Motivate and Select: Relational Contracts with Persistent Types

Radoslawa Nikolowa†

December 12, 2016

Abstract

We develop a model of relational contracts with moral hazard and asymmetric persistent information about an employee’s type. We find that the form of the optimal contract depends on the job characteristics and the distribution of employees’ talent. Bonus contracts are more likely to be adopted in complex jobs and when high talent is not too common or too rare. Firms with ‘normal’ jobs are more likely to adopt termination contracts. In labor market equilibrium, different contracts may be adopted by ex ante identical firms. Hence, we offer an explanation for the co-existence of different employment systems within the same industry.

Keywords: Relational contracts, Job characteristics, Employment systems, Labor market segmentation

†School of Economics and Finance, Queen Mary University of London; r.nikolowa@qmul.ac.uk, Mile End Road, London E1 4NS, United Kingdom; tel. +44(0)20 7882 3996

I would like to thank the editor Professor Casadesus-Masanell, the associate editor, and two anonymous referees for their thorough and helpful comments. The paper has benefited from helpful comments by Daniel Ferreira, Luis Garicano, David Martimort, Jeannine Miklos-Thal, Bernard Salanié, and the participants of the third edition of the conference “Industrial Organization: Theory, Empirics and Experiments”, the Econometric Society meeting in Toulouse, and the Microeconomics seminar in Tilburg. The usual disclaimer applies.
JEL classification:  D82, M5, L14
1 Introduction

Employment contracts are often incomplete and informal reputational mechanisms are prominently used in organizations where informal codes of conduct affect the behavior of agents. A growing empirical literature documents employers’ use of firing, promotions and bonuses based on informal agreements to motivate their employees.\textsuperscript{1} Furthermore, MacLeod and Parent (2015) empirically show that reputational mechanisms can provide insight into the form of compensation and that job characteristics affect the nature of the compensation schemes offered to employees.

In this paper, we develop a model of self-enforceable relational contracts when the employee’s type is his private information and the job requires handling tasks with different levels of difficulty. The model allows us to understand how the characteristics of the job, the labor market, or the firm affect the compensation package used by employers and their hiring and firing policies. We also show that in labor market equilibrium, a priori identical firms may choose to offer different yet equally profitable contracts that feature different hiring and firing policies. Hence, the paper also helps explain why different employment systems may co-exist within the same industry. In the US manufacturing industry, some firms have adopted flexible organizational structures with employees’ involvement and teamwork, while others have adopted a more rigid job structure aimed at cutting costs (Osterman, 1994). Hunter (1999) offers evidence of the co-existence of different employment models in the US retail banking sector, and Bailey and Sandy (1999) examine the

\textsuperscript{1}MacLeod and Parent (1999) analyze different incentive schemes used in the US and show that 14% of the individuals in their data set received informal bonus pay. Hayes and Shaefer (2000) provide evidence for the use of subjective performance measures when boards of directors decide the salaries and bonuses of chief executives. Cappelli and Chauvin (1991) show that termination contracts are used as incentive devices. Altmann et al. (2014) provide experimental evidence regarding the link between contractual incompleteness and unemployment. They also find that contractual incompleteness leads to the co-existence of different job types, i.e., jobs that offer different wages and performance requirements.
different practices in the apparel industry.

In the model, a continuum of heterogeneous workers and a continuum of homogeneous employers play an infinitely repeated game. Each employee privately observes his talent. An employee’s job consists of handling one task in each period, and the firm owner’s problem is to motivate the employee to successfully perform this task. To succeed in a task, an employee needs ability and effort. The more difficult a task is, the more able an employee must be to successfully perform it. Performance is observable within the relationship but cannot be verified by a third party such as a court of law. An employer will be able to credibly promise a performance-based bonus only when the rent from retaining an employee is sufficiently high. Because we focus on cases where jobs are in short supply, the only rent for the employer comes from the potential difference in talent between the current employee and the average unemployed person. Consequently, a bonus can be used to motivate employees with high talent to exert effort only when high-talented employees are sufficiently rare.

Alternatively, the employer may offer a pooling termination contract with a fixed wage and a performance requirement that the employee must achieve to remain employed. In that case, the employee exerts effort to retain his job. The vacancy is filled in every period; however, if an employee is unable to reach the performance requirement, he is fired. The tenure of employees in jobs where an efficiency wage contract is used is shorter than the tenure in jobs in which a bonus contract is offered.

Using a bonus contract comes at the cost of not being able to fill the vacancy when matched with a low-talent employee, and a termination contract comes at the cost of a rent that the firm needs to offer to the employees. It follows that the higher the likelihood is to be matched with a high-talent employee, the higher the benefit is from adopting a
bonus contract. However, if the fraction of high-talent employees is too high, the principal cannot credibly promise to pay the bonus. Therefore, a bonus contract is more likely to be adopted when high talent is not too common or too rare.

For a given distribution of talent, if agents with low talent are able to solve fewer tasks, then a company is more likely to adopt a bonus contract. Indeed, in jobs for which most of the tasks require the expertise of a highly talented employee, the opportunity cost of not hiring low-talent employees is lower. Hence, bonus contracts are more likely to be found in ‘complex jobs’. In ‘normal’ jobs, i.e., those in which low-talent employees are able to solve a larger fraction of tasks, an efficiency wage contract is more likely to be used. Whether a job should be labeled as ‘normal’ or ‘complex’ is not exogenously set, and the model allows for the nature of a job to change if a change in the technology, for example, affects the distribution of tasks that an employee needs to handle.

The decision to adopt a bonus or a termination contract depends on the quality of the pool of available employees. However, the quality of that pool is endogenous and depends on the contracts offered by the employers. Because a bonus contract is only accepted by high-talent employees, if all firms offer a bonus contract, the probability of finding a high-talent employee in the pool of unemployed is very low. The low probability of being matched with a high-talent employee makes the termination contract more appealing to employers, and a market equilibrium where both bonus and termination contracts co-exist may emerge.

**Related literature.** Specifically, the paper contributes to the literature on relational contracts. The early literature on relational contracts focuses on frameworks with symmetric information (see Shapiro and Stiglitz (1984) and Bull (1987), among others). MacLeod
and Malcomson (1989) proposed the first treatment of the problem. They show that provided the surplus created by the continuation of the employment relationship is sufficiently high, either a performance-based bonus or a termination contract can be sustained in equilibrium.\(^2\) MacLeod and Malcomson (1998) go further and stress the importance of labor market conditions in the choice between a termination contract and a bonus contract. When agents are homogeneous in skills and there is unemployment, the principal can replace an employee with an unemployed agent at no cost. There is no rent for the employer to continue the employment relationship. Thus, the bonus is not credible, and a termination contract is the only way to motivate employees.\(^3\) Conversely, when the latter are in short supply, vacancies cannot be immediately filled, which creates rent from retaining an agent. Hence, whenever the cost of unoccupied jobs is not too high, an equilibrium with full employment and bonus could emerge. Differently from these papers, we consider that employees are heterogeneous in their ability, and this ability is unknown to the employer. 

Levin (2003) analyzes relational contracts when the agent’s type is private information. However, he assumes that the type is non-persistent, i.e., it changes from period to period, and the focus of his analysis is on the restrictions on revelation due to the self-enforcing character of the agreement. Yang (2013) considers a setting of relational contracts with asymmetric persistent information about the employee’s type and non-contractible performance. The focus of his paper is on demonstrating that the optimal pooling contract is non-stationary and on deriving implications in terms of wage progressions over time. In his setting, the separating contract is always optimal, and the arguments against it being used by employers are exogenous to the model. Differently from his paper, we show

\(^2\)Because both the principal and employees are risk-neutral, the reward is a pure transfer.

\(^3\)Yang (2008) shows that a bonus contract can be offered in a market with homogeneous employees when firms incur exogenous turnover costs. In our paper, the turn-over cost is endogenous and depends on the talent of the average unemployed agent.
that depending on the characteristics of the job - complex versus normal - and the labor market conditions, a firm may optimally choose a separating bonus contract or a pooling termination contract. Additionally, in the present paper, different types of contracts may coexist in labor market equilibrium.

Halac (2012) studies relational contracts with persistent private information in a different setting. Her paper considers the dynamics of a relationship in which the principal has private persistent information about her outside option.

More generally, the present work is also related to the literature on the co-existence of employment contracts. Assignment models, such as, for example, Besley and Ghatak (2005), Alonso-Paulí and Pérez-Castrillo (2012) and Macho-Stadler et al. (2014), assume that the firms (and in some cases the workers as well) are heterogeneous, and the adoption of different types of contracts results in the assignment of specific worker types to specific jobs. In our model, firms are homogeneous, and workers and firms are randomly matched. A firm’s decision regarding the type of contract to offer to a potential employee depends on the firm’s beliefs about the employee’s talent. A bonus contract is preferred when there is a high probability of being matched with a high-talent employee. However, as more firms offer bonus contracts, a firm with a vacancy becomes less likely to be matched with a high-talent employee, which makes the adoption of a termination contract more profitable.

Board and Meyer-ter-Vehn (2015) show that wage dispersion may arise in equilibrium in a relational contract setting when employees are allowed to search for new employment opportunities while on the job. Differently from our paper, they assume that

\[\text{Burdett and Mortensen (1998) also show that labor market segmentation can arise in a model with homogeneous workers and firms when workers search while on the job. However, the focus of their paper is on the possibility of attracting more or less workers rather than on providing incentives to exert more or less effort.}\]
employees are homogeneous in terms of talent. In equilibrium, firms offer different wages associated with different performance requirements, and employees exert effort to stay in employment. Because we assume that employees are heterogeneous in their ability, performance-based bonuses can be used by firms to incentivize effort and select talented employees. Hence, our insight concerns the co-existence of employment systems such as performance-based bonus contracts and termination contracts.\footnote{In an experimental study, Bartling et al. (2012) show the existence of complementarity between high wages, effort discretion, and the possibility of observing information about an employee's past performance, and the authors offer insight into the bundling of employment practices into 'control' vs. 'trust' strategies. However, the co-existence of these practices under a specific experimental treatment results from sub-optimal choices made by some of the agents.} On the one hand, we have better-selected employees who work in order to receive performance-based bonuses and have longer tenure in their job. On the other hand, the employees are less selected \textit{ex ante}, they only receive fixed wages and exert effort in order to keep their job, they are subject to higher turnover, and they may handle a more narrow set of tasks.

Finally, our paper also relates to Lazear (2000). He shows that piece-rate contracts can be effective in inducing self-selection. We bring this idea to the relational contract setting and show that bonus contracts can be used in order to hire only the best employees.

The rest of the paper is organized as follows. Section 2 presents the setup of the model and a benchmark case in which performance is verifiable. Section 3 characterizes the optimal relational contract adopted by an individual employer for given characteristics of the pool of unemployed agents and given the outside option of the employees. Section 4 presents the equilibrium in the labor market and shows that different types of contracts may coexist in equilibrium. Then, we conclude. All proofs can be found in the Online Appendix.
2 Model

2.1 The setting

We consider a continuum of firms (employers, principals) with mass $n$ and a continuum of potential employees (agents) with a mass of one. Each firm has one position to fill and jobs are in short supply $n < 1$. The firms and the agents live forever, are risk neutral, and share a common discount factor, $\delta$. Time is discrete. Firms are identical and each firm has one job. Each agent is endowed with ability (talent) $\theta$, where $\theta \in \{\bar{\theta}, \tilde{\theta}\}$, with $\bar{\theta} > \tilde{\theta}$. An employee is privately informed about his ability. An employer observes the distribution of types in the pool of available agents: $\text{Prob}(\theta = \bar{\theta}) = p$ and $\text{Prob}(\theta = \tilde{\theta}) = 1 - p$.

To introduce the idea of task heterogeneity, we use a production function in the spirit of Garicano (2000). The production function describes particularly well those jobs that require handling clients’ requests, such as customer service jobs, loan approval processing jobs in a retail bank, etc.

Production requires tasks (problems) to be performed successfully. These tasks are equally valuable to the firm but differ in terms of their difficulty. The revenue for a firm from a successfully performed task is normalized to one. In each employment period, an employee receives a task with difficulty $x$, drawn from a commonly known distribution $F(\cdot)$ on the bounded support $[0, \bar{\theta}]$, i.i.d. over time. An employee who receives a task $x$ chooses his level of effort $e$. The choice is binary; the employee either works $e = e$ or shirks $e = e$.

An employee succeeds in resolving a problem with a probability $e$, only if the problem is

---

If employees were in short supply $n > 1$, there would be a rent from retaining an employee due to the scarcity of the workforce. The impact of this effect on the optimal choice of contract is discussed by MacLeod and Malcomson (1998). By focusing on the cases with unemployment, any rent from retaining an employee is driven by the fact that employees are heterogeneous.

In Garicano (2000), succeeding in a task requires only knowledge, and the paper’s focus is on the optimal organizational structure.
within the employee’s ability set \((x \in [0, \theta])\). That is, an employee cannot solve a task with difficulty \(x > \theta\). Working is costly \(c(x, \theta, \bar{e}) = c\) for \(x \leq \theta\), while shirking is costless \(c(x, \theta, e) = 0\). For simplicity, we normalize \(\bar{e} = 1\) and \(e = 0\).

The modeling adopted here is appealing for several reasons. First, it enables us to introduce and discuss the impact of job characteristics on the choice of the optimal relational contract. Second, it allows us to simplify the analysis without losing in terms of generality. Throughout the paper, we discuss the assumptions and their implications on the generality of the results.

At the beginning of a period, if a job is vacant, the employer is randomly matched with an agent from the pool of the unemployed. In each employment period \(t\), the principal offers a contract to the employee. The employee decides to accept or reject the contract. If the contract is rejected, the position remains unoccupied for that period and the profit is zero, and the employee returns to the unemployment pool and receives his inter-temporal outside utility \(V_{out}\). The contract consists of a fixed wage \(w_t\) that is paid independently of the worker’s performance, a bonus schedule \(b_t(x)\) that the firm agrees to pay for any successfully performed task with difficulty \(x\), and possibly a minimum performance requirement, \(y_t\) that the employee is expected to achieve to retain his job. The firm is committed only to the wage of the current period. An employee who has accepted a contract receives a task with difficulty \(x\) and decides the level of effort \(e \in \{0, 1\}\).

At the end of each employment period, the employee and the employer observe the difficulty of the task and whether it has been successfully performed. Once payments have been made (or not), each party can decide to leave the current relationship. With probability \((1 - \alpha)\), an employer-employee match becomes unprofitable for exogenous
reasons. We assume that $\alpha$ is independent of the number of periods an employee has been in a relationship with the employer, and it also does not depend on past performance. After separation, the employee joins the pool of unemployed agents and receives his outside option, while the employer has a vacant position to fill. The timing of one period of employment is summarized in Figure 1 below.

![Figure 1: Timing of the employment relationship](image)

We assume that an agent who has been separated from an employer cannot be immediately re-employed and spends one period in the unemployment pool.\(^8\) For each period of unemployment, an agent receives an unemployment benefit $k$. At the end of an unemployment period, an agent is matched with an employer with probability $\lambda$, where $\lambda$ is

---

\(^8\)Assuming that an agent remains unemployed for at least one period is a convention that we adopt. Because jobs are in short supply, this is without loss of generality.
endogenously determined in the market equilibrium.

The expected one-period surplus from employing an agent with ability \( \theta \) is \( F(\theta)(1 - c) - k \). For the remainder of the paper, we will focus on those cases in which the surplus created from employing an agent with type \( \theta \) is positive.

**Assumption 1.** \( F(\theta)(1 - c) - k > 0 \)

It follows that it is socially optimal to hire any employee with whom a firm is matched (at least) for the current period.

Before analyzing the incomplete contracting case, we offer a benchmark where each employer can offer a legally enforceable contract based on the employee’s performance.

### 2.2 Benchmark: The market equilibrium with commitment

Let us assume that a principal can credibly offer a contract contingent on the employee’s performance in the current period, i.e., the contract is contingent on the difficulty of a task an employee has successfully resolved or not. Each employer is too small to have an individual impact on the market conditions and solves for the optimal contract taking the distribution of talent in the unemployment pool and the employee’s outside option as given. Focusing only on short-term contracts is without loss of generality because - as we will see - a menu of contracts with short-term commitments implements the socially optimal solution and allows each principal to extract the entire surplus created by the employment relationship.

**Proposition 1.** *When a principal can commit to paying a bonus conditioned on output,*
the following menu of contracts is optimal:

\[
\begin{aligned}
C &= (w = (1 - \delta)V^{out}, b(x \in [0, \theta]) = c, b(x \in [\theta, \bar{\theta}]) = 0) \\
\bar{C} &= (w = (1 - \delta)V^{out} - (1 - F(\theta))\epsilon, b(x \in [0, \theta]) = c, b(x \in [\theta, \bar{\theta}]) = c + \epsilon),
\end{aligned}
\]

where \( \epsilon \geq 0 \). An employee with type \( \theta \) selects \( C \), and an employee with type \( \bar{\theta} \) selects \( \bar{C} \).

In the case with commitment, the employer’s choice is independent of the distribution of talent in the unemployment pool. It is in the interest of each firm to offer a menu of bonus contracts that will motivate any employee to exert effort on all tasks within his ability set and to truthfully reveal his type and that satisfies the employee’s participation constraint with equality. There is no profitable deviation for an individual employer; therefore, all firms offer a menu of contracts as in Proposition 1.

3 The optimal relational contract

In what follows, we assume that performance (i.e., whether a problem has been solved or not and the difficulty of that problem) is observable by the employer and the employee but cannot be verified by a third party, and the principal cannot write a legally enforceable contract based on performance. However, a principal and an agent can agree on a relational contract enforced by the possibility of future actions for each of them. A contract is self-enforceable if it is in the interest of both parties to abide by it.

The solution concept is perfect public equilibrium. Public strategies require that each party’s strategy be dependent on the public history within the current relationship. For any public history, strategies form a Nash equilibrium from that point on.

We make the following assumption:
Assumption 2. The labor market is anonymous.

The reasons that a particular employer-employee match is terminated are not observed by any party in the labor market, whether employers or employees. The idea behind this is that monitoring the reasons for a separation is difficult. The implication of Assumption 2 is that it is impossible to build an external reputation. If either party is allowed to build an external reputation, then not respecting the contract leads to a loss of cooperation of current and future partners. This loss of cooperation would increase the contractor’s incentives to respect the terms of the agreement.

The public history of an employee’s performance within a relationship boils down to knowing whether the employee has been able to solve a problem with difficulty \([0, \theta]\) and/or \((\theta, \bar{\theta}]\) conditional on the contract that he has been offered.\(^9\) An employee’s strategy consists of deciding whether to accept the contract and, after receiving a task, whether to exert effort. A strategy for the firm specifies whether to fire a worker and what spot contract to offer, with both as a function of public history.

Finally, no principal is ‘large enough’ to have an impact on market conditions. We concentrate first on solving the problem of an individual employer who takes the employees’ outside option \(V^{out}\) and unemployed agents’ distribution \(\bar{p}\) as given. The focus of the analysis is on stationary market equilibria, so we assume that the contracts offered by other firms are such that the outside conditions \(\bar{p}\) and \(V^{out}\) are stationary. The endogenous values for \(V^{out}\) and \(\bar{p}\) are then consistently determined in market equilibrium. In this section, we assume that \(V^{out}\) is independent from an employee’s type. In Section 4, we show that in our setting, the equilibrium inter-temporal utility of unemployed agents

\(^9\)In general terms, the employee’s history can be summarized by the most difficult problem the employee has resolved and the least difficult one he has failed to resolve, conditional on the contract that he has been offered.
is independent of their talent.

3.1 Bonus contracts and \textit{ex ante} selection

A performance-based bonus is the optimal way to motivate employees to exert effort in the complete contract setting. Thus, we start by considering the use of a bonus contract when performance is not verifiable.

Lemma 1. In equilibrium, a bonus contract cannot be used to motivate employees with talent $\theta$ to exert effort.

A performance-based bonus effectively motivates an employee to exert effort only if it is credible that the employer will pay it after observing the employee’s performance. For this to be the case, the rent the employer obtains from continuing the relationship with a successful employee must be sufficiently high. In our framework, the only rent for the employer comes from the potential difference in talent between the current employee and the average unemployed agent. There is no such rent from retaining an employee with revealed ability $\theta$ or with expected ability $p\theta + (1-p)\bar{\theta}$.\footnote{If the principal offers a pooling bonus contract, observing that a newly hired employee has successfully resolved a problem $x \in [0, \bar{\theta}]$ does not reveal the employee’s type, and the principal’s updated belief about the expected ability of this employee is $p\theta + (1-p)\bar{\theta}$. Because the latter is equal to the expected talent of an unemployed agent, there is no rent from retaining such an employee, and therefore, the bonus cannot be credibly offered.}

It follows that to use a bonus to motivate an employee with talent $\bar{\theta}$ to exert effort on all tasks $x \in [0, \bar{\theta}]$, the principal must separate the employees \textit{ex ante} by offering a contract that is only acceptable for high-talent employees. An employee with talent $\theta$ rejects the contract and takes his outside payoff $V^{\text{out}}$.\footnote{Full separation is specific to the assumption that an employee’s talent can only be of two types. In a model where the talent $\theta$ can be of more than two types, the equilibrium will exhibit some degree of pooling.}
Lemma 2. Subject to self-enforceability, the contract \(C_b = (b(x \in [0, \overline{\theta}]) = c + \epsilon, w = (1 - \delta)V^{out} - \epsilon\), where \(\epsilon > 0\), is only accepted by employees with talent \(\overline{\theta}\) and motivates them to exert effort on all tasks \(x \in [0, \overline{\theta}]\).

With a separating contract, the firm learns the type of a new worker in the first period of employment. In that case, the contracting environment between the firm and the worker is stationary, and therefore, focusing on stationary contracts is without loss of generality, as shown by Levin (2003), for example.

The self-enforceability constraint must guarantee that the principal is strictly better off by paying the bonus to an employee with talent \(\overline{\theta}\) who has successfully performed a task \(x \in [0, \overline{\theta}]\) rather than not paying it:

\[
(1) \quad \Pi_b(\overline{\theta}) - \Pi^{dev}_b \geq \frac{c + \epsilon}{\alpha \delta}.
\]

\(\Pi_b(\overline{\theta})\) is the inter-temporal expected profit from employing a worker with talent \(\overline{\theta}\)

\[
(2) \quad \Pi_b(\overline{\theta}) = (1 - c) - (1 - \delta)V^{out} + \alpha \delta \Pi_b(\overline{\theta}) + (1 - \alpha)\delta \Pi_b,
\]

where \(\Pi_b\) is the principal’s inter-temporal expected profit from offering the contract \(C^b\) to an agent he is matched with:

\[
(3) \quad \Pi_b = \bar{p}\Pi_b(\overline{\theta}) + (1 - \bar{p})\delta \Pi_b.
\]

\(\Pi^{dev}_b\) is the inter-temporal expected profit from deviating (i.e. not paying the bonus
after an employee exerts effort and solves a problem):

\begin{equation}
\Pi_b^{dev} = \overline{\pi}(1 + \epsilon - (1 - \delta)V^{out}) + \delta \Pi_b^{dev}.
\end{equation}

To effectively deter reneging by the firm, we assume that the employees follow the trigger strategy: exert effort in \( t \) only if the company has paid the bonus in period \( t - 1 \); otherwise, quit immediately.\(^\text{12}\) This strategy corresponds to the most severe punishment, and the penalty for the employer is given by the alternative available in the market.\(^\text{13}\)

Note that the higher the bonus is (i.e., the higher \( \epsilon \)), the more difficult it is to satisfy the self-enforceability constraint.\(^\text{14}\) Therefore, the largest set of parameters for which the contract of Lemma 2 can be credibly implemented is obtained for \( \epsilon \to 0 \).

**Proposition 2.** If \( \overline{\pi} < \pi^{SE} \) (where \( \pi^{SE} \equiv 1 - c \sigma \delta V^{out} \)), then a contract \( C_b = (b(x \in [0, \overline{\theta}]) = c + \epsilon, w = (1 - \delta)V^{out} - \epsilon) \) with \( \epsilon \to 0 \) is self-enforceable.

The bonus can be credibly used to motivate high-talent employees only if their talent is sufficiently rare. If the principal can be easily matched with another employee with high talent, then the rent from continuing the ongoing employment relationship is not sufficiently high, and the principal is likely to renege on the payment of the bonus.

\(^{\text{12}}\)It is in the interest of an employee to stop exerting effort if he believes that the employer will continue to cheat in the future. If the contract \( C_b \) is such that \( \epsilon > 0 \), then an employee who is not exerting effort is strictly better off by quitting the firm. If \( \epsilon = 0 \), the employee is indifferent between staying or leaving, but the principal is better off firing the current employee and hiring a new one from the pool of unemployed.

\(^{\text{13}}\)We recall that under Assumption 2, a principal cannot build an external reputation.

\(^{\text{14}}\)If we were to consider a more general cost function that increases with the difficulty of the task, for example, then the principal could offer a bonus that depends on the level of difficulty of the task performed by an employee, and the strongest self-enforceability constraint would be the one for the highest bonus offered by the principal. In that case, the principal may be able to motivate an agent with talent \( \overline{\theta} \) to only solve a subset of tasks for which the bonus is credible.
3.2 Efficiency wages and *ex post* selection

An alternative mechanism to motivate employees to exert effort is to guarantee sufficiently high gains for the employee from continuing the employment relationship. In that case, the employee exerts effort to retain his job. The termination contract in our setting is the pair \( (w_t(y_t, \theta), y_t) \), where \( w_t(y_t, \theta) \) is the fixed wage offered to employees with talent \( \theta \) at the beginning of period \( t \) of the employment relationship and \( y_t \) is a minimum performance requirement, i.e., employees are expected to solve any task \( x \leq y_t \) that they receive. If such a task is not resolved, the employee is fired. Firing an employee who fails at a task \( x < \min\{\theta, y_t\} \) is in the principal's interest if he believes that the employee will continue shirking in the future.\(^\text{15}\) If an employee has failed at a task \( x \in (\theta, y_t] \), the principal updates his beliefs about the employee, and firing the employee again becomes the best response.

Let \( V_t(\theta, w_t(y_t), y_t) \) be the inter-temporal expected utility of an employee with talent \( \theta \) who exerts effort on tasks \( x \leq \min\{\theta, y_t\} \) in period \( t \) of the employment relationship, when the efficiency wage contract is \( (w_t(y_t, \theta), y_t) \). The incentive compatibility constraint in employment period \( t - 1 \) of an agent \( \theta \) who has received a task \( x \leq \min\{\theta, y_{t-1}\} \) is as follows:

\[
-c + \alpha \delta V_t(\theta, w_t(y_t, \theta), y_t) \geq \alpha \delta V^{out}
\]

From equation (5), we notice that in period \( t - 1 \), an employee is motivated by future rents, i.e., the fixed wage of period \( t \). However, if the firm owner decides to renegotiate and offer a wage \( w'_t \) lower than the perceived 'fair wage' \( w_t \) in the beginning of \( t \), the

\(^{15}\)Any employee is able to solve tasks \( x < \min\{\theta, y_t\} \).
worker would respond by shirking. It is indeed in the employee's interest to do so if he anticipates that an employer who reduces the wage in this period will also do so in future periods.

**Lemma 3.** To motivate both types of agents to exert effort, the principal offers a pooling termination contract.

The intuition for this result is as follows. An agent exerts effort to retain his job. An employee with revealed talent $\theta$ is less able than the average unemployed worker; therefore, a low-talent employee would only be retained if the principal’s expected profit from such an employee is the same as his expected profit from a high-talent employee, which in turn implies that the inter-temporal expected profit of the principal when using a separating menu of contracts will never exceed his expected profit from an agent with talent $\theta$. It follows that a pooling contract dominates a menu of separating termination contracts. The formal proof is in Section 4 of the Online Appendix.

**Proposition 3.**

i) The minimum performance requirement $y_t$ is never lower than $\theta$.

ii) If $\bar{p} > c$, then the optimal efficiency wage contract is stationary, the performance requirement is $y = \bar{\theta}$, and the fixed wage is $w(\bar{\theta}) = \frac{c}{\alpha \delta} + (1 - \delta) V^{out}$.

$\bar{p} > c$ is a sufficient condition for $y = \bar{\theta}$. When $\bar{p}$ becomes very low, the optimal efficiency wage contract is still stationary, but the minimum performance requirement is $y = \theta$ and $w(\theta) = \frac{c(1 - \alpha \delta)}{\alpha \delta} + F(\theta)c + (1 - \delta) V^{out}$. The stationarity of the optimal termination (and bonus) contract is driven by the specifics of the model, the fact that talent can only be of two types, and the way in which information is revealed to the
principal over time. If we consider a setting with more than two types, the optimal contract may exhibit some non-stationarity.

### 3.3 The optimal contract

Here, we present the optimal relational contract for an individual employer – for a given employees’ outside option – depending on the value of \( p \). We consider that \( F(\theta) > w(\theta) \); otherwise, a credible bonus contract is always more profitable than a termination contract.

**Proposition 4.**

i) For \( p \in [\bar{p}^{SE}, 1] \), the optimal relational contract is \((w(\bar{\theta}), \bar{\theta})\).

ii) For \( p \in [\max\{p'', \hat{p}\}, \bar{p}^{SE}] \), the optimal relational contract is a bonus accepted only by the employees with talent \( \bar{\theta} \).

iii) For \( p \in [\min\{p', \hat{p}\}, \max\{p'', \hat{p}\}] \), the optimal relational contract is \((w(\bar{\theta}), \bar{\theta})\).

iv) For \( p \in [0, \min\{p', \hat{p}\}] \), the optimal relational contract is \((w(\bar{\theta}), \bar{\theta})\).

\( p'' \) is given by: \( \Pi(w(\bar{\theta})) = \Pi_b \), \( p' \) is determined by: \( \Pi(w(\bar{\theta})) = \Pi(w(\bar{\theta})) \), \( \hat{p} \) is obtained from \( \Pi(w(\bar{\theta})) = \Pi_b \), and \( p^{SE} \) is the threshold defined in Proposition 2 and, below which the bonus contract is self-enforceable.

A bonus contract is more likely to be used in a context where employees with high talent are neither too rare nor too common. Indeed, if the probability of being matched with a high-talent employee is high, then an employer has poor incentives to pay the bonus after observing high effort. Therefore, a bonus contract cannot be credibly offered. On the other hand, the cost of using a bonus contract relative to a termination contract is that low-talent employees are never hired and the position remains vacant unless the
employer is matched with a high-talent employee. The lower the probability is of being matched with a high-talent employee, the higher the cost of offering a bonus contract will be. Hence, for low values of $\overline{p}$, an employer is more likely to offer a termination contract.

The choice between a termination contract with a high ($y = \overline{\theta}$) or low ($y = \underline{\theta}$) performance requirement depends on whether the cost of paying a higher fixed wage to all employees is lower than the benefit of motivating high-talent employees to solve all tasks in their skill set.

**Corollary 1.** $p', \overline{p}'$, and $\tilde{p}$ increase with the fraction of tasks that can be solved by an employee with talent $\theta$.

The results of Proposition 4 and Corollary 1 are illustrated through an example in Figure 2.

For a given distribution of talent in the unemployment pool, if an employee with talent $\underline{\theta}$ is able to solve a lower fraction of tasks, then a company is more likely to adopt a bonus contract. A low $F(\underline{\theta})$ corresponds to low productivity of the employees with talent $\underline{\theta}$. Therefore, the opportunity cost of not hiring low-talent employees is lower, and a bonus contract is more likely to be adopted by an employer. The opposite is true for jobs where $F(\underline{\theta})$ is high. In that case, a termination contract is more likely to be offered.

$F(\theta)$ captures both the characteristics of the job and the characteristics of the employees. For example, for a given talent of employees $\theta$, a higher $F(\theta)$ would correspond to a job with a higher fraction of basic tasks, while a lower $F(\theta)$ would imply that the job is characterized by a prevalence of exceptional tasks. For a given distribution of tasks, a lower $F(\underline{\theta})$ corresponds to a higher ability gap between skilled and unskilled agents.\footnote{In the latter case, the lower $F(\underline{\theta})$ is only driven by a decrease in $\underline{\theta}$.}

Hence, a bonus contract is more likely to be adopted in complex jobs and when the skill...
gap among employees is higher, while a termination contract is more likely to be observed in ‘normal’ jobs and when the ability gap among employees is lower.

The choice of employment contract also has implications for the employer’s hiring and firing practices. In ‘normal’ jobs for which a termination contract is adopted, vacancies are filled more quickly, but the turnover is higher because employees are selected on-the-job. In contrast, in complex jobs for which a bonus contract is used, filling a vacancy takes longer, but the turnover rate is lower. MacLeod and Parent (2015) provide empirical support for the existence of a longer tenure in complex jobs for which bonus contracts are used.

In our model, the distinction between ‘normal’ and ‘complex’ jobs - and therefore the
optimal relational contract - can evolve with changes in technology and in the characteristics and supply of skills. For example, a new technology that decreases the fraction of tasks that less qualified employees are able to solve would affect the nature of the job, and the employer could move from termination to a bonus contract. Identifying such changes could allow us to empirically test some of our theoretical results.

When introducing the model, we made the assumption that all tasks are equally valuable. If we were to relax this assumption, depending on which ‘easy’ or ‘difficult’ tasks are more valuable to the firm, the choice of a termination contract will be more likely in the former case, and the choice of a bonus contract will be more likely in the latter.

Another possible extension of the model would be to consider that there is a continuum of employee types. This would make the analysis significantly more complex. In that case, the results of stationary contracts or full separation of types when firms offer bonus contracts would no longer hold. However, it is still true that a bonus contract can be credibly offered only to high-talent employees. Therefore, the main trade-offs driving an employer’s choice of one contract over the other would still depend on the distribution of talent and the nature of the job that the principal offers. For example, it should still be true that in jobs for which most tasks can be solved by low-talent employees, the cost of adopting a bonus contract is too high, and therefore a termination contract is the likely choice for an employer.
4 Market equilibrium

We have characterized the set of contracts that can be supported by self-enforcing agreements for given market conditions. In this section, we characterize the stationary labor market equilibrium.

Employees’ outside option. The inter-temporal expected utility of an unemployed agent is as follows: \(^{17}\)

\[
V^{out}(\theta) = k + \delta \lambda V_C(\theta) + \delta (1 - \lambda) V^{out}(\theta),
\]

where \(V_C(\theta)\) is the inter-temporal expected utility of an employee with talent \(\theta\) who is employed and offered a contract \(C\). In Section 7 of the Online Appendix, it is shown that in this setting, the equilibrium \(V^{out}\) is independent of \(\theta\) when the firms offer a bonus or a termination contract. The result that, in equilibrium, the outside option is independent of an employee’s type is straightforward in the cases where all firms offer a bonus contract or all firms offer a low-requirement termination contract. Thus, here, we only discuss the intuition for this result when all firms offer a high-requirement termination contract. In that case, even though low-talent employees suffer from endogenous separation (unlike high-talent employees), they are also less likely (relative to a high-talent employee) to exert high effort. In equilibrium, these two effects exactly cancel each other out, and therefore, both types have the same inter-temporal expected utility as employed or unemployed.

The re-employment probability \(\lambda\) guarantees the equilibrium of flows between the firm and the labor market. The steady-state flow into the unemployment pool is \(n \times X\), where

\(^{17}\)Recall that an unemployed agent stays in the unemployment pool for at least one period.
$X$ is the fraction of jobs that become vacant during the period. This fraction of jobs can be decomposed as follows: $X = (X_R + X_S)$. $X_R$ is the fraction of agents who enter the unemployment pool at the beginning of the period because they have rejected a contract offered by an employer; and $X_S$ is the fraction of agents who enter the pool at the end of the period due to exogenous or endogenous separations. These fractions depend on the type of contracts adopted by the firms and are formally defined in Table 1 below. Because the flow out of the unemployment pool is $\lambda(1 - n + n \times X_R)$, the re-employment probability is\(^{18}\)

\[
\lambda = \frac{n \times X}{1 - n + n \times X_R}.
\]

**Talent distribution.** The fraction of high-talent employees in the overall population is $\overline{s}$. Let $\overline{q}$ be the fraction of firms employing an agent with talent $\overline{\theta}$. Let $\overline{g}$ be the fraction of agents with talent $\overline{\theta}$ in the pool of unemployed at the beginning of period $T$ after vacancies have been filled but before the employees who have rejected offers have returned to the pool of unemployed. In equilibrium, the following condition needs to hold.

\[
n \times \overline{q} + (1 - n) \times \overline{g} = \overline{s}
\]

Once some agents decide to reject the contract they have been offered, they return to the unemployment pool and are available to be re-employed at the beginning of $T + 1$.\(^{19}\)

\(^{18}\)We implicitly assume that the values of the parameters are such that $\lambda < 1$; that is, all firms with vacant positions are matched with a new employee. We take this approach in order to concentrate on the case where jobs are always in short supply, and it does not affect our qualitative results. If there is a shortage of agents in the unemployment pool, then a principal is less likely to renege on the payment of the bonus because it is more likely that if he does so, the position may remain unfilled. The effect of an employee shortage on the choice of relational contract in the context of homogeneous employees is discussed in MacLeod and Malcomson (1998).

\(^{19}\)Recall that in equilibrium, only low-talent employees may reject a contract.
The probability of finding an agent with talent \( \tilde{\theta} \) among those available for employment at the beginning of \( T + 1 \) is denoted by \( \bar{p} \) and is given by:

\[
\bar{p} = \frac{\bar{y}(1 - n)}{1 - n + n \times X_R}.
\]

We can now define the fractions of agents who leave at the beginning or at the end of a period conditionally on the contract offered by the employers.\(^{20}\)

<table>
<thead>
<tr>
<th>Contract Type</th>
<th>( X_R )</th>
<th>( X_S )</th>
</tr>
</thead>
<tbody>
<tr>
<td>bonus contract</td>
<td>((1 - \bar{p})X)</td>
<td>((1 - \alpha)(1 - X(1 - \bar{p})))</td>
</tr>
<tr>
<td>termination contract (( w(y), y ), where ( y \in {\theta, \tilde{\theta}} ))</td>
<td>0</td>
<td>((1 - \alpha) + (1 - \bar{y})\alpha(F(y) - F(\theta)))</td>
</tr>
</tbody>
</table>

Table 1: The fractions of employees returning to the unemployment pool during a given period of time conditional on the contract adopted by the mass of employers.

In a stationary equilibrium, the fraction of \( \tilde{\theta} \) agents in the pool of employed workers should be the same at the beginning and at the end of a period. For this to be the case, the fraction of high-talent agents who leave the firm for exogenous reasons \((1 - \alpha)\bar{q}\), should be offset by the fraction of high-talent agents among the new hires, \( \bar{p}X \), that is,

\[
(10) \quad (1 - \alpha)\bar{q} = \bar{p}X.
\]

Our aim is to derive the conditions under which different employment systems, namely termination contracts and bonus contracts, co-exist in the market equilibrium. Thus, let \( \beta \) be the fraction of firms that offer a bonus contract in equilibrium and \((1 - \beta)\) be the

\(^{20}\)A detailed explanation of Table 1 is provided in Section 8 of the Online Appendix.
fraction of firms that offer a termination contract \((w(y), y)\), where \(y \in \{\underline{\theta}, \bar{\theta}\}\).\(^{21}\)

**Lemma 4.**

i) The equilibrium probability for an employer with a vacancy to be matched with a high-talent employee decreases with the fraction of firms that offer a bonus contract, i.e. \(\frac{\partial p}{\partial \beta} < 0\).

ii) For a given \(\beta > 0\), the equilibrium probability for an employer with a vacancy to be matched with a high-talent employee decreases with the performance requirement of the termination contract, i.e. \(\frac{\partial p}{\partial y} < 0\).

When matched with employers offering a bonus contract, agents with low talent reject that contract and immediately return to the pool of unemployed, while when matched with a firm offering a termination contract, low-talent employees remain employed until the employment relationship is interrupted for endogenous or exogenous reasons. Therefore, when more firms offer a bonus contract, the agents who are employed are more thoroughly selected, which in turn implies that it becomes less likely to find a high-talent agent in the pool of unemployed. In what follows, \(\overline{p}(\beta)\) denotes the equilibrium probability for an employer with a vacancy to be matched with a high-talent employee as a function of \(\beta\).

The higher the performance requirement is, the more likely it will be for a firm offering a termination contract to identify and fire a low-talent employee. Hence, a termination contract with \(y = \bar{\theta}\) allows the firms offering that contract to select employees on the job, increases the fraction of high-talent agents among the employed and therefore decreases the probability of finding a high-talent agent in the pool of unemployed.

\(^{21}\)\(\beta = 1\) corresponds to the case where all firms offer a bonus contract and \(\beta = 0\) to the case where all firms offer a termination contract.
An implication of Lemma 4 is that the probability of finding a high-talent employee in the pool of unemployed is the lowest when all firms offer a bonus contract. Conversely, this probability is the highest when all firms offer a low-requirement termination contract.

Before characterizing the market equilibrium, it is important to emphasize that the value of the employees’ outside option also depends on the contracts offered by the firms. If more firms offer a termination contract, an unemployed agent is more likely to be matched with such a firm and earn a rent, which in turn implies that his outside option is higher. Therefore, the thresholds derived in Proposition 4 also depend on the fraction of firms adopting a bonus contract, we adopt the following notations $p_{SE}(\beta), p'(\beta)$ and $\hat{p}(\beta)$ to reflect that these thresholds depend on $\beta$.

**Proposition 5.**

i) If $p_{SE}(1) > p(1) > \max\{p'(1), \hat{p}(1)\}$, then in equilibrium all firms offer a bonus contract.

ii) If $p(0) < \max\{p'(0), \hat{p}(0)\}$, then all firms offer a termination contract.

iii) If $p(0) > \max\{p'(0), \hat{p}(0)\}$ and $p(1) < \max\{p'(1), \hat{p}(1)\}$, then in equilibrium, bonus and termination contracts co-exist, and the fraction of firms offering a bonus contract is uniquely determined by $p(\beta) = \max\{p'(\beta), \hat{p}(\beta)\}$.

As more firms adopt a bonus contract, the fraction of high-talent employees in the pool of unemployed workers decreases. Furthermore, $\beta$ also affects the threshold $p'(\beta)$ (resp. $\hat{p}(\beta)$) that determines the choice between a bonus contract and a high-requirement (resp. low-requirement) termination contract. Indeed, as more firms adopt a bonus contract (i.e. higher $\beta$), the employee’s outside option decreases and the threshold above which
an individual firm is willing to adopt the bonus contract is higher. Both effects make the termination contract more appealing for an individual employer. Hence, the main result that we want to emphasize is that for some values of the parameters, a priori identical companies adopt different employment systems in equilibrium. While some firms adopt strong *ex ante* selection standards, offer bonus payments and have low turnover, other firms are less selective, offer fixed wages and experience higher turnover. Note that the fixed wage offered as part of a bonus contract increases as the fraction of firms offering termination contracts increases.

The *ex ante* inter-temporal expected profit of the employers is the same regardless of the contract that is adopted. However, if the position is filled, firms offering a bonus contract enjoy a higher inter-temporal profit than firms offering a termination contract.

Finally, we also note that when bonus and low-requirement termination contracts co-exist in equilibrium, firms do not solve the same sets of tasks. Indeed, the firms that choose a bonus contract can offer solutions to a broader set of tasks compared to firms offering a termination contract \((w(\theta), y = \theta)\).

Osterman (1994), Hunter (1999), and Bailey and Sandy (1999) investigate the adoption of high-performance work systems in different US industries. All these studies show that the adoption of high-performance work systems is not generalized and that different employment systems indeed actually do co-exist. For example, Hunter (1999) studies the US retail banking industry and shows that two employment systems co-exist: inclusive and segmented. The inclusive system is characterized by better employee selection, by a broader set of tasks to be addressed, and by being more likely to be involved in team-work with high-powered incentives. The segmented system is characterized by a large set of

\footnote{Note that when all firms offer termination contracts, it may be that all firms offer the same contract or that high and low-requirement termination contracts co-exist.}
employees who are specialized in handling a narrow set of tasks, low-powered incentives with rare possibilities for promotion, and high turnover.

One of the justifications for the co-existence of the employment systems examined in the papers discussed above is that firms have different product strategies. If there were heterogeneity among firms in our paper in terms of the distribution of tasks that their employees must perform, then this heterogeneity would certainly drive the adoption of different employment systems. However, setting aside this level of heterogeneity allows us to emphasize the importance of employee characteristics for the adoption of a specific employment system, or for the co-existence of different types of jobs.

5 Conclusion

Informal reputational mechanisms play a crucial role in a large range of contractual relationships. The aim of this paper is to show how the characteristics of the job, the labor market, or the firm affect the compensation packages used by employers and their hiring and firing policies. We accomplished this in a setting with heterogeneous employees whose ability is persistent over time and is the employee’s private information. We showed that in ‘bonus jobs’, selection occurs ex ante, and hence a vacant position may remain unoccupied for a longer period; however, once an employee has been hired, he will remain for a longer period in that position. In contrast, firms offering efficiency wage contracts fill their vacant positions more quickly. However, because employees are selected on-the-job, the tenure of the average employee will be shorter. We have also offered a theoretical justification for the adoption of different contracts and hiring/firing policies by identical firms operating in the same industry.
6 References


