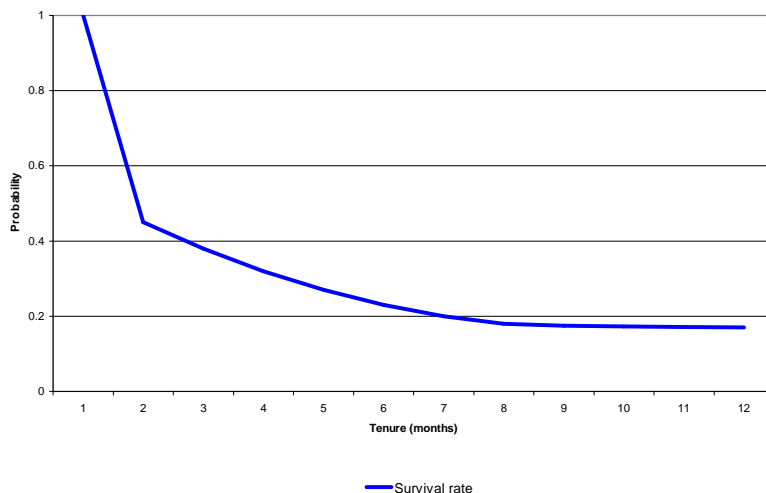
Figure 9: Compensation schedule of the optimal history- and tenure-dependent contract



profits of \$256.84, which is approximately 48% higher than the optimal linear contract. This difference is mainly due to the fact that history dependence and non-linearity allow the firm to provide incentives to achieve the first best, which gains are mostly captured by the firm through the conditioning on the average of past signals. In addition, to this source, there is also an increase in the average match quality in the firm to 3.2 from 2.9 under the optimal linear contract. Interestingly, average tenure in the firm decreases slightly relative to the average tenure under the optimal linear contract. This finding reveals that the firm does not use the greater flexibility of the nonlinear contracts to ensure that it retains for sure very high quality workers. Actually, it makes profits from extracting almost all of the surplus from low and average quality matches. In addition to the high share of the surplus, such a compensation schedule ensures that low and average quality workers quit fast and free their positions for potentially highly productive workers. Therefore, what attracts candidates to the firm is the prospect that, if they are some of the lucky few who match well, they will capture a sizable portion from the employment surplus. Nevertheless, the attractiveness of the job relative to the alternatives makes workers quit eventually even if they are at the top of the distribution.

Next, I consider the optimal contract in the family of history- and tenure-dependent contracts. Due to the speed of learning and the computational intensiveness of searching for the optimal sequence of pay schedules, I limit the time dependence only to the first 6 months. After that, I maintain that the compensation schedule remains unchanged. The sequence of compensation schedules by tenure are reported in Figure 9. As a benchmark, in red, I also plot the time-invariant contract discussed in the preceding paragraphs. The results indicate that the optimal tenure-dependent schedule for the first month is translated to the left and has generally flatter slope than the tenure-invariant schedule. These features capture the fact that after the first performance signal both the firm and the worker have very noisy posterior beliefs, so both the option value of staying is high and the value of experimentation for the firm is low. As tenure increases, beliefs become more precise and both the worker and the

Figure 10: Survival rate under the optimal history- and tenure-dependent contract.



firm are more inclined to terminate the relation.¹³ This dynamics is captured by the shift to the right of the pay schedules for the following months and by their increasingly steeper slope. The qualitative features of each pay schedule are, however, very similar to those of the previously discussed optimal tenure-invariant contract.

Finally, table 7 shows that the greater flexibility in conditioning the pay schedule on tenure has not led to radical changes compared to the tenure-invariant but history dependent contract: profits increase by only about 6 per cents or \$16. The results reveal that this difference in profits is largely due to an improvement in the mechanism of the firm to select workers of high quality through high turnover of low and medium quality matches. Evidence for this last point are presented in Figures 10, which plots the probability of survival of a worker as tenure increases. Not surprisingly, the survival rate is monotonically declining. It also drops sharply in the first 6 months, so that practically by the 12th month since entry more than 80 per cent of the employees have already quit. The rate continues to decline, but

¹³Note that this increasing hazard of quitting conditional on tenure is a feature of the underlying normal distributon. For more details, see Ericson and Pakes (1999).

at a much lower rate. Thus, the annual turnover at the firm under the optimal history- and tenure-dependent contract is similar in magnitude to the turnover under the very inefficient hourly wage. Yet, the causes, the characteristics, and the implications of turnover under the two contracts are very different.

6 Conclusion

To summarize, this paper presented a structural model of effort choice, learning about match quality, and turnover. It showed how such a model can be estimated with a two-step procedure that borrows ideas from the literature on estimation of dynamic structural models. Thus, the paper presented a simple and easy to implement estimation method for armed-bandit problems. The results indicated that employees are very responsive to pay incentives, and impatient to postpone future consumption. Furthermore, workers accumulated experience during the first six months on the job which improved performance. Still, variability in the quality of the employer-employee match accounted for most variation in performance across individuals under a given set of pay incentives. The paper examined a variety of contracts to find that the firm maximizes profits by selecting and keeping the high match quality employees, even sometimes at the expense of inducing low effort. It also showed that most gains from switching to the optimal pay regime from an hourly wage could be traced to an improvement in the match quality of the workforce. It also characterized some of the complicated ways in which effort choice and the quality mix interact to shape the compensation policy of firms. In particular, the results indicated that turnover, rather than effort, is the main channel through which pay incentives affect profits and welfare. Relatedly, as it presented evidence for both efficient and inefficient turnover, the paper cautioned that the correct interpretation of turnover always depends on the incentives environment that generates it. Finally, the analysis of history- and tenure-dependent compensation schedules confirmed the qualitative results obtained for the optimal linear contract. In this last part, the analysis proceeded under the additional assumption that effort is observable but not verifiable, which assumes away the

possibility for strategic manipulations of the productivity signal by the workers. While such an assumption was plausible in the current context, I believe that the incorporation of career concerns in the presented framework is a promising avenue for future research.

7 Appendix A

The crucial feature of additive separability of the technology in effort and match quality can be tested nonparametrically. as follows. ¹⁴ If workers start with a common prior and learn the quality of their match with the employer over time, the distribution of match quality does not vary across different pay regimes. Since match quality does not interact with effort, only the mean of performance in the first period varies across regimes. That is, the distributions of performance across regimes are the same up to a location parameter. Observation 1 states this argument formally.

Observation 1. *Consider the model defined by (1) and (2) and suppose that the workers share a common prior at the time of hiring. Then the demeaned distribution of performance at $t = 1$ is the same across piece rates:*

$$F(y_1^0|R) = F(y_1^0|R'),$$

for any R and R' , where $y_1^0 = y_1 - E(y_1|R)$.

Proof: Since effort enters additively in the stochastic technology and optimal effort does not vary with θ_i or θ_{it} across i , the pay regime affects only the first moment of the conditional distribution of performance. Furthermore, under the assumption of a common prior belief at the time of hiring, the entry decision is not affected by the pay regime in place, so $F(\theta_i|R) = F(\theta)$. Therefore, $F(y_1^0|R) = F(y_1^0|R')$. ■

The proof depends crucially on the assumption of common prior: if some workers had a more accurate belief about the quality of the match than others, the probability of staying more than one period will differ with beliefs leading to differences in the distribution of newly hired employees. Thus, for single-peaked distributions with non-zero probability for every possible match quality this property is also enough to distinguish between Bayesian learning with a common prior and known ability.

¹⁴The subscript i is omitted where no confusion arises.

Table 8: Mann-Whitney tests for equal distributions of perf. under different regimes in the first month: reported prob. of identical distributions.

	Regime 1	Regime 2	Regime 3
Std. Dev.	0.465	0.454	0.457
Regime 1		Pr=0.907	Pr=0.564
Regime 2			Pr=0.408

Table 9: Summary statistics for the first 6 months (includes only workers who start and work under the same regime).

Variable:	Regime 1	Regime 2	Regime 3
β (in \$)	3.3	2.5	2.5
α (in \$)	3.8	3.5	3.5
Avg. Perf. (call/hr.)	2.74	2.6	2.66
Std. Dev. (Perf.)	0.68	0.77	0.7
Avg. Perf., stay \geq 6	2.91	2.76	2.71
Std. Dev. (Perf., stay \geq 6)	0.65	0.67	0.6
Turnover	0.52	0.68	0.67
Obs., stay $>$ 6	113	59	9

Observation 1 imposes necessary restrictions on the observed performance series that are strong. Table 5 presents the results from testing for the technology restrictions implied by Observation 1. A casual look at the standard deviations of performance in period 1 under the different regimes verifies the plausibility of the hypothesis of equal variance: the standard deviations vary between 0.45 and 0.47. This observation is confirmed by the results of the Mann-Whitney tests for equality of the demeaned distributions of performance under regimes 1, 2, and 3 in the first month: the tests fail to reject the hypothesis of equality of the demeaned distributions across regimes.

8 Appendix B¹⁵

This appendix present a slightly more general version of the model presented in the body of the paper. The main ingredients of the model are: a continuous individual time-invariant productivity parameter θ with a CDF denoted by F_θ ; an outside offer ξ , with a CDF denoted by F_ξ , equal to the utility that the worker will receive if she quits; if she stays, then she must decide on effort l_t , $l_t \in L \subset R_+$, on the basis of the piece rate $R_{it} = (\alpha_{it}, \beta_{it})'$ and the other relevant parameters of the utility function defined as

$$u(R_{it}, l_t, y_t) = \alpha_t + \beta_t y_t - \psi(l_t),$$

where y_t is the performance signal at t and ψ represents the disutility of labor; upon observing the performance signal y_t the agent updates her belief about θ denoted θ_t , whose CDF is denoted F_θ^t . The piece rate is taken to be exogenous by the employee and is not expected to change. The noise ε_t in the performance signal is continuous and iid over time with CDF denoted by F_ε .

The following assumptions on individual behavior are maintained throughout.

AA1. *The stochastic technology governing y_t is defined by (1) and is strictly increasing in its arguments, bounded, and jointly continuous.*

AA2. *$u(R_{it}, l_t, y_t)$ is strictly increasing in its first argument and strictly decreasing and strictly concave in its second argument, bounded and jointly continuous. L is compact.*

AA3. *ε_t, θ , and ξ are continuous. Furthermore, F_ε and F_θ are log-concave and have full support.*

AA4. *The sequence of signals $\{y_{ik} - g(k) + l_{ik}\}_{k=1}^t$ are ordered in the sense of the likelihood ratio property.*

The continuity assumptions on the production function and the distributions are necessary for the proof of existence. The monotonicity assumptions ensure monotonicity of the value

¹⁵In this section, I drop the individual subscript "i"

function and the optimal policy. The last assumption establishes a link between signals and beliefs. In this role, it has a crucial role in the characterization of the solution and the identification of learning.

$F_{\theta_{t+1}}$ provides the update of Bayesian beliefs given an output realization y_t . Since the stochastic technology is additively separable in θ and effort, effort choice does not affect the precision of posterior beliefs, so the transitional map $Q : P(\Theta) \times L \rightarrow P(P(\Theta))$ does not depend on effort choice in t and is the beliefs in the following period $t + 1$, conditional on the available information at t . The expected utility at time t becomes

$$U(R_t, l_t, t) = \int_{\Theta} \int_Y u(R_t, l_t, y_t) f(y_t | \theta_t, t, R_t, l_t) \gamma(dy_t) F_{\theta_t}(d\theta_t).$$

Then, the worker's dynamic program (P) is :

$$v(\theta_t, R_t, t) = \int_{\Theta} \left[\max[\xi_t, \max_{l_t \in L} [U(R_t, l_t, t) + \beta \int_{\Theta} v(\theta_{t+1}, R_t, t + 1) F_{\theta_{t+1}}(d\theta | \theta_t, t)]] \right] F_{\xi}(d\xi)$$

Under AA1-AA4, a unique continuous solution to this problem exists and the value function is convex in the appropriate sense, and the optimal policy is unique. These results are summarized in the following two propositions.

Proposition 1. *Under Assumptions AA1-AA4:*

i. *The functional equation (P) has a unique continuous solution $V(\theta_t, R_t, t)$ and the optimal policy*

$$A(\theta_t, R_t, t) = \{l_t \in L \mid (P) \text{ holds.}\}$$

is a continuous function.

ii. *Optimal effort, $l(R_t) > l(R'_t)$ if $R_t > R'_t$.*

iii. *$V(\theta_t, R_t, t) > V(\theta_t, R'_t, t)$ if $R_t > R'_t$, and $V(\theta_t, R_t, t)$ increases θ_t in the sense of the likelihood ratio property.*

Proof of Proposition 1:

Part (i). $B(\cdot)$ and $Q(\cdot)$ are continuous by Lemma 1 and 2 in Easley and Kiefer (1988), so the proof of existence is reduced to a problem which can be solved using Blackwell (1965). Define the operator T by

$$(Tw)(\theta_t, R_t, t) = \int_{\Theta} \left[\max[\xi_t, \max_{l_t \in L} [U(R_t, l_t, t) + \beta \int_{\Theta} w(\theta_{t+1}, R_t, t+1) F_{\theta_{t+1}}(d\theta|\theta_t, t)]] \right] F_{\xi}(d\xi)$$

Let C denote the set of bounded functions on $P(\Theta)$. Under the supnorm metric, $\|\cdot\|$, C is a Banach space. By the contraction mapping theorem, a contraction operator $T : C \rightarrow C$ has a unique fixed point and by Blackwell's contraction mapping lemma, T is a contraction if

(1). (Monotonicity) $w_1 \geq w_2$ implies $Tw_1 \geq Tw_2$ and

(2). (Discounting) there exists $\beta \in (0, 1)$, such that $T(w + c) \leq Tw + \beta c$, for any constant $c \geq 0$.

Consequently, to prove existence it is sufficient to show that **(i)** the operator T is a contraction and that **(ii)** T maps continuous bounded functions into the space of continuous bounded functions, C .

(i). This result follows by establishing that conditions (1) and (2) of the Blackwell's contraction mapping lemma are satisfied. It is obvious that if $w_1 \geq w_2$ uniformly, then $Tw_1 \geq Tw_2$. Furthermore, for discount factor β

$$\begin{aligned} T(w + c) &= \int \left[\xi, \max_{l \in L} (U + \beta w + \beta c) \right] dF_{\xi} \\ &< \int \left[\xi, \max_{l \in L} (U + \beta w) \right] dF_{\xi} + \beta c \\ &= Tw + \beta c \end{aligned}$$

(ii). Suppose that $w(\theta_{t+1}, R_t, t+1)$ is continuous. Since L is compact-valued,

$$\tilde{w}(\theta_t, R_t, t) = \max_{l_t \in L} \left[U(R_t, l_t, t) + \beta \int_{\Theta} w(\theta_{t+1}, R_t, t+1) F_{\theta_{t+1}}(d\theta|\theta_t, t) \right]$$

has a solution. Observe that L is a constant correspondence, so it is a continuous correspon-

dence and the theorem of the maximum applies, so $\tilde{w}(F_{\theta_t}, h_t)$ is continuous. The function $\max(a, b)$ is continuous if a and b are continuous, and the integral over ξ is also continuous if ξ is continuous. Thus, T is a contraction that maps bounded continuous functions into bounded continuous functions. The proofs of (i) and (ii) imply that a unique solution $V(\theta_t, R_t, t)$. By the theorem of the maximum, the optimal policy correspondence $A(\theta_t, R_t, t)$ is upper-hemicontinuous, and since $u(\cdot)$ is concave in l_t , the optimal policy is a continuous function.¹⁶

Part (ii). The absence of interaction between effort and beliefs makes the problem of choosing optimal effort static. Since the utility function obeys increasing differences in (β, l_t) , optimal effort $l(R_t) > l(R'_t)$ if $R_t > R'_t$.

Part (iii). Suppose that $R_t > R'_t$ and $V(\theta_{t+1}, R_t, t+1) > V(\theta_{t+1}, R'_t, t+1)$. Since ξ_t is constant in R_t , and $l(R_t) > l(R'_t)$, the first part of the statement follows. Finally, suppose that $V(\theta_t, R_t, t)$ increases in θ_{t+1} in the sense of the likelihood ratio property; then, the integral of $V(\theta_t, R_t, t)$ over the distribution of θ_{t+1} conditional on θ_t and t is also increasing in θ_t in the sense of the likelihood ratio property. Similarly, $u(\theta_t, R_t, l_t, t)$ increases in θ_t in the sense of the likelihood ratio property. Thus, $V(\theta_t, R_t, t)$ increases in θ_t in the sense of the likelihood ratio property. ■

¹⁶Note that nothing in the proof of this section requires that θ must be unidimensional. All proofs hold for an arbitrary finite number of unknown parameters. The assumption that θ is unidimensional is important in the following section, since vectors are only partially ordered.

9 Appendix C

There are 18 moment conditions and 5 parameters to be estimated. The conditions are nonlinear which complicates identification. By the assumption on the approximation of $G(\mu_{it}, R_{it}, t)$, the minimum distance problem has a solution. The following discussion addresses the uniqueness of that solution. The discount factor is identified from variation in $G(\mu_{it}, R_{it}, t + 1)$ with t that is associated with changes in the precision of beliefs. The results from step 1 show that by $t = 12$ the accumulation of experience has come to an end. Therefore, conditional on the information available at t , $t > 12$,

$$U(\mu_{it+1}, R_{it}, t, \gamma, \psi) - U(\mu_{it}, R_{it}, t, \gamma, \psi) = 0.$$

Consequently,

$$\begin{aligned} & G(\mu_{it+1}, R_{it}, t + 1) - G(\mu_{it}, R_{it}, t) \\ &= \delta [E_{\mu_{it+1}}(\lambda(G(\mu_{it+1}, R_{it}, t + 1))) - E_{\mu_{it+2}}(\lambda(G(\mu_{it+2}, R_{it}, t + 2)))] \end{aligned}$$

Conditional on the information available at t and $R_{it} = R_{it+1}$, variation in $G(\mu_{it}, R_{it}, t)$ across periods originates from changes in the precision of posterior beliefs. Therefore, variation in the first differences on the left-hand and the right-hand side of the condition above identifies the discount factor.

Given δ , the structural parameters γ and ψ are identified from $\gamma = \gamma(\psi, \widehat{\Delta l})$ and from variation in $G(\mu_{it}, R_{it}, t)$ with t that is associated with changes in the pay regime.

References

- [1] Abraham, K. (1988), "Flexible Staffing Arrangements and Employers' Short-term Adjustment Strategies", in *Robert A. Hart (ed.) Employment, Unemployment, and Labor Utilization*, Boston: Unwin Hyman.
- [2] Abraham K., Farber, H. (1987), "Job Duration, Seniority, and Earnings", *American Economic Review* **77** (3), 278 - 97.
- [3] Aghion, P., Bolton, P., Harris, C., and Jullien B. (1991), "Optimal Learning by Experimentation", *Review of Economic Studies* **58**, 621 - 654.
- [4] Aguirregabiria, V., Mira, P. (2007), "Sequential estimation of dynamic discrete games", *Econometrica* **75**, 1—53.
- [5] Altonji, J. G., Williams, N. (2005), "Do Wages Rise with Job Seniority? A Reassessment", *Industrial and Labor Relations Review*, **58**(3), 370-97.
- [6] Arcidiacono, P., Miller R. (2011), "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," *Econometrica*, **79** (6), 1823–1867.
- [7] Autor, D. (2008), "The Economics of Labor Market Intermediation: An analytical Framework", *NBER Working Paper 14348*.
- [8] Autor, D., Houseman, S. (2005), "Do Temporary Help Jobs Improve Labor Market Outcomes for Low-Skilled Workers? Evidence from Random Assignments", *NBER Working Paper No. 11743*.
- [9] Baker, G., Gibbs M., and Holmstrom, B. (1994), "The Internal Economics of the Firm: Evidence from Personnel Data", *Quarterly Journal of Economics* **109**(4), 881-919.
- [10] Bandiera, O., Barankay, I., and Rasul, I. (2007), "Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment", *Quarterly Journal of Economics* **122**(2), 729-73.
- [11] Bandiera, O., Barankay, I., and Rasul, I. (2005), "Social Preferences and the Response to Incentives: Evidence from Personnel Data", *Quarterly Journal of Economics* **120**(3), 917-62.
- [12] Bandiera, O., Barankay, I., and Rasul, I. (2009), "Team Incentives: Evidence from a Field Experiment", Working paper.
- [13] Bellman, R., Kalaba, R., and Kotkin, B. (1963), "Polynomial Approximation – a New Computational Technique in Dynamic Programming", *Mathematical Computation* **17**(8), 155–161.
- [14] Camargo, B., Pastorino, E., (2010), "Career Concerns and Learning-by-Doing", Working paper.

- [15] Bojilov, R.; 2010. "Estimating the Effects of Incentives When Workers Learn about Their Ability" Working paper.
- [16] Chiappori, P. A., Salanie, B., Valentin, J. (1999), "Early Starters or Late Beginners", *Journal of Political Economy* **107**(4), 731 - 762.
- [17] Easley, D., Kiefer, N. (1988), "Controlling a Stochastic Process with Unknown Parameters", *Econometrica* **56**(5), 1045-1064.
- [18] Farber, H. (1999), "Stability and Mobility: The Dynamics of Job Change in Labor Markets", *Handbook of Labor Economics* **3**, 2439 - 2483.
- [19] Ferrall, C., Shearer, B. (1994), "Incentives, Team Production, Transaction Costs and the Optimal Contract: Estimates of an Agency Model using Payroll Records", *CIRANO papers 94-12s*.
- [20] Fernie, S., Metcalf, D. (1996), " It's Not What You Pay", *Centre for Economic Performance Discussion Paper 295*.
- [21] Florens, J. P.; Heckman, J., Meghir, C., Vytlacil, E. (2008), "Identification of Treatment Effects Using Control Functions in Models with Continuous, Endogenous Treatment and Heterogeneous Effects", *NBER Working Papers: 14002*.
- [22] Gibbons, R., Katz, L. F., Lemieux, T., Parent, D. (2005), "Comparative Advantage, Learning, and Sectoral Wage Determination", *Journal of Labor Economics* **23**(4), 681-723.
- [23] Haley, M. (2003), "The Response of Worker Effort to Piece Rates: Evidence from the Midwestern Logging Industry", *Journal of Human Resources* **38**, 881–890.
- [24] Hansen, L. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica* **50**(4), 1029-54.
- [25] Harris, M., Holmstrom, B. (1982), "A Theory of Wage Dynamics", *Review of Economic Studies* **49**(3), 315 - 333.
- [26] Hausman, J., Wise, D. (1979), "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment", *Econometrica* **47**(2), 455-73.
- [27] Hoffmann, F. (2010), "An Empirical Model of Life-Cycle Earnings and Mobility Dynamics", Working paper.
- [28] Hotz, J., R.A. Miller (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models", *Review of Economic Studies* **60**, 497-529.
- [29] Holmstrom, B. (1979), "Moral Hazard and Observability", *Bell Journal of Economics* **13**, 74–91.
- [30] Holmstrom, B., Milgrom, P. (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives" *Econometrica* **55**(2), 303-28.

- [31] Houseman, S. (2001), "Why Employers Use Flexible Staffing Arrangements: Evidence from an Establishment Survey", *Industrial and Labor Relations Review* **55**(1), 149–170.
- [32] Jovanovic, B. (1979), "Job Matching and the Theory of Turnover", *Journal of Political Economy* **87**(5), 972 - 990.
- [33] Jovanovic, B., Nyarko, Y. (1994), "Learning by Doing and the Choice of Technology", *Econometrica* **64**(6), 1299 - 1310.
- [34] Kanemoto, Y., MacLeod, W. (1992), "The Ratchet Effect and the Market for Secondhand Workers", *Journal of Labor Economics* **10**(1), 85-98.
- [35] Kahn L., Sherer, P. (1990), "Contingent Pay and Managerial Performance", *Industrial and Labor Relations Review* **43**(3), 107-120.
- [36] Kveton, B., Hauskrecht, M. (2004), "Heuristic Refinements of Approximate Linear Programming for Factored Continuous State Markov Decision Processes", *ICAPS Working Paper 14*.
- [37] Lemieux, T., MacLeod, W., Parent, D. (2009), "Performance Pay and Wage Inequality", *Quarterly Journal of Economics* **124**(1), 1-49.
- [38] Lazear, E. (2000), "Performance Pay and Productivity", *American Economic Review* **90**(5), 1346 - 1361.
- [39] Lazear, E. (1995), *Personnel Economics*, MIT Press: Cambridge, MA.
- [40] Lazear E., Oyer, P. (2011), "Personnel Economics", *Handbook of Labor Economics*, forthcoming.
- [41] Lazear, E., Shaw, K. (2008), "Tenure and Output", *Labor Economics* **15**(4), 704-723.
- [42] MacLeod, W., Malcomson, J. (1989), "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment", *Econometrica* **57**(2), 447-80.
- [43] Mas, A., Moretti, M., (2009), "Peers at Work", *American Economic Review* **99**(1), 112-45.
- [44] McMillan, J., Whalley, J., Zhu, L. (1989), "The Impact of China's Economic Reforms on Agricultural Productivity Growth", *Journal of Political Economy* **97**(4), 781-807.
- [45] Miller, R. (1984), "Job Matching and Occupational Choice", *Journal of Political Economy* **92**, 1086-1120.
- [46] Mortensen, D. (1988), "Wages, Separations, and Job Tenure: On-the-job Specific Training or Matching?", *Journal of Labor Economics* **6**(4), 445 - 471.
- [47] Nagypal, E. (2007), "Learning by Doing and Learning about Match Quality: Can We tell Them Apart?" *Review of Economic Studies* **74**, 537 - 566.
- [48] Neumann, M., von Sachs, R. (2000), "A Wavelet-Based Test for Stationarity", *Journal of Time Series Analysis* **21**, 597–613.

- [49] Nijman, T., Verbeek, M. (1992), "Testing for Selectivity Bias in Panel Data Models", *International Economic Review* **33**(3), 681-703.
- [50] Pastorino, E. (2009), "Learning in Labor Markets and Job Mobility", Working Paper.
- [51] Paarsch, H., Shearer, B. (1999), "The Response of Worker Effort to Piece Rates: Evidence from the British Columbia Tree-Planting Industry", *Journal of Human Resources* **34**, 643-667.
- [52] Paarsch, H., Shearer, B. (2000), "Piece Rates, Fixed Wages, and Incentive Effects: Statistical Evidence from Pay Roll Records", *International Economic Review* **41**, 59-92.
- [53] Paarsch, H., Shearer, B. (2009), "The Response to Incentives and Contractual Efficiency: Evidence from a Field Experiment", *European Economic Review* **53**(5), 481-494.
- [54] Pakes, A., Ericson, R. (1999), "Empirical Implications of Alternative Models of Firm Dynamics", *Journal of Economic Theory* **79**, 1-45.
- [55] Pesendorfer, M., Schmidt-Dengler, P. (2003), "Identification and Estimation of Dynamic Games", *NBER working paper 9726*.
- [56] Prendergast, C. (1999), "The Provision of Incentives in Firms," *Journal of Economic Literature* **37**(1), 7-63.
- [57] Prendergast, C. (2002), "The Tenuous Trade-Off between Risk and Incentives" *Journal of Political Economy* **110**(5), 1071-1102.
- [58] Priestley, M., Rao, T. (1969), "A test for Non-Stationarity of Time-Series." *Journal of the Royal Statistical Society, Series B*, **31**(1), 140-149.
- [59] Pugatch, T. (2012), "Bumpy Rides: School to Work Transitions in South Africa." Working Paper.
- [60] Shearer, B. (2004), "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment", *Review of Economic Studies* **71**(2), 513-34.
- [61] Stiglitz, J., (1991), "Symposium on organizations and economics", *Journal of Economic Perspectives* **5**, 15-24.
- [62] Superb Staff Services (2011), "Study of Call Center Costs", available through <http://www.sashacorp.com/turnframe.html>.
- [63] Taylor, F. W. (1911), *The Principles of Scientific Management*, New York, NY: Harper & Brothers.
- [64] Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press: Cambridge, MA.