Relational Contracts, Limited Liability, and Employment Dynamics*

Yuk-fai Fong
Kellogg School of Management
Northwestern University
y-fong@kellogg.northwestern.edu

Jin Li
Kellogg School of Management
Northwestern University
jin-li@kellogg.northwestern.edu

April 23, 2009

Abstract

This paper develops a tractable model of relational contract with imperfect public monitoring where the agent has limited liability. In the optimal relational contractort, both monetary reward for good performance and punishment for bad performance through termination are used. Both are postponed as much as possible yet termination always occurs with positive probability.

The optimal relational contract sheds light on a number of important patterns of employment dynamics. First, employment relationship sometimes starts with a probation phase, after which the agent either receives permanent employment or is terminated. Second, the sensitivity of wages to performance increases with experience and wages are backloaded. Third, turnover rates can be inverse-U shaped with seniority. Fourth, earlier successes are more important for future wage growth.

The tractability of the model also allows us to carry out several comparative statics that are typically difficult to obtain in discrete-time dynamic incentive models. Our technique of obtaining comparative statics uses properties of functional operators and may be of independent interest. Some of the comparative statics results shed new light on important policy issues. For example, minimum wage may harm workers who are already employed because they are more likely to be terminated.

JEL Classifications: C61, C73, J33, L24
Keywords: Relational Contract, Limited Liability, Efficiency Wage

*We thank Guy Aries for excellent research assistance. We especially thank Marina Halac for detailed comments. We also thank Ricardo Alonso, Attila Ambrus, Dan Bernhardt, Odilon Camara, George Deltas, Bob Gibbons, Michihiro Kandori, John Matsusaka, Kevin Murphy, Tony Marino, Marco Ottaviani, Peter Schnabl, and seminar participants at Northwestern, University of Delaware, UIUC, USC, and IIOC for helpful comments and discussions. All remaining errors are ours.

1 Introduction

There are three salient features of the labor market. First, employment relationships involve repeated interactions. Hall (1982) finds that the average tenure of U.S. workers is 8 years and a quarter of the workforce has jobs that last more than 20 years. Beaudry and DiNardo (1992) show that long-term contracts better describe the labor market than spot markets. Second, efforts of the workers can be hard to monitor, and it is often important to provide incentive to induce efforts from the workers. Moral hazard problems have been extensively studied in the literature; see for example Holmstrom and Hart (1987), Gibbons (1997), Gibbons and Waldman (1999) and Prendergast (1999) for reviews. Third, worker has limited liability: worker is liquidity-constrained and in general cannot pay the firm more than what he owns. There can be many forms of the limited liability constraint: firms cannot take money out of the worker's pocket; minimum wage laws can set a wage floor; institutional reasons can attach wages to jobs; psychological costs of the workers can prevent firms from setting (nominal) wages that are lower than previous years.[1]

These three features are the key elements underlying the famous efficiency-wage model of Shapiro and Stiglitz (1984). Moral hazard together with limited liability of the workers generates rents in jobs. Repeated interactions between firms and workers imply that future rents can be used as an incentive device to induce effort from the workers. In the Shapiro and Stiglitz (1984) model, wages paid to the workers are assumed to be constant over time. It has been recognized, however, that constant-wage is not the optimal: firms can improve their payoffs through alternative contractual arrangements; see for example Akerlof and Katz (1989).[2]

In this paper, we study the optimal contract for the firms in an infinitely repeated principal-agent model where the agent has limited liability. In particular, we focus on an environment in which a) output is stochastic in agent's effort and b) output is observable but not contractible. The first assumption implies that low output can occur even if the agent puts in effort. This feature allows us to study turnover dynamics that has been missing in existing efficiency wage models.[3]

The second assumption implies that there is no explicit contract in this model. Instead, we study relational contracts. Relational contracts are contracts enforced not by the rule of the court but by the concerns of the parties for their future interests. Each relational contract in this model

[1] For incentive models with limited liability; see for example Sappington (1983); Innes (1990), and Jewitt, Kadan, Swinkels (forthcoming).

[2] Carmichael (1985) shows that if the worker does not have limited liability and the firm can be trusted to act honestly, then the optimal wage contract is to have the worker post a bond and forfeit it if caught shirking. When there is no limited liability and the firm cannot be trusted, then Levin (2003) implies that the optimal contract can be implemented by a sequence of stationary contracts.

[3] If the output is the principal's private information, Fuchs (2007) shows that turnover will occur. However, the equilibrium contract in that case is sufficiently complicated that little is known about the turnover dynamics.

corresponds to a Public Perfect Equilibrium (PPE), which is the standard solution concept in this setting. The existing literature on efficiency wages in general assume that firms will behave honestly (be able to commit to a contract) and justify it by invoking the argument that they care about the future. Our setup illustrates the conditions under which this argument is valid. We also characterize the optimal contract when firms can commit and study its implications on wage dynamics in Section 5.

The optimal relational contract sheds light on a number of important patterns of employment dynamics. First and foremost, the beginning of the employment relationship is characterized by a "probation phase" during which the agent receives constant wage. Depending on the output sequence, the worker is either terminated or receives permanent employment. Probation is an important feature for many professions, and its existence is typically attributed to worker selection; see for example Bull and Tedeschi (1989), Sadanand et al. (1989), Weiss and Wang (1990), and Wang and Weiss (1998). Here, we show that the probation phase can also serve as an incentive device.

Second, the sensitivity of wages to performance increases with experience and wages are backloaded. It is well-known that backloaded wage payment can induce effort from the workers; see for example Becker and Stigler (1974), Lazear (1981), and Akerlof and Katz (1989). This model extends previous models to stochastic production, and it predicts that not only wage increases with time, but also that wage becomes more responsive to performance over time. This prediction appears to fit the incentive structure of some occupations; see for example Lin (2005) and Coughlan et al (2005).

Third, this model generates a variety of turnover patterns. In particular, turnover rates can be inverse-U shaped with years on the job. This is the celebrated prediction of Jovanovic (1979) model of learning and matching, which has been a workhorse model guiding empirical research; see for example Farber (1999). Moreover, when the probability of high output is small, the optimal contract implies that there will be a fixed date such that the worker will be terminated if he does not receive permanent employment prior to it. Such up-or-out turnover pattern is a distinctive feature in professional service firms.

Finally, there will be path-dependency in both wage and turnover dynamics: earlier successes are more important for both job security and future wage growth. This prediction on wage dynamics is related with the idea of "fast track" in the internal labor market literature, which states that workers who experience high wage growth in the past are more likely to have high rates of wage growth in the future; see for example Baker, Gibbs, Holmstrom (1994a, b) and Treble et al (2001).

The employment dynamics above are direct consequences of the optimal relational contract, which efficiently utilizes the rent in the job to induce effort from the worker. In particular, for a

worker to exert effort, he must be rewarded for high outputs. The principal can provide the agent the incentives to exert effort either by a carrot, i.e., bonus or a stick, i.e., positive probability of termination. Relying exclusively on the bonus can guarantee efficiency but when the agent has limited liability, doing so requires the principal to offer the agent a high continuation payoff right at the beginning of the game. The high payoff to the agent is partly reflected by the fact that even if he shirks forever, he will never be terminated. Alternatively, the principal can take away some of the agent's continuation payoff by terminating him after a long streak of failure and rewarding him with permanent employment only after multiple periods of success. Under the new scheme, the agent is motivated by the benefit of permanent employment and the fear of termination, so the principal does not have to pay any bonus until the agent earns his permanent employment status. Since termination can be pushed back to a distant future if the principal only takes away a small amount of the agent's payoff, it is optimal for the principal to at least allow for some possibility of (postponed) termination even when efficient production is sustainable as an equilibrium outcome.

One distinguishing feature of this model is its tractability. Models of dynamic moral hazard are typically complicated.[4] Here, we show that for a wide range of parameters, the Pareto frontier of the PPE payoff set is completely determined by a functional equation, and we develop an algorithm for finding it. The tractability of the model also allows us to carry out several comparative statics that are difficult to obtain in discrete-time dynamic incentive models. Our technique of obtaining comparative statics by using properties of functional operators may be of independent interest. Moreover, some of the comparative statics results shed new light on important policy issues. For example, while most of the debates on minimum wages focus on the employment margin, our analysis indicates that minimum wage may also harm workers who are already employed through an increase in the probability of involuntary turnovers in the future.

Moreover, we obtain closed-form solutions for the Pareto frontier of the relational contract payoffs for some parameter values. These close-form solutions are not only facilitates the calculation of the Pareto frontier, but also illustrate important theoretical properties. For example, we show that the Pareto frontier is in general not differentiable.[5] More interestingly, we provide an example where the number of non-differentiable points is infinite, and these points converge to a point which is differentiable.

In addition to the efficiency wage literature, this paper is related to three literatures. First, this paper belongs to a growing literature that focuses on dynamics of the relational contracts. In

[4] The analytical difficulty of dynamic moral hazard in the discrete-time setting has been an important motivation for continuous-time models; see for example DeMarzo and Sannikov (2006). In continuous-time models, however, additional assumptions (either difference in discount rates or bounds on the bonus) are needed to guarantee the existence of the optimal contract.

[5] Thomas and Worrall (1994) is the first to point out that the Pareto frontier may not be differentiable.

the classical models of relational contracts, such as Bull (1987), MacLeod and Malcomson (1988), Baker, Gibbons and Murphy (1994), and Levin (2003), the optimal relational contracts can be implemented by stationary contracts. Recently, a few interesting papers bring dynamics into relational contracts. The most related paper is Thomas and Worrall (2007). They investigate a dynamic relational contract between two partners under limited liability. Their model is more general by having a continuum of effort choices and outputs. Consequently, Thomas and Worrall (2007) generate interesting distortions on the effort side: in particular it is possible for efforts to be overprovided. The key difference between our model and that of Thomas and Worrall (2007) is that monitoring is imperfect in our model, and this allows us to analyze turnover dynamics while turnover does not occur in their setting.[6]

There are also several papers that generate interesting dynamics in relational contracts of by incorporating additional features of the labor market. Yang (2005) studies relational contracts where workers are heterogeneous and their types are private information. In equilibrium, low types leave the relationship over time and wage is increasing in tenure. Chassang (2008) studies relational contracts with experimentation. The agent can experiment on new technologies but is not always successful. His model helps explain why there is significant amount of dispersion in productivity in otherwise similar firms/plants. Fuchs (2007) studies relational contracts in which the output is the principal's private information. He shows that the optimal contract can be implemented by termination contracts. Halac (2008) studies relational contracts where the principal's outside option is his private information. In equilibrium, principals with high outside option may renege on bonus payments and terminate the relationship.

Second, this paper is related to the vibrant literature of dynamic moral hazard with limited liability when the principal can fully commit. This literature takes the optimal contracting view and studies its implications in diverse economic situations. In finance, Biais et al (2007), DeMarzo and Fishman (2007), and DeMarzo and Sannikov (2006) study the implication of optimal contract on the use of debt, equity and credit line in standard financial contracts. In industrial organization, Clementi and Hopenhayn (2006) show that the optimal contracting view gives rise to realistic patterns firm dynamics. In the political economy context, Myerson (2008) examines how the optimal contract of a prince affects the career of the governors serving him. In the search context, Lewis (2009) and Lewis and Ottaviani (2008) characterize the optimal dynamic contract when the moral hazard problem is related to the agent's search intensity.

This paper adds to this literature by studying the implication of optimal contracts on employment dynamics. Different from these papers, we consider both full-commitment and relational contracts, and we are able to explore how the lack of commitment affects the optimal contract.

[6] Another difference is that Thomas and Worrall (2007) is a model of partners and the moral hazard is two-sided. Here, we study an employment relationship and there is one-sided moral hazard.

We show that when the surplus in the relationship is large, lack of commitment imposes no cost on the principal and does not affect the employment dynamics. On the other hand, when the surplus of the relationship is small, the principal is hurt by the lack of commitment and there will be significant differences in some aspects of the employment dynamics. For example, the agent receives permanent employment with positive probability when firms can commit, and with relational contract, the relationship terminates with probability 1. We explore these differences in Section 5.3.

Finally, this paper contributes to the literature of formal models that explain multiple employment patterns in labor markets; see for example Harris and Holmstrom (1982); Bernhardt (1995); and Gibbons and Waldman (1998). The existing models assume that workers differ by their types, which can be learned by the employer over time. In other words, worker heterogeneity is important for generating multiple employment patterns in these models. Our model complement the models above by showing that even with homogeneous workers, it is possible to simultaneously explain features of employment dynamics, including probation phase, deferred compensation, tenure, and fast track through concerns for dynamic incentives.

The rest of the paper is organized as follows. We set up the model in Section 2. We analyze the model in Section 3. Properties of the optimal relational contract are derived in Section 4. Section 5 analyzes the model under alternative assumptions. Section 6 concludes.

2 Setup

Time is discrete and indexed by $t \in \{1, 2, ..., \infty\}$.

2.1 Players

There is one principal and one agent. Both are risk neutral, infinitely lived, and have a common discount factor of δ. The agent's per period outside option is $\underline{u}$; the principal's per period outside option is $\underline{v}$.

2.2 Production

If the principal and the agent engages in production together, the agent will be asked to perform a task. The agent can choose either to work or shirk. If the agent works, the outcome of the task Y is y with probability $p \in (0, 1)$ and 0 with probability $1 - p$. If the agent shirks, the outcome is y with probability $q < p$. If he chooses to work, the agent incurs an effort cost of c.

We assume that it is efficient for the agent to work. Moreover, the value of the relationship is less than the outside options if the outside options if the agent choose to shirk.

$$py - c > \underline{u} + \underline{v} \geq \underline{v} > qy.$$

These are standard assumptions in the literature. The production technology here is a special case of Levin (2003) and is identical to Fuchs (2007).

2.3 Time line and Information

We follow the timing in Levin (2003) with one change that simplifies the exposition. At the beginning of each period t, the principal decides whether to offer a contract to the agent, $d_t^P \in \{0, 1\}$. If the principal chooses not to offer the contract ($d_t^P = 0$), then the two parties receive their outside options in this period. If the contract is offered, it specifies a wage w_t, which can be legally enforced. We assume that

$$w_t \geq \underline{w},$$

where $\underline{w}$ is an exogenously given wage floor.

The agent chooses $d_t^A \in \{0, 1\}$, and if he rejects the contract ($d_t^A = 0$), the two parties receive their outside options. Otherwise, the relationship starts. The principal pays out the wage. The agent chooses effort $e_t \in \{0, 1\}$, and the output Y_t is realized. Finally, we assume that there is a public randomization device, so that a random variable $x_t \in [0, 1]$ will be drawn at the end of the period. This is a standard assumption in the literature made to ensure that the set of perfect public equilibrium payoffs are convex; see for example Fuchs (2007).

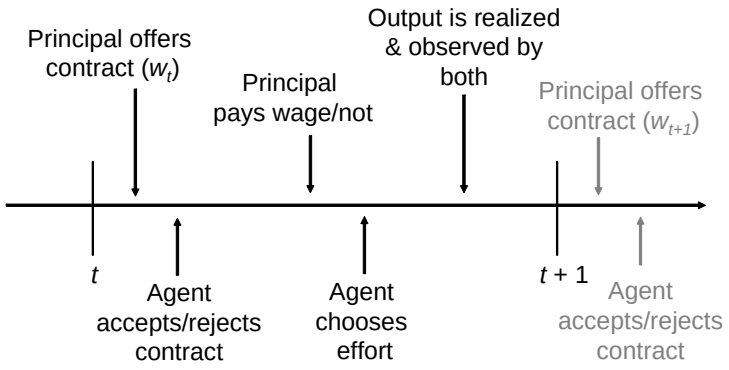

Figure 1: Timeline

We assume that the effort is the agent's private information. The output is publicly observed by the principal and the agent, but it cannot be contracted upon. The non-contractbility of the output makes the contract relational. To induce effort from the agent, it is often modelled that there is a (non-contractible) performance bonus paid out at the end of a period.[7] This modelling choice is very helpful in illustrating the connection between explicit and relational contracts; see for example Levin (2003).

In our model, we do not have a performance bonus at the end of period. Instead, the "performance bonus" is paid out at the beginning of the following period. In other words, the agents are incentivized through "efficiency wage" as opposed to "bonus pay". Since all the outputs and bonuses are publicly observed, it is well-known in the literature (see for example Macleod and Malcomson (1988)) that these two setups give rise to the same equilibrium payoff set.[8] We choose our setup to to highlight the effect of limited liability in the efficiency wage literature.

Finally, We refer to the constraint $w_t \geq \underline{w}$ as the limited liability constraint, and there are several ways in interpreting it. When $\underline{w} = 0$, this constraint can be literally thought of as the limited liability constraint so that the principal cannot take money out of the agent's pocket. When $\underline{w}$ is equal to the minimum wage, this constraint specify that the wage offered has to exceed the minimum wage. In cases where wages are often tied to jobs, this constraint specifies that there are lower bounds of wages associated with jobs.[9] The limited liability constraint, in the sense that there is a limit to how much one can punish the agent for bad performance, is inherent in many important efficiency wage models of labor market; see for example Shapiro and Stiglitz (1984) and Akerlof and Katz (1989).

2.4 Strategy and Equilibrium Concept

2.4.1 History

We denote $h_t = \{d_t^P, w_t, d_t^A, y_t, x_t\}$ as the public events that happen in period t. Denote $h^t = \{h_n\}_{n=0}^{t-1}$ as a public history path at the beginning of period t, and $h^1 = \varnothing$. Let $H^t = \{h^t\}$ be the set of public history paths till time t, and define $H = \cup_t H^t$ as the set of public histories.

2.4.2 Strategy

We restrict ourselves to public strategies. This means that the actions of a player will be contingent only on events that are publicly observable. In particular, in period t, the action of

[7] It is possible that the bonus is negative in those models, and in such case, the agent has a discretion to make payments to the principal.

[8] MacLeod and and Malcomson (1998) shows that efficiency wage is more likely to arise than pay-for-performance when there are fewer workers than firms.

[9] Since there is only one job in this model, this constraint should be interpreted as the lower bound for this particular job. Prendergast (1993) gives one reason why firms want to tie wages to jobs.

the principal is to choose D_t^P from H^t to $\{0,1\}$ and W_t from H^t to $[\underline{w},\infty)$. The public strategy of the principal is $\{D_t^P, W_t\}_{t=1}^{\infty}$.

In period t, the agent chooses D_t^A from $H^t \cup \{w_t\}$ to $\{0,1\}$ and E_t from $H^t \cup \{w_t\}$ to $\{0,1\}$. And the strategy of the agent is $\{D_t^A, E_t\}_{t=1}^{\infty}$.

We allow the players to play mixed strategies in this game. Denote σ^A to be the mixed public strategy of the agent and σ^P to be the mixed public strategy of the principal. They can be defined in the standard way, so we omit their precise definitions here.

2.4.3 Public Perfect Equilibrium

We analyze the public perfect equilibrium (PPE) of the game. A strategy profile is a PPE if a) the players use public strategies, and b) following every public history, the strategy is a Nash equilibrium. PPE is the standard equilibrium concept in repeated games of imperfect public monitoring. Moreover, this model is a game of imperfect monitoring with "product structure", in the sense that the output depends on the agent's effort alone. It follows that our restriction to PPEs is without loss of generality; see Fudenberg and Levine (1994).

To check whether a strategy profile is PPE in this model, it suffices to check one-stage deviation. In particular, a mixed strategy profile σ is a PPE if and only if

- following any history h^t, any (d_t^P, w_t) in the support of σ^P is the best response for the principal, holding all of the rest of the strategy fixed.
- following any history $h^t \cup \{d_t^P, w_t\}$, (d_t^A, e_t) in the support of σ^A is the best response for the agent, holding all of the rest of the strategy fixed.

3 Analysis

We solve the model in this section. In Subsection 3.1, we show that the PPE payoff set is completely determined by its Pareto frontier. In Subsection 3.2, we characterize the Pareto frontier. In Subsection 3.3, we show that the optimal relational contract is inefficient. In Subsection 3.4, we provide closed-form expressions for the Pareto frontier for a range of parameters.

3.1 Reduction to Pareto Frontier

Abreu, Pierce, Stacchetti (1990) (APS hereafter) develop a powerful technique to characterize the PPE set. The basic idea of APS is that instead of focusing on the set of strategies, more information can be obtained by looking at PPE payoff sets. Under the assumptions that there are finite number of actions for the players and that there is a public randomization device, APS

show that the PPE payoff set E is convex and compact. Moreover, APS develop an algorithm in finding the PPE payoff set E. Since E is a multidimensional set, its characterization is in general difficult.

In this model, two features of the relational contract setup help simplify the characterization of the PPE payoff set. First, both the principal and the agent can choose to take their outside option, so any PPE payoff must give the principal at least $\underline{v}$ and the agent at least $\underline{u}$. Moreover, taking their outside options $(\underline{u}, \underline{v})$ is a PPE payoff, supported by the belief that the principal will not offer contract to the agent in the future and the agent will always shirk.

Second, the principal can make transfers to the agent. This implies that, if (u, v) is a PPE payoff, then by asking the principal to transfer extra money to the agent at the beginning of period 1 and keep the rest of the strategies fixed can produce another PPE payoff (u', v') with $u' + v' = u + v$, $u' > u$, and $v' \geq \underline{v}$. In particular, $(u + v - \underline{v}, \underline{v})$ is a PPE payoff.

These two features imply that any payoff that a) gives the agent at least $\underline{u}$, b) gives the principal at least $\underline{v}$, and c) lies below the Pareto frontier of the PPE payoff set, is again a PPE payoff. In other words, the PPE payoff set is completely characterized by its Pareto frontier.

We denote $f(u)$ as the maximum PPE payoff of the principal if the agent's payoff is u. From APS, we know that f is well-defined because the PPE set is compact. The lemma below states the result above formally.

Lemma 0: Let $u_{\max}$ be the maximum PPE payoff of the agent. Then the PPE payoff set E is given by

$$E = \{(u, v) : u \in [\underline{u}, u_{\max}], v \in [\underline{v}, f(u)]\}$$

Proof. First, note that the payoff pair $(\underline{u}, \underline{v})$ (meaning that the agent's normalized expected payoff is $\underline{u}$ and the principal's normalized payoff is $\underline{v}$) is in the PPE payoff set. This payoff is supported by an equilibrium in which on the equilibrium path, the principal and the agent does not start a relationship, and off the equilibrium path, the agent never puts in effort and both the principal and the agent do the start the relationship in the future.

Second, it follows by convexity of the PPE payoff set that any payoff on the line segment between $(\underline{u}, \underline{v})$ and $(\underline{u}, f(\underline{u}))$ can be supported as a PPE payoff, and this is the left boundary of the PPE payoff set. Third, because there is no limit in the amount of transfer the principal can make to the agent at the beginning of period 1, it is easily seen that the lower boundary of the PPE set is given by the horizontal line at $\underline{v}$. Finally, convexity implies that any equilibrium payoff between $(u, \underline{v})$ and $(u, f(u))$ can be obtained. ■

3.2 Characterizing the Pareto Frontier

In this subsection, we characterize the Pareto frontier f. We show that the Pareto frontier can be classified into three regions. To the left of a low threshold, the relationship terminates immediately with positive probability, and the Pareto frontier is a straight line with a positive slope that starts out at $(\underline{u}, \underline{v})$. To the right of a high threshold, the Pareto frontier is a straight line with a slope of -1. In this region,the relational contract attains the unconstrained Pareto efficiency and termination never occurs. Between the two thresholds, the Pareto frontier satisfies a functional equation. There is a unique solution to the functional equation, and we show that it can be found using a two-step procedure where each step involves solving a fixed point of a contraction mapping.

In Subsection 3.2.1, we list the necessary incentive constraints to induce effort. In Subsection 3.2.2, we characterize and compare the Pareto frontiers of the PPE payoffs both with and without limited liability. In Subsection 3.2.3, we provide an algorithm of finding the Pareto frontier. In Subsection 3.2.4, we show that the optimal relational contract is inefficient. With the exception of Proposition 1, all of the proofs are collected in the Appendix.

In this section, we analyze the model by assuming that a) the surplus in the relationship is high and b) wage floor is high. These two assumptions makes the analysis particularly tractable and help us highlight the difference of relational contract with and without limited liability. We analyze the model under alternative assumptions in Section 5. The main features of the Pareto frontier remain the same.

Assumption 1:

$$py - \frac{[1-\delta(1-p)]c}{\delta(p-q)} \geq \underline{v} + \underline{w}.$$

This assumption guarantees that there will be a stationary PPE in which the worker puts in effort every period and the relationship is efficient outcome. The main purpose for this assumption is expositional. In particular, the existence of this equilibrium helps us compare how the Pareto frontier of the PPE set with limited liability differs from one that does not have limited liability.

Assumption 2:

$$\underline{w} \geq \underline{u}.$$

This assumption states that the wage floor is higher than the outside option of the agent. A common special case of this assumption is one in which both the wage floor and the outside option of the agent are normalized to zero. Since $\underline{w}$ is the lowest possible wage the agent can receive, this assumption implies that there will be substantial amount of rent in the relationship for the agent.

This assumption therefore perhaps better describe jobs that attract many applicants and are hard to get. In addition, this assumption is more likely to be satisfied if the matching process of the labor market is inefficient. In this case, Assumption 2 means that it will take a worker a substantial amount of time to find a new employer when the previous employment relationship ends. For example, many people queue up for minimum-wage jobs during recession times.

There are also many economic environments in which this assumption does not fit. These cases are technically more challenging and are analyzed in Section 5. In this section, we use Assumption 2 because it leads to a simple characterization of the optimal relational contract, which allows us to highlight the key mechanism of how to induce incentive in a dynamic setting when rent is present in a job: termination should be used in the earlier stage of the relationship and bonus payment should be delayed as much as possible. This mechanism remains valid when the wage floor is lower than is assumed in this section.

3.2.1 Incentive Constraints

In the efficient outcome of the game, the agent chooses to work in each period. For the agent to have incentive to work in any period, we must have

$$(1-\delta)(w-c)+\delta((1-p)u_L+pu_H)\geq(1-\delta)w+\delta((1-q)u_L+qu_H),$$

where u_H corresponds to the agent's continuation payoff after a good outcome ($Y=y$) and u_L corresponds to the agent's continuation payoff after a bad outcome ($Y=0$). We can rewrite expression above can be rewritten as

$$u_H-u_L\geq\frac{(1-\delta)c}{\delta(p-q)}\equiv k.$$

Here, the k is the minimum reward for high output to sustain effort. In other words, to induce the agent to put in effort, the continuation payoffs following good and bad outcomes should differ by at least k.

In this model, the principal's IC for willing to stay in the relationship at the beginning of period t is simply

$$v_t\geq\underline{v},$$

where v_t is the principal's expected payoff at the beginning of period t (following some public history h^t).

Note that in models where the bonus is paid out at the end of a period, the IC constraint of

the principal (often called the non-reneging constraint) is given by

$$(1-\delta)(-b_{t-1}) + \delta v_t' \geq \delta \underline{v},$$

where b_{t-1} is the bonus paid to the agent in period $t-1$ and v_t' is the principal's payoff at the beginning of period t after the bonus is paid out. Note that these two IC constraints are identical. In our model, we have

$$v_t = v_t' - \frac{(1-\delta)}{\delta} b_{t-1},$$

accounting for the fact that the bonus is paid out at the beginning of period t.

3.2.2 Comparing PPE with and without Limited Liability

In this subsection, we investigate the shape of the Pareto frontier with limited liability. First, we show that the Pareto frontier of the relational contract with free transfers (Levin (2003)) is a negative 45 degree line. Then we show that the limited liability constraint truncates the negative 45 degree line and turn it into three regions.

When the transfer is free, we can use the transfer at the beginning of the first period to reallocate the payoffs between the principal and agent (and keep the rest of the strategies fixed), as long as the resulting payoffs are both greater than or equal to their outside options. More formally, if (u, v) is a PPE payoff, then all the payoffs in the set

$$\{(u', v') | u' + v' = u + v,\ u' \geq \underline{u},\ v' \geq \underline{v}\}$$

are again PPE payoffs.

This implies that every PPE payoff generates a -45 degree line segment such that all points on the line segment are again PPE payoffs. In particular, if we take the PPE that maximizes the joint surplus,[10] then every point on the -45 degree line segment generated by this PPE are again PPE payoffs. Moreover, these points are on the Pareto frontier because by definition they maximize the joint surplus and any payoff to the upper-right of the line segment gives higher joint surplus and thus cannot be a PPE payoff.

Assumption 1 implies that the efficient outcome can be sustained as a PPE. Consequently, the Pareto frontier in the case of free transfer is given by

$$\{(u, v) | u + v = py - c,\ u \geq \underline{u},\ v \geq \underline{v}\}$$

[10] There can be many PPE payoffs that maximize the joint surplus. The compactness of the PPE payoff set guarantees that we have at least one.

It follows that the PPE payoff vector that gives the principal the highest payoff is $(\underline{u}, py - c - \underline{u})$. Levin (2003) shows that the optimal relational contract can be sustained by stationary contracts. In the case of no limited liability, payoff vector can be supported by the following. The principal pays the agent a base wage w_s, and in case of a good outcome, a bonus of k will be paid out at the beginning of next period. In other words, the next period wage of the agent is w_s in case of a bad outcome and $w_s + k$ in case of a good outcome. The bonus induces the agent to put in effort in each period. It can be checked that when the base wage satisfies

$$w_s = \underline{u} + c - \frac{\delta p k}{1 - \delta} = \underline{u} - \frac{q}{p - q} c,$$

the agent's expected payoff is $\underline{u}$, and this leads to a PPE payoff of $(\underline{u}, py - c - \underline{u})$. Figure 2 below illustrates the Pareto frontier of PPE payoff with free transfers.

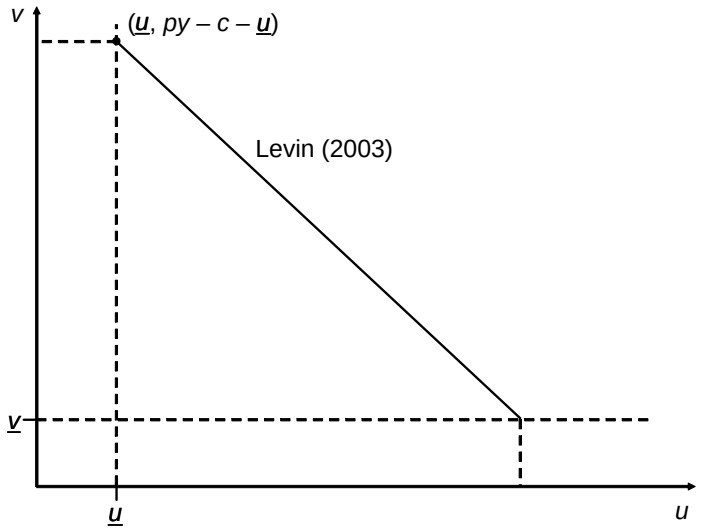

Figure 2: Pareto Frontier, w/o Limited Liability

Note that

$$w_s < \underline{u} \leq \underline{w},$$

so the above PPE with free transfer violates the limited liability constraint in Assumption 2. In this case, the Pareto frontier is no longer a -45 degree line. Instead, it has three regions, divided by the two thresholds u_0 and u_e. We provide below the basic intuitions of why the Pareto frontier taking this form, a rigorous derivation can be found in the appendix.

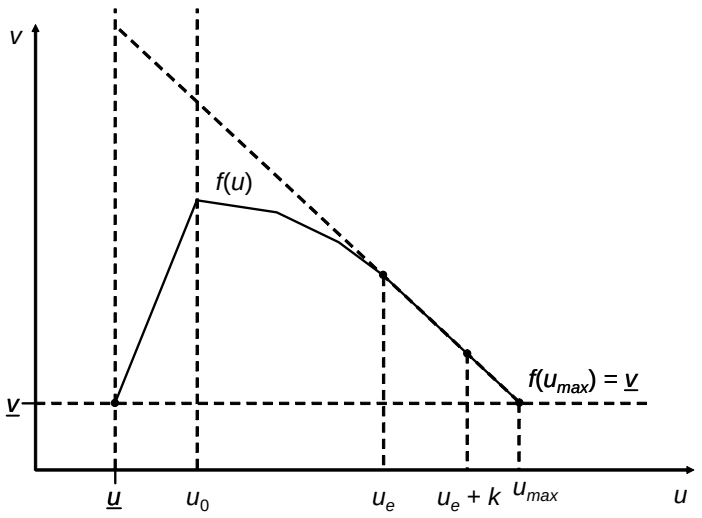

Figure 3: Possible Pareto Frontier, w/ Limited Liability

First, to the right of a threshold (u_e) the Pareto frontier is again a -45 degree line. Here, u_e is the unique payoff of the agent such that

$$u_e = (1-\delta)(\underline{w} - c) + \delta(u_e + pk). \tag{1}$$

In particular, the agent receives a payoff of u_e when he is paid a base wage $\underline{w}$, puts in effort, and receives a bonus k (to be paid out at the beginning of the next period) if the outcome is good. It can be shown that this can be supported as an equilibrium that results a payoff of u_e for the agent and $py - c - u_e$ for the principal. Moreover, this equilibrium maximizes the joint surplus of the two parties, so the Pareto frontier is a -45 degree line for $u > u_e$. And this can be achieved by making additional first period transfer *from the principal to the agent* and keeping the remaining strategies intact.

At u_e, the base wage of the agent is $\underline{w}$, so the limited liability constraint binds. Therefore, for $u < u_e$, the Pareto frontier can no longer lie on the -45 degree line because this would require the agent receive wage lower than $\underline{w}$ and thus violates the limited liability constraint.

To the left of u_e, the Pareto frontier can be classified into two regions. In particular, define u_0 such that

$$u_0 = (1-\delta)(\underline{w} - c) + \delta(\underline{u} + pk). \tag{2}$$

The expression states that if the agent receives a base wage of $\underline{w}$, puts in effort, and receives a continuation payoff of $\underline{u}$ for a bad outcome and $\underline{u} + k$ for a good outcome, then the agent's current expected payoff is u_0. Since the agent cannot receive a payoff less than $\underline{u}$, this implies that u_0 is the lowest expected payoff required for the agent to exert effort.

Therefore, for $u < u_0$, pure strategies cannot induce the agent to exert effort. The Pareto frontier for $u < u_0$ is obtained by a randomization between $(\underline{u}, \underline{v})$ and $(u_0, f(u_0))$, where recall that $f(u_0)$ is the highest PPE payoff for the principal when the agent receives u_0. The randomization implies that the Pareto frontier is a straight line. In other words, for $u < u_e$, there is some positive probability that the relationship will be terminated. And the termination probability corresponds to the weight of $(\underline{u}, \underline{v})$ in the randomization.

For $u \in [u_0, u_e]$, we note that at the Pareto frontier $((u, f(u)))$, the wage of the agent in the first period must equal to $\underline{w}$. Otherwise, the principal can lower his first period transfer by some small amount, so the slope of the Pareto frontier (f) at u is at most -1. This cannot be possible because a) the PPE payoff set is convex so f is concave and b) the slope of f is strictly less than -1 for $u < u_e$.

Since the agent is paid wage $\underline{w}$ in period 1, we introduce the following linear operator. Define L as the unique linear operator such that

$$u = (1-\delta)(\underline{w} - c) + \delta[(1-p)L(u) + p(L(u)+k)]. \tag{3}$$

In other words, $L(u)$ corresponds the to the agent's continuation payoff next period if he is paid $\underline{w}$ this period, puts in effort, but the outcome is $Y = 0$.

When the agent's continuation payoff is $L(u)$, a payoff pair $(L(u), v)$ can be supported as a PPE payoff as long as the principal's payoff $v \in [\underline{v}, f(L(u))]$. Similarly, any payoff pair $(L(u)+k, v')$ can be supported as a PPE payoff as long as $v' \in [\underline{v}, f(L(u)+k)]$. Now if $((u, f(u)))$ is on the Pareto frontier, this implies that the continuation payoffs must lie on the Pareto frontier as well. In other words, the continuation payoffs corresponding to the good and bad outcomes must be $(L(u)+k, f(L(u)+k))$ and $(L(u), f(L(u)))$. This gives rise to the following functional equation that the Pareto frontier must satisfy:

$$f(u) = (1-\delta)(py - \underline{w}) + \delta[pf(L(u)+k) + (1-p)f(L(u))]. \tag{4}$$

To summarize, we see that the Pareto frontier of the PPE payoff satisfies

$$f(u) = \begin{cases} \underline{v} + \frac{u-\underline{u}}{u_0-\underline{u}}(f(u_0) - \underline{v}) & \text{if } u \in [\underline{u}, u_0] \\ (1-\delta)(py - \underline{w}) + \delta[pf(L(u)+k) + (1-p)f(L(u))] & \text{if } u \in [u_0, u_e] \\ f(u_e) + u_e - u & \text{if } u \in [u_e, u_e + f(u_e) - \underline{v}], \end{cases} \tag{5}$$

where $u_e = (\underline{w} - c) + \frac{\delta pk}{(1-\delta)}$, $f(u_e) = py - c - ((\underline{w} - c) + \frac{\delta pk}{(1-\delta)})$, *and* $u_0 = (1-\delta)(\underline{w} - c) + \delta(\underline{u} + pk)$.

In the next section, we show that there is a unique solution that satisfies the functional equation and how we can find the solution. In Subsection 3.3, we constructed an explicit

solution for some cases. Readers who are more interested in the applications of the model can directly jump to Section 4.

3.2.3 Algorithm for Finding the Pareto Frontier

In this subsection, we describe how one can find an f that satisfies (5). The key is to note that the right hand side of the middle equation in (5) can be thought of as a contraction mapping on f, and the middle equation essentially states that f is a fixed point to this contraction mapping. This contraction mapping, however, depends on the value of f to the left of u_0.

The key is to note that $f(u_0)$ is determined both by the top equation in (5) that governs the region to the left of u_0 and by the contraction mapping in the middle equation. To find the value of $f(u_0)$, we first take a guess that $f(u_0) = Z$, which will give value to $f(u)$ to the left of u_0, and this defines the contraction mapping in the middle region, which we call T_Z. We can find the unique fixed point to T_Z, and this will give a value of $f(u_0|Z)$. We compare this value to Z, and adjusts Z upwards if $f(u_0|Z)$ is bigger than Z and adjust Z downwards if $f(u_0|Z)$ is smaller. We finds a solution to (5) if $Z = f(u_0|Z)$.

Formally, we take the following two-step procedure. In the first step, we define an operator T_Z on the space of bounded functions on $[u_{0,}u_e]$ as:

$$T_Z g(u) = \begin{cases} (1-\delta)(py-\underline{w}) + \delta(pg(L(u)+k) + (1-p)g(L(u)) & u_0 \leq L(u) < u_e - k \\ (1-\delta)(py-\underline{w}) + \delta(pg(L(u)+k) + (1-p)(\underline{v} + \frac{L(u)-\underline{u}}{u_0-\underline{u}}(Z-\underline{v}))) & L(u) < \min\{u_e - k, u_0\} \\ (1-\delta)(py-\underline{w}) + \delta(pf(L(u)+k) + (1-p)g(L(u)) & L(u) > \max\{u_e - k, u_0\} \\ (1-\delta)(py-\underline{w}) + \delta(pf(L(u)+k) + (1-p)(\underline{v} + \frac{L(u)-\underline{u}}{u_0-\underline{u}}(Z-\underline{v}))) & u_0 > L(u) \geq u_e - k \end{cases}, \tag{6}$$

where note that when $L(u) + k \geq u_e$, $f(L(u)+k) = py - c - (L(u)+k)$.

Note that if $Z = f(u_0)$, then f restricted to $[u_0, u_e]$ is a fixed point of T_Z :

$$\begin{aligned} T_Z f(u) &= (1-\delta)(py-\underline{w}) + \delta(pf(L(u)+k) + (1-p)f(L(u)) \\ &= f(u). \end{aligned}$$

But we can also define T_Z for arbitrary Z. It can be checked that T_Z is a contraction mapping in the sense that, if we take any two bounded functions g_1 and g_2 on $[u_0, u_e]$, we have

$$||T_Z g_1 - T_Z g_2|| \leq \delta ||g_1 - g_2||,$$

where $||g|| = \sup_{u\in[u_0,u_e]} |g(u)|$. Standard fixed point theorem imply that for each T_Z, there exists a unique g_Z such that

$$T_Z g_Z = g_Z.$$

In other words, take any Z, we can define T_Z and find g_Z.

The second step involves how to find Z^* such that $Z^* = f(u_0)$. Now for each Z, the first step induces a fixed point $g_{Z\cdot}$ And in particular, this gives a value of $g_Z(u_0)$. In the appendix, we show that

$$0 \leq \frac{dg_Z(u_0)}{dZ} < 1.$$

In other words, the function from Z to $g_Z(u_0)$ is again a contraction mapping. And it has a unique solution such that

$$Z^* = g_{Z^*}(u_0) = f(u_0) = g_{f(u_0)}(u_0). \tag{7}$$

And in particular, the contraction mapping above is monotone, so Z^* can be found using standard numerical procedures.

In summary, $f(u_0)$ can be found as follows. First, for each Z, we map Z into $g_Z(u_0)$ through finding g_Z as a fixed point of the contraction mapping T_z (defined in (6)). Second, we find $f(u_0)$ by noting that it is the fixed point of the contraction mapping from Z into $g_Z(u_0)$. We call this procedure "finding a double fixed point".

Theorem 1: *When* $py - \frac{[1-\delta(1-p)]c}{\delta(p-q)} \geq \underline{v} + \underline{w}$ *and* $\underline{w} \geq \underline{u}$, *the Pareto frontier of the PPE payoff is the unique function that satisfies*

$$f(u) = \begin{cases} \underline{v} + \frac{u-\underline{u}}{u_0-\underline{u}}(f(u_0) - \underline{v}) & \text{if } u \in [\underline{u}, u_0] \\ T_{Z^*} f(u) & \text{if } u \in [u_0, u_e] \\ f(u_e) + u_e - u & \text{if } u \in [u_e, u_e + f(u_e) - \underline{v}], \end{cases}$$

where $u_e = (\underline{w} - c) + \frac{\delta pk}{(1-\delta)}$, $f(u_e) = py - c - ((\underline{w} - c) + \frac{\delta pk}{(1-\delta)})$, *and* $u_0 = (1-\delta)(\underline{w} - c) + \delta(\underline{u} + pk)$. *Finally,* $f(u_0) = Z^*$, *where* Z^* *satisfies* $Z^* = g_{Z^*}(u_0)$ *and* g_Z *is the fixed point of the operator defined in (6).*

3.2.4 Inefficiency in Optimal Relational Contract

In this subsection, we show that the optimal relational contract under limited liability is inefficient. The optimal relational contract is the PPE that maximizes the principal's payoff. While this is a theoretical point, this inefficiency result also helps characterize the employment dynamics in the next section.

Before proving the inefficiency, we first state a corollary of Theorem 1.

Corollary 1: *For almost all* $u \in [u_0, u_e]$,

$$f'(u) = pf'(L(u) + k) + (1 - p)f'(L(u)). \tag{8}$$

Proof. Because f is concave, the left derivative is equal to the right derivative almost everywhere. (Royden, Chapter 5). The equation follows directly Lemma 6. ■

Proposition 1: *There exists $u < u_e$ such that*

$$f(u) > f(u_e).$$

Proof. Take $u_0' < u_e$ such that $f(u_0')$ is differentiable and $L(u_0') + k > u_e$. If $f'(u_0') < 0$, then we are done, because

$$f(u_0') - f(u_e) = -\int_{u_0'}^{u_e} f'(u)du,$$

where this is a Lebesgue integral, and we have $f'(u) < 0$ almost everywhere due to the concavity of f.

Otherwise, we take $\widetilde{u}_1$ such that $L(\widetilde{u}_1) = u_0'$. According to (8) and the fact that $u_0' < u_e < u_0' + k$, within the interval of $[\widetilde{u}_1, \widetilde{u}_1 + u_e)$, we can find u_1' such that

$$f'(u_1') \leq -p + (1-p)f'(u_0').$$

This procedure can continue forever, i.e., we can find $u_{n+1}' \in (u_n', u_e)$ such that

$$f'(u_{n+1}') \leq -p + (1-p)f'(u_n')$$

It is then clear that there exists an N such that $f'(u_N') < 0$, and we are done. ■

In Theorem 1, we know that the right derivative of $f(u_e)$ is -1. Proposition 1 essentially shows that the left derivative of $f(u_e)$ is also -1, even if f has infinitely many kinks in any open neighborhood of u_e. Since u_e is defined as the smallest payoff of the agent that maximizes the joint surplus, Proposition 1 implies that the optimal contract cannot be efficient.

The characterization implies that the optimal relational contract, the PPE that maximizes the principal's payoff, is inefficient in the sense that termination occurs with a positive probability. This is because when there is limited liability, there is rent in the relationship for the agent. Consequently, the principal can use the threat of termination to induce effort. Since output depends both on effort and luck in this model, termination can occur even if the agent always put in effort. Of course, termination is costly to the principal as well. It is not clear that the principal will prefer using termination over giving bonus to provide incentive at the margin. In fact, when high output is unlikely and we restrict our attention to stationary contracts, the optimal relational contract in this class involves no termination.

The more efficient way to provide incentive is to backload the reward of the agent. When a good outcome appears, the agent can be rewarded either by a performance bonus or an

increased chance/probability of permanent employment. By rewarding the latter, the principal not only saves the wage bill, but also reduces the probability of termination, which is costly to the principal. Therefore, the principal strictly prefers rewarding the agent with the increased probability of permanent employment, and this implies that no performance bonus will be paid out until the agent receives permanent employment.

Under this efficient way of incentive provision, the marginal cost for the principal of using termination to provided approaches zero when no termination is used. This is because when the probability of termination is very small, termination tends to happen very late in the relationship. By then, agent has internalized most of the cost of termination (because the principal does not pay the agent any bonus before termination occurs even if the outputs are good). This point is illustrated more clearly in the example in the next subsection.

3.3 Pareto Frontier with Countable Number of Kinks

In Subsection 3.2, we show that f can be found by solving for a double fixed point problem. In this subsection, we show that, for some parameters, it is possible to calculate f explicitly.

The condition for the closed-form expression for f is that

$$\underline{u} + k = L(u_0) + k \geq u_e = (\underline{w} - c) + \frac{\delta p k}{(1-\delta)}.$$

Or equivalently,

$$\underline{u} + \frac{(1-\delta-\delta q)}{\delta(p-q)} c \geq \underline{w}. \tag{9}$$

This condition states that the expected payoff of the agent is greater than or equal to u_e following any good outcomes. And the condition is more likely to be satisfied when the probability of success (p) is small and when the discount factor (δ) is small.

Now applying Corollary 1 here, we have that for $u \in [u_0, u_e]$

$$f'(u) = -p + (1-p) f'(L(u)).$$

This formula enables us to partition the interval of $[u_0, u_e]$ into countable number of sub-intervals and then calculate the value of $f'(u)$ sub-interval by sub-interval.

In particular, define u_1 such that

$$L(u_1) = u_0.$$

In other words, u_1 is the agent's payoff such that his low continuation payoff falls to u_0. It is clear that

$$L(u) \in [\underline{u}, u_0] \quad \text{for all } u \in [u_0, u_1].$$

Some algebra gives that

$$u_1 - u_0 = \delta(u_0 - \underline{u}) = \delta((1-\delta)(\underline{w} - c - \underline{u}) + \delta pk).$$

Next define u_{n+1} such that

$$L(u_{n+1}) = u_n.$$

In other words, if the agent's expected payoff is u_{n+1}, his continuation payoff following a bad outcome is u_n. This construction implies that, for an agent with expected payoff u_n, he is guaranteed to stay in the relationship for at least n more periods.

It can be checked that

$$L(u) \in [u_{n-1}, u_n] \text{ for all } u \in [u_n, u_{n+1}].$$

$$u_{n+1} - u_n = \delta^{n+1}(u_0 - \underline{u}).$$

Moreover, we can show that

$$\lim_{n\to\infty} u_n = u_e.$$

In other words, if the agent has an expected payoff exceeding u_e, he is guaranteed that he will never be terminated.

This partition of the agent's payoff into countable number of intervals helps illustrates that why the optimal relationship contract is necessarily inefficient. Note that if the principal decreases the agent's payoff u_e by ε, we that $u_e - \varepsilon \in (u_{N(\varepsilon)}, u_{N(\varepsilon)+1}]$ for some $N(\varepsilon)$. As ε goes to 0, $N(\varepsilon)$ approaches infinity. In other words, by moving away a bit from the efficient equilibrium, termination can only occur in the very distant future. In particular, the efficiency loss of the relationship is bounded by $\delta^{N(\varepsilon)}$, which goes to zero in an exponential speed. As ε goes to 0, it can be shown that the ratio of $\delta^{N(\varepsilon)}$ to ε goes to zero. This implies that when the principal decreases the agent's payoff to $u_e - \varepsilon$, the principal's payoff increases to $f(u_e) + \varepsilon$. Therefore, the efficient contract cannot be optimal.

The next corollary shows that the Pareto frontier is formed by a sequence of line segments, where each line segment joining $(u_n, f(u_n))$ and $(u_{n+1}, f(u_{n+1}))$ for some n. The proof of this result can be found in the appendix.

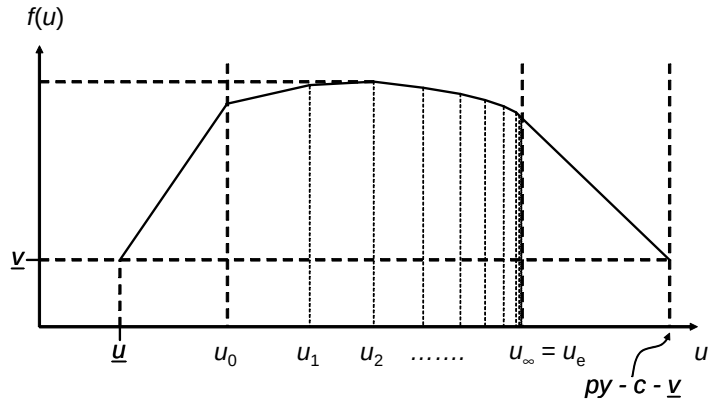

Figure 4: Pareto Frontier w/ Countable Number of Kinks

Corollary 2: *If Assumption 1 and 2 holds, and also* $\underline{u} + \frac{(1-\delta-\delta q)}{\delta(p-q)}c \geq \underline{w}$*, then*

$$f(u) = \begin{cases} \underline{v} + s_0(u - \underline{u}) & \text{if } u \in [\underline{u}, u_0] \\ \underline{v} + \frac{(u_0-\underline{u})}{1-\delta(1-p)}(s_0 - \frac{p\delta(1-\delta^{n+1})}{1-\delta} - \delta^{n+1}s_n(1-p)) + s_{n+1}(u - u_n) & \text{if } u \in [u_n, u_{n+1}] \\ f(u_e) + u_e - u & \text{if } u \in [u_e, u_e + f(u_e) - \underline{v}], \end{cases}$$

where $u_0 = (1-\delta)(\underline{w} - c) + \delta(\underline{u} + pk)$, $s_0 = \frac{(1-\delta)(py-\underline{w})+\delta((1-p)\underline{v}+p(py-c-(\underline{u}+k)))-\underline{v}}{(1-\delta)(\underline{w}-c)+\delta(\underline{u}+pk)-\underline{u}}$, $u_n = u_0 + \frac{\delta(1-\delta^n)}{1-\delta}(u_0 - \underline{u})$, $s_n = s_0 - (1+s_0)(p + (1-(1-p)^{n+1}))$, $u_e = (\underline{w} - c) + \frac{\delta pk}{(1-\delta)}$, $f(u_e) = py - c - ((\underline{w} - c) + \frac{\delta pk}{(1-\delta)})$.

Corollary 2 implies that the graph of f has a kink at each of the u_n. Therefore, f is not differentiable at each point. Interestingly, contrary to most examples in the literature where the PPE payoff set typically has finite or a continuum of extremal points, we have a countable number of extremal points here, and they converge to u_e.

4 Empirical Implications

In this section, we derive some properties of the optimal relational contract using the characterization of the PPE frontier in the previous section. In Subsection 4.1, we characterize the employment dynamics of the relationship. In Subsection 4.2, we derive comparative static results on how the outside options and the minimum wage affects the principal, agent's payoffs and the overall efficiency. All these implications are derived under Assumption 1 and 2, and we discuss in the next section how the predictions change under more general conditions.

4.1 Patterns of Employment Dynamics

In this subsection, we derive the empirical implications of optimal relational contract. In Subsection 4.1.1, we show that the optimal relational contract specifies a probation phase. In Subsection 4.1.2, we show that the employee's compensation is deferred in the sense that his expected compensation increases over time even if his expected productivity remains constant. In Subsection 4.1.3, we explore the turnover dynamics. We show that the optimal relational contract can generate "tenure": there is a fixed time in which the employee will be terminated if no permanent employment is obtained prior to that. Subsection 4.1.4 shows show that earlier successes matter more for receiving permanent employment.

4.1.1 Probation

The next proposition characterizes the employment dynamics of the optimal relational contract. While the optimal relational contract with limited liability is not stationary, it is still relatively simple. Essentially, the agent's employment dynamics can be categorized into three phases. The agent starts with Phase 1, in which he is paid the wage floor $\underline{w}$ per period. Depending on the outcomes, the agent either moves into Phase 2, in which the relationship is terminated and the two parties receive outside option $(\underline{u}, \underline{v})$ forever. Or the agent enters Phase 3, in which the relationship is never terminated. In this phase, the remaining relational contract can be implemented stationarily: the agent receives a fixed base wage, and he receives a bonus (to be paid out at the beginning of next period in our setting) every time the output is high. Moreover, Phase 2 and Phase 3 are absorbing in the sense that, as the time goes to infinity, the agent will either be in one of the two phases with probability 1, and once the agent is in that phase, he stays there forever.

Proposition 2: *In the optimal relational contract, the set of histories can be partitioned into $H = H_1 \cup H_2 \cup H_3$, such that*

(i): ($H_1$is the probation phase): When $h^t \in H_1$, $w(h^t) = \underline{w}$.

(ii): (H_2 is the termination phase): When $h^t \in H_2$, both the principal and the agent receive their outside options $(\underline{u}, \underline{v})$.

(iii): (H_3 is the permanent employment phase. Optimal relational contract in H_3 may be stationary): When $h^t \in H_3$, the optimal relational contract can be implemented in the following way:

$$\begin{aligned} w(h^{t+1}) &= \underline{w} \qquad \text{if } y_t = 0; \\ &= \underline{w} + k \quad \text{if } y_t = y. \end{aligned}$$

(iv): (Once the agent is in Phase 2 or 3, he stays there forever): If $h^t \in H_i$, for $i = 2, 3$, then $h^{t+k} \in H_i$ if $h^{t+k}|t = h^t$.

(v): (Employment starts with the probabtion phase)

$$h^1 \in H_1.$$

(vi): (Phase 2 and 3 are absorbing)

$$\lim_{t\to\infty} \Pr(h^t \in H_2 \cup H_3) = 1.$$

Proof. Define H_3 as the set of histories such that the agent's continuation payoff $u \geq u_e$. Define H_2 as the set of histories such that the agent's continuation payoff $u = \underline{u}$. By Theorem 1, it is clear $H_2 \cap H_3 = \Phi$. Define $H_1 = H \backslash (H_2 \cup H_3)$.

By Proposition 1, we know that the game starts in H_1, so (v) is proved. Theorem 1 also directly gives (i). Since $(\underline{u}, \underline{v})$ is the unique PPE payoff in which the agent's payoff is $\underline{u}$ and it is supported by taking outside option forever, (ii) is proved. Now if $u \geq u_e$, then $f(u)+u = py-c$, so the continuation payoff must be $py - c$ as well. Moreover, $(u, f(u))$ can be implemented by paying the agent $\underline{w} + \frac{u-u_e}{1-\delta}$, in this period, and uses $(u_e, f(u_e))$ and $(u_e + k, f(u_e) - k)$ as continuation payoffs forever. This proves (iii) and (iv). (vi) follows because take any u, there exists an N such that with a fixed probability bounded away from 0, the agent either have continuation payoff (u, v) or $u \geq u_e$. Then (vi) follows from standard statistics arguments. ■

The three phases in the employment dynamics correspond to the three regions in the PPE payoff set. Proposition 1 implies that the optimal relational contract is in the middle region, and the continuation payoffs bounce around according to the outcomes. If the outcome is good, the continuation payoff moves to the right. Otherwise, it moves to the left. When the continuation payoff moves across the right threshold (u_e) the agent receives permanent employment, and the remaining optimal contract can be implemented in a sequence of stationary contracts (cf Levin (2003)). When the continuation payoff cross the left threshold (u_0), then termination occurs with some probability.

Phase 1 of the employment resembles the probation period in employment contracts, which are often defined as the period during which the employee can be fired without cause. The probation phase is an important feature for many occupations, including lawyers, doctors, professors, and government officials. The probation periods have received some attention from labor economists; including Bull and Tedeschi (1989), Sadanand et al. (1989), Weiss and Wang (1990), and Wang and Weiss (1998). Most of these models assume that workers differ by their types and the probation period is used as a sorting device to screen out the bad types. And no doubt that this is an important function of the probation period; see for example Loh (1994).

Our analysis suggests that the probation period can also serve as an incentive device and may arise even when agents are homogeneous. Workers exert effort in the probation phase in

hope of receiving permanent employment as reward. Riphahn and Thalmaier (1999) find significant responses of white collar employees and public sector workers to probation periods: once employment probation is completed and individuals enter into regular employment contracts, the probability of work absences takes discrete jumps and is significantly above previous levels.[11]

In addition, the previous theoretical literature implies that there is a fixed duration for the probation period, and our analysis suggests that the probation period can be random. Casual empiricism suggests that while the probation periods in many employment contracts have a fixed duration, the employer often reserve the right to change the probation duration, and in some jobs, the probation duration is not explicitly written down.

4.1.2 Deferred Compensation

The employment structure in Proposition 2 immediately implies that the compensations of the workers are backloaded. In addition, the expected profit of the employer on the employee decreases over time.

Corollary 3: *If the optimal relational contract is implemented by a sequence of stationary contracts once the worker receives permanent employment, then*

(i): The expected wage of the agent within the employment relationship is nondecreasing over time.

(ii): The expected payoff of the principal is nonincreasing over time.

Proof. (i) is immeidate. For (ii), the principal's expected payoff is a constant when she's in Phase 2 or Phase 3. When in Phase 1, the principal is always getting $py - \underline{w}$ per period, which is higher than her normalized payoff. ■

Backloaded compensation has receives much attention from economists; see for example Salop and Salop (1979) for a screening argument, Lazear (1981) for an incentive argument, and Harris and Holmstrom (1982) for a learning and insurance argument.

Perhaps closest to the current paper is Akerlof and Katz (1989), who characterize the optimal dynamic moral hazard under limited liability in an efficiency wage framework. They show that the optimal contract takes the form of a trust fund, which represents the agent's loss if caught shirking. When worker cannot post a bond, which is one form of limited liability, Akerlof and Katz (1989) shows that rent in the relationship for the agent is necessary to induce effort for all periods.

[11] Of course, our model does not predict this phenomenon since the effort level in our model is binary. Riphahn and Thalamier (1999)'s result suggests that efforts are higher during the probation phase.

Two important assumptions in Akerlof and Katz (1989) are a) there is perfect monitoring for the worker when he puts in effort, i.e. a worker who puts in effort will never be caught shirking (corresponding to $p = 1$ in this model, and b) the firm can commit to wage payments. In many economic situations, it may be argued that efforts can be difficult to evaluate and firms may find it difficult to contract on outputs.

This model can be thought of as an extension of the Akerlof and Katz (1989) model by a) allowing for imperfect monitoring of effort and b) assuming that the firm cannot contract on outputs (and thus cannot commit). This extension preserves the basic feature of wage dynamics in Akerlof and Katz (1989) in the sense that compensation is deferred. Moreover, it enriches Akerlof and Katz (1989) by highlighting the optimal use of termination and bonus as incentive device and thus generates turnover dynamics (and the associated inefficiency) in the relationship. Finally, the lack of commitment power of the firm adds restrictions to the optimal wage profiles: for example, the firm cannot postpone the bonus payment to the worker indefinitely.[12]

In addition to the backloaded compensation, the model gives more specific prediction on the composition of compensation: the bonus to total compensation ratio increases over employment duration as well. This prediction appears to fit the incentive structure of some occupations. In sales job in particular, it is well-known that the commission rate increases with seniority, see for example Lin (2005) and Coughlan et al (2005). It will be interesting to test this prediction more broadly across occupations using large datasets.

The flip side of the backloaded compensation is that the profitability of the principal over the agent decreases over time. Of course, this prediction can be reversed if we assume that the worker can accumulate firm-specific human capital over time. Nevertheless, it is not impossible that firms may make less money on their more senior employees. For example, Circuit City studied the profitability of their salesforce and discovered that their more experienced salespersons actually delivered lower profit.

It should be emphasized that the dynamic incentive studied here implies that "firing workers with the lowest marginal productivity/wage ratio" may not be the optimal firing policy. Firing more senior (and more expensive) workers may be interpreted as a violation of the implicit contract, and anticipation of such action may result in lower efforts from the workers. This appears to be what has happened in Circuit City. Philp Schoonover, who resigned as CEO in Sept 2008, has been criticized for his blunder in firing the more experienced workers:

> "Schoonover was slammed for his cost-cutting decision in early 2007 to fire 3,400 of Circuit City's most experienced employees. The company said at the time that

[12] While there are still some uncertainty in the wage dynamics when the surplus is large, the uncertainty diminishes as the surplus decreases and disappears completely for the parameters studied in Section 5.1.

they were earning too much money and could be replaced with cheaper workers. But analysts said the move devastated morale and led to a decline in service." Business Week, Sept 22, 2008

4.1.3 Turnover Dynamics and Tenure

Due to the uncertainty in output, this model provides rich turnover dynamics. In particular, the turnover rate may not be monotone with respect to the employment duration. In some cases, the optimal relational contract implies a turnover pattern that has features of "tenure": there is a fixed time in which the employee will be terminated if no permanent employment is obtained prior to that.

Corollary 4: *When* $\underline{u} + \frac{(1-\delta-\delta q)}{\delta(p-q)}c \geq \underline{w}$, *there exists* T^* *such that the turnover rate is* 0 *for* $t < T^*$ *and is again* 0 *for* $t > T^* + 1$. *Generically, turnover happens only in* T^*.

Proof. When $\underline{u} + k \geq (\underline{w} - c) + \frac{\delta pk}{(1-\delta)}$, Corollary 2 gives an explicit formula of f. There are two cases to consider. In case 1, there's a unique u_n that maximizes $f(u)$. In this case, if any of the output in the first $n+2$ periods is positive, the agent receives permanent employment. Otherwise, the agent's continuation payoff moves to u_{n+1-t} in period t, and is terminated at time $t = n + 2$.

In case 2, there exists n such that $f(u)$ is maximized in $[u_{n-1}, u_n]$. In this case, if no positive outcome has been generated, the agent's continuation will be in $[u_{n-t}, u_{n+1-t}]$ in time t. And the agent will be terminated either in time $t = n + 1$ or $t = n + 2$. ■

In this example, the condition for the parameters is exactly that in Corollary 2. Therefore, the Pareto frontier has countable number of extremal points in this example. The optimal relational contract corresponds to one of the kinks (u_n for some n). One success moves the agent to permanent employment. And one failure moves the continuation payoff to the kink adjacent to the left until $(u_0, f(u_0))$. And termination occurs when failures occurs at $(u_0, f(u_0))$.

Note that the condition in the example is more likely to be satisfied when the probability of success is small. One labor market that fits this assumption and also has tenure as its key feature is the academia: writing a few home-run papers can bring an assistant professor over the tenure hurdle, and failure to do so before a fixed date leads to termination of the employment contract.

The turnover pattern from the tenure system is an example of (degenerate) inverse-U shaped turnover patterns with respect to employment duration. The inverse-U shaped turnover pattern appears to be hard to generate theoretically but empirically relevant, see Jovanovic (1979) for

a model and Farber (1994) for an empirical investigation. It will be interesting to investigate the turnover patterns for other parameter values of this model. One possibility is to look at the continuos time limit of this discrete time model, and then the turnover pattern may be obtained through simulation or even analytically.

It should be pointed out that while this model generates rich turnover patterns, there is still a commonality among all possible turnover patterns. The model predicts that, as time goes to infinity, the turnover rate goes to zero. This is because as time goes to infinity, for a worker remaining in the employment relationship, the probability that he has received permanent employment has approached 1.

Finally, we should note that termination is not renegotiation proof in the sense that both parties would rather keep the relationship (and start anew for example) when termination occurs. However, termination is useful ex ante to generate higher profits for the principal. In the case of tenure, in particular, this model provides a rationale for up-or-out contracts.

4.1.4 Path Dependency

We finish this subsection by noting that earlier successes benefit the agent more. In particular, the agent prefers having a high output followed by a low output than the other order. Earlier success is more important both in terms of increasing future expected compensation and in terms of reducing the probability of termination. We state the result formally in the next corollary.

Corollary 5: *For an agent with expected payoff* $u \in [u_0(\underline{u}), u_e]$, *define* $U_h(u)$ *as his continuation payoff following a high output. Also define* $U_{hl}(u)$ *and* $U_{lh}(u)$ *similarly. If* $U_h(u) \in [u_0(\underline{u}), u_e]$, *then*

$$U_{hl}(u) \geq U_{lh}(u).$$

Proof. Direct calculation. ■

The path dependency results from the way the incentive is structured. Since the reward for success in the probation phase in this model is deferred, this implies that the agent will receive a bigger "interest payment" for earlier success. The interest payment takes the form of a higher probability of higher future expected payoffs, and in particular, a lower probability of termination.

This prediction is related to the idea of "fast track" in the internal labor market literature, which states that workers who experience high wage growth in the past are more likely to have high rates of wage growth in the future. More formally, changes of wages are positively serially correlated; see for example Baker, Gibbs, Holmstrom (1994a, b). To explain this phenomenon,

the existing literature has relied on heterogeneity in individual abilities; see for example Meyer (1991), Bernhardt (1995), and Gibbons and Waldman (1998).

In this model, while there is no fast track result for wages because there are multiple wage paths consistent with the optimal relational contract once the worker receives permanent employment, Corollary 5 implies a fast track result concerning turnover. In particular, termination is less likely for workers who have good performances earlier than those whose success come later. It will be interesting to test this prediction if data on performance evaluations and turnover rates are available. The academic labor market is an example in which this prediction may be tested.

4.2 Outside Options, Wage Floor, and Efficiency

In this subsection, we study how the optimal relational contract is affected by the outside options of the players and by the wage floor. We show that first, as the agent's outside option improves, the principal's expected payoff decreases, the agent's expected payoff decreases, and the overall efficiency increases. Second, as the principal's outside option improves, the principal's expected payoff increases, the agent's expected payoff decreases, and the overall efficiency decreases. Moreover, the aggregate turnover rate of the agent increases. Finally, as the wage floor increases, the principal's payoff decreases, but the change in agent's payoff is ambiguous.

While most of the results of our comparative statics are intuitive, our proof method may be of independent interest. We prove our results by exploiting certain monotonicity properties of functional operators. Difficulties in obtaining explicit comparative statics results on dynamic incentive problems in discrete time setting is one motivation for studying continuous time models; see for example DeMarzo and Sannikov (2007). Here, comparative statics results can be obtained without knowing the exact value of the optimal PPE payoff. Moreover, some of the comparative statics results shed new light on important policy issues. For example, while most of the debate on minimum wages focus on the total number of employment. Our analysis indicates that minimum wage may harm workers who are already employed through an in increase in the probability of firing.

We collect our proofs in the appendix. We give detailed proofs on the effects of the agent's outside option on the payoffs of the players and the efficiency, and we sketch out the proofs of the effects of the principal's outside option and the wage floor because the proofs are similar.

We first examine the effect of the agent's outside option. Define $F(u, \underline{u}, \underline{v}, \underline{w})$ as the Pareto frontier of the PPE payoff when the agent's value is u, his outside option is $\underline{u}$, the principal's outside option is $\underline{v}$, and the wage floor is $\underline{w}$. For notational simplicity, we omit $(\underline{v}, \underline{w})$ and define $u_0(\underline{u})$ as the agent's payoff such that $L(u_0(\underline{u})) = \underline{u}$. Define $u^P(\underline{u}, \underline{v}, \underline{w})$ as the maximum

equilibrium payoff of the principal when the agent's outside option is $\underline{u}$. Define $u^A(\underline{u}, \underline{v}, \underline{w})$ as the associated payoff of the agent. We have the following results.

Proposition 3: *If* $\max_u\{F(u, \underline{u}, \underline{v}, \underline{w})\} > F(u_0(\underline{u}), \underline{u}, \underline{v}, \underline{w})$, *then*

$$\begin{aligned} \frac{\partial u^P(\underline{u}, \underline{v}, \underline{w})}{\partial \underline{u}} &< 0; \\ \frac{\partial u^A(\underline{u}, \underline{v}, \underline{w})}{\partial \underline{u}} &> 0; \\ \frac{\partial (u^P(\underline{u}, \underline{v}, \underline{w}) + u^A(\underline{u}, \underline{v}, \underline{w}))}{\partial \underline{u}} &> 0. \end{aligned}$$

Proposition 3 shows that as the agent's outside option increases, his equilibrium payoff increases, the principal's payoff increases, and the overall efficiency increases. The intuition for these results is as follows. When the agent's outside option increases, the rent in the relationship for the agent is smaller so the threat of termination becomes a less effective way for the principal to induce effort. Therefore, the principal will use bonus more as incentives. Since bonus gives rents to the agent, the principal's payoff decreases, and this establish the first inequality. When the optimal relational contract results in less terminations, it follows directly that the overall efficiency increases, which is the third inequality. Finally, in this change, the agent's expected payoff increases both because he's less likely to be terminated and because even if he's terminated, his expected payoff is higher, and thus establishes the second inequality above.

A similar reasoning explains these comparative statics effects are reversed when the principal's outside option increases. In this case, the principal finds the threat of termination becomes more cost effective. Therefore, the principal will use termination more often, and the efficiency of the relationship decreases even if the payoff of the principal increases. Finally, the agent's payoff decreases because he's more likely to be terminated. Proposition 4 states these results formally.

Proposition 4: *If* $\max_u\{F(u, \underline{u}, \underline{v}, \underline{w})\} > F(u_0(\underline{u}), \underline{u}, \underline{v}, \underline{w})$, *then*

$$\begin{aligned} \frac{\partial u^P(\underline{u}, \underline{v}, \underline{w})}{\partial \underline{v}} &> 0; \\ \frac{\partial u^A(\underline{u}, \underline{v}, \underline{w})}{\partial \underline{v}} &< 0; \\ \frac{\partial (u^P(\underline{u}, \underline{v}, \underline{w}) + u^A(\underline{u}, \underline{v}, \underline{w}))}{\partial \underline{v}} &< 0. \end{aligned}$$

A direct consequence of Proposition 4 is that the overall turnover rate increases with the principal's outside option. While the principal's option is treated as an abstract parameter,

both local market conditions or the prestige of the employer may be used as proxies. It will be interesting to test this prediction empirically.

Corollary 6: *The aggregate turnover probability increases with* $\underline{v}$.

Proof. This follows from that u_0 does not change with $\underline{v}$ and the agent's payoff decreases, so a standard coupling argument works. ■

Finally, we examine the effect of the wage floor. We show that the wage floor decreases the principal's payoff and the overall efficiency.

Proposition 5: *If* $\max_u\{F(u,\underline{u},\underline{v},\underline{w})\} > F(u_0(\underline{u}),\underline{u},\underline{v},\underline{w})$, *then*

$$\begin{aligned} \frac{\partial u^P(\underline{u},\underline{v},\underline{w})}{\partial \underline{w}} &< 0 \\ \frac{\partial (u^P(\underline{u},\underline{v},\underline{w}) + u^A(\underline{u},\underline{v},\underline{w}))}{\partial \underline{w}} &< 0. \end{aligned}$$

The intuition that the wage floor hurts the principal's payoff is straightforward because more rents need to go to the agent's pocket. It follows that termination becomes a relatively more attractive method of inducing incentives because termination is now relatively less costly for the principal. Therefore, the principal will use more of termination as an incentive device, and termination is more likely to occur. This lowers the efficiency of the relationship.

The effect of wage floor on the agent's expected payoff is ambiguous because there are two forces at work. On the one hand, wage floor raises the base wage of the agent, and this helps raises his expected payoff. On the other hand, the agent is more likely to be terminated and this damages his welfare. We suspect that if the relationship is less likely to be terminated, then the wage floor may increase the agent's welfare. On the other hand, if termination is frequent, then wage floor may actually decrease the agent's payoff.

Finally, we note the following caveat. We derive the welfare property of the wage floor by holding the principal and the agent's outside options fixed. In general, the wage floor may affect the outside options of the players as well. To the extent that higher wage floor may increase the agent's outside option, then it may be welfare enhancing. We think that multiple equilibrium is possible and it will be interesting to explore this important issue further.

5 Alternative Assumptions

In this section, we study the properties of our model under alternative assumptions. Section 5.1 studies the properties of the optimal relational contract when the amount of surplus is too

low to sustain full efficiency. Section 5.2 studies the case where the wage floor is lower than the outside option of the agent. Section 5.3 discuss how the results of the paper change when the outputs can be contracted on and the principal can commit.

5.1 Low Surplus

In this subsection, we study the property of the optimal relational contract when there is insufficient amount of surplus in the relationship so that the efficient outcome (the agent puts in effort in each period) cannot be supported as a PPE. In other words, Assumption 1 fails here so that

$$py - \frac{[1-\delta(1-p)]c}{\delta(p-q)} < \underline{v} + \underline{w}.$$

When there is insufficient surplus, the analysis remains similar to the case with sufficient surplus and the resulting Pareto frontier is also similar. In particular, the Pareto frontier continues to have three regions, where a) termination happens stochastically to the left of a threshold, b) bonus is paid out to the right of another threshold, and c) no bonus is paid out and no termination occurs between the two thresholds. Moreover, the value of the left threshold remains unchanged in this case. Formally, among the lemmas that determine the Pareto frontier in Section 3, Lemma 1 and Lemmas 3-7 continue to hold.

On the other hand, the value of the right threshold will be different in this case. In particular, since the efficient output is not possible in this case, the formula for right threshold u_e in the sufficient surplus case (Lemma 2) is no longer valid. There is no closed-form solution for the value of u_e. However, we have the following lemma that relates u_e and $f(u_e)$.

Lemma 10: *Suppose* $py - \frac{[1-\delta(1-p)]c}{\delta(p-q)} < \underline{v} + \underline{w}$, *then*

$$f(L(u_e) + k) = \underline{v}.$$

Proof. Recall that u_e is the minimum payoff the agent obtains among the constrained efficient PPEs. We know that the slope of f for $u > u_e$ is -1 and the agent is paid $\underline{w}$ in period 1. Moreover, when $py - \frac{[1-\delta(1-p)]c}{\delta(p-q)}$, we must have

$$L(u_e) < u_e,$$

because otherwise we would have an efficient PPE.

Let $u_{\max}$ be the maximum payoff of the agent such that $f(u_{\max}) = \underline{v}$. It is clear that $L(u_e) + k \leq u_{\max}$.

If $L(u_e)+k < u_{\max}$, then consider the equilibrium profile $(u_e+\varepsilon,\, f(u_e+\varepsilon))$. The discussion above implies that

$$f(u_e+\varepsilon) = f(u_e) - \varepsilon.$$

On the other hand,

$$f(u_e+\varepsilon) \geq (1-\delta)(py-\underline{w}) + \delta((1-p)f(L(u_e+\varepsilon)) + pf(L(u_e+\varepsilon)+k)).$$

Now for small enough ε, we have $L(u_e+\varepsilon) < u_e$. Now since $f'(u) > -1$ for $u < u_e$, the above implies that $f'(u_e+\varepsilon) > -1$ for small enough ε. Therefore, we have

$$f(u_e+\varepsilon) > f(u_e) - \varepsilon,$$

and this is a contradiction. ■

Lemma 10 shows that the value of $f(u_e)$ is completely given by the value of u_e. This allows us to characterize the Pareto frontier completely. As in Section 3, we know that the two thresholds (u_0 and u_e) divide the Pareto frontier into three regions, where both the left and right regions are line segments, and the middle region is determined by the functional equation in Lemma 6. Essentially, the Pareto frontier is solved if we know the value of the two thresholds and the associated equilibrium payoffs for the principal.

As in Section 3, we know the value of u_0 but not that of $f(u_0)$. Since the value of $f(u_0)$ is determined both by the left region (as the end of a line segment) and the middle region (as the solution of a functional equation), we solve for $f(u_0)$ by equating the value of the two, just as in Section 3. Different from Section 3, we don't know the value of u_e. But Lemma 10 shows that $f(u_e)$ is determined once u_e is chosen. In addition, the value of $f(u_e)$ is determined by solving the functional equation in Lemma 6. It turns out that the value of $f(u_e)$ can be solved by equating these two value.

Numerically, this can be done in a three step procedure. First, we choose a candidate $u_e = Z'$, with the initial value being possibly the one in Section 3. We then perform the two-step procedure in Section 3 by finding $f(u_0|Z')$ whose value satisfy both the equation that governs the left region and the functional equation in Lemma 6 that governs the middle region. This two-step procedure generates a value of $f(u_e|Z')$, given by the functional equation in Lemma 6. If the value of $f(u_e|Z')$ is smaller than the value of $f(u_e)$ given by Lemma 10, we move our next guess of u_e to be less than Z', and otherwise we make a guess larger than Z'. Similar argument as in Lemma 7 shows that this is a contraction mapping, and we can find a unique Z^* such that $u_e = Z^*$ leads to unique solution to the functions governing the Pareto Frontier.

We summarize our discussion in the following theorem:

Theorem 2: *When* $py - \frac{[1-\delta(1-p)]c}{\delta(p-q)} < \underline{v} + \underline{w}$, *and* $\underline{w} \geq \underline{u}$, *the Pareto frontier of the PPE payoff is the unique function that solves the following equation*

$$f(u) = \begin{cases} \underline{v} + \frac{u-\underline{u}}{u_0-\underline{u}}(f(u_0) - \underline{v}) & \text{if } u \in [\underline{u}, u_0] \\ (1-\delta)(py - \underline{w}) + \delta[pf(L(u)+k) + (1-p)f(L(u))] & \text{if } u \in [u_0, u_e] \\ f(u_e) + u_e - u & \text{if } u \in [u_e, u_e + f(u_e) - \underline{v}], \end{cases} \tag{10}$$

where $f(L(u_e)+k) = \underline{v}$, *and* $u_0 = (1-\delta)(\underline{w} - c) + \delta(\underline{u} + pk)$.

The discussion above shows that the main features of the Pareto frontier remain the same, but the technical aspects of finding the Pareto frontier become more challenging. The basic logic is again that the optimal relational contract will involve an efficient combination of termination and bonus to provide incentive. And in particular, the optimal relational contract starts with a "probation phase," and bonus will be paid out only if the agent's continuation payoff exceeds a threshold.

While the shape of the Pareto frontier remains the same, there are also differences. The most important difference is that there will no longer be permanent employment as part of the optimal relational contract in this case. The reason is that, when future surplus is low, it is not possible to sustain an efficient equilibrium (in which the agent puts in effort each period) because the bonus required to reward good output exceeds the total future surplus of the relationship. It follows that when the surplus is small, in order to induce effort from the agent it is always necessary to punish the agent with some probability of termination.

It should be emphasized that such termination may not be carried out immediately, but instead take the form that "if many low outputs occur in a row, then there will be positive probability of termination at some point". Nevertheless, even if there may be zero probability immediate termination for some periods, termination will be carried out eventually as the next proposition shows.

Proposition 6: *If* $py - \frac{[1-\delta(1-p)]c}{\delta(p-q)} < \underline{v} + \underline{w}$, *then as* $t \to \infty$, *the relationship dissolves with probability* 1.

Proof. It can be seen that $(\underline{u}, \underline{v})$ is the only absorbing state of the stochastic process given by $(u, f(u))$, so the result follows from standard arguments in stochastic process. ■

The second significant change is in the dynamics of wages. When there is sufficient surplus in the relationship, recall that we show that the optimal relationship can be implemented by a sequence of stationary contracts once the agent receives permanent employment. On the other hand, there can be non-stationary contracts (in the stage of permanent employment) that are

optimal. And this indeterminacy of wage dynamics increases as the surplus in the relationship increases.

When the surplus is low, there is no flexibility of moving compensations around because acceleration or deferral of the bonuses will make the principal's reneging constraint harder to sustain. The following proposition gives an exact expression for the wage of the agent.

Proposition 7: *If* $py - \frac{[1-\delta(1-p)]c}{\delta(p-q)} < \underline{v} + \underline{w}$, *then the wage at the beginning of period* t *for an agent with expected payoff* u_t is given by

$$w_t = \underline{w} + \max\{u_t - u_e, 0\},$$

where u_e is determined in Theorem 2 by *(10).*

5.2 Low Wage Floor

In this subsection, we explore how the PPE payoff set changes when the wage floor $\underline{w}$ decreases and falls below $\underline{u}$ so that the limited liability constraint is less binding than the case studied in Section 3.

The analysis here becomes more complicated because the Pareto frontier may look qualitatively different from the one in Section 3. On the one hand, it remains true that the slope of the Pareto frontier is -1 to the right of a threshold u_e, determined by Lemma 2. On the other hand, it is no longer true that the Pareto frontier to the left of u_0 is a line segment joining $(\underline{u}, \underline{v})$ and $(u_0, f(u_0))$.

In particular, there are three cases to consider. First, when $\underline{w}$ is sufficiently low so that

$$\underline{w} - c + pk \leq \underline{u},$$

the limited liability has no bite because the principal can offer a base wage sufficiently low to extract all of the rand we can use stationary contract to provide incentive just as in Levin (2003). In this case, $f(\underline{u}) = py - c - \underline{u} \neq \underline{v}$.

Second, when the wage floor is not so small so that

$$\frac{\underline{u}}{\delta} \leq u_0 = (1-\delta)(\underline{w} - c) + \delta(\underline{u} + pk),$$

we can show that f remains a straight line segment in $[\underline{u}, u_0]$, and we have

$$f(\underline{u}) = \max\{\underline{v}, \frac{\underline{u} f(u_0)}{u_0}\}.$$

In this case, the basic structure of the PPE set remains and the basic structure of the employment dynamics remains.

Third, when $\underline{w}$ is in the intermediate range, we can no longer show that f is a straight line between $\underline{u}$ and u_0, and the analysis becomes considerably more difficult. It will be interesting to explore the property of the optimal relational contract for this range of wage floor.

While we cannot characterize the Pareto frontier to the left of u_0 for some ranges of $\underline{w}$, we can show that the functional equation that governs the region between the two thresholds (u_0 and u_e) is again the one in Lemma 6. In other words, as long as $\underline{w} - c + pk > \underline{u}$ so that the wage floor has a bite in the sense that the principal cannot lower the base wage enough to extract all of the surplus, we again have the basic intuition that the most efficient way to induce effort is to use the threat of termination before paying out bonus as reward. It follows that the optimal relational contract is also inefficient in this case. Moreover, the structure of employment relationship is similar to that in Section 3 in the sense that the agent starts the employment relationship in a probation phase and receives no bonus until his continuation payoff exceeds u_e.

5.3 Full Commitment

We have assumed in our model that the outputs are not contractible. In this subsection, we assume instead that the output is contractible and the principal is able to commit to the long-term contract offered.[13] The purpose of studying this case is to understand the role of commitment in affecting the efficiency of the employment relationship and the resulting wage dynamics.

Theorem 3 characterizes the Pareto frontier of the long-term contracts under full commitment.

Theorem 3*: When firms can commit to long-term contract, he Pareto frontier is given by the unique function that solves the following equation*

$$f(u) = \begin{cases} \underline{v} + \frac{u-\underline{u}}{u_0-\underline{u}}(f(u_0) - \underline{v}) & \text{if } u \in [\underline{u}, u_0] \\ (1-\delta)(py - \underline{w}) + \delta[pf(L(u)+k) + (1-p)f(L(u))] & \text{if } u \in [u_0, u_e] \\ f(u_e) + u_e - u & \text{if } u \in [u_e, u_e + f(u_e) - \underline{v}], \end{cases} \tag{11}$$

where $u_e = (\underline{w} - c) + \frac{\delta pk}{(1-\delta)}$, $f(u_e) = py - c - ((\underline{w} - c) + \frac{\delta pk}{(1-\delta)})$, *and* $u_0 = (1-\delta)(\underline{w} - c) + \delta(\underline{u} + pk)$.

Proof. When the agent's expected payoff is u_e, the long-term contract cannot do better than $py - c - u_e$. The rest of the theorem is proved in exactly the same way as Theorem 1. ■

[13] In other words, we do not require the contract to be renegotiation-proof. Moreover, the full-commitment here is one-sided: the principal can commit to the contract, and the agent may leave the relationship at any time.

Theorem 3 shows that when there is sufficient surplus in the relationship (so Assumption 1 holds), the Pareto frontier of long-term contracts with commitment is identical to that of relational contracts . In particular, this implies that the optimal long-term contract and the optimal relational contract is equally efficient. The literature on efficiency wages have often assumed that the firms can commit and justify this assumption by the reputation concern of the firms. This result confirms this intuition that when the surplus in the relationship is large enough, one may assume that firms can commit.

It follows that the results for optimal relational contract carry through for the optimal long-term contract under full commitment. In particular, workers start in the employment relationship in a probation phase, and depending on the outputs, he will either receive permanent employment or is terminated. One caveat is that even if the efficiency and the structure of the optimal contract with and without full commitment are the same here, there can be some differences in wage dynamics.

Specifically, once the agent receives permanent employment, there is little control over the ways the wages are paid out. For example, there is no limit on how much the principal can push back the payment to the agent, as long as the discounted expected payment remains the same. Such schemes may not be possible under relational contracts because the amount of bonus paid out by the principal cannot exceed that of the future surplus in the relationship.

This difference takes an more explicit form when the surplus of the relationship is small (so Assumption 1 fails). In this case, turnover dynamics is different as well. When surplus is small, Subsection 5.1 indicates that the optimal relational contract eventually terminates with probability 1. Moreover, the wage is completely determined. In contrast, Theorem 3 indicates that under the optimal contract with full commitment, there is positive probability that the agent receives permanent employment. And once permanent employment is obtained, there is again little control on how wages are paid out.

This difference in employment dynamics arises because under the optimal long-term contract with commitment, the principal in fact incurs a loss by staying in the relationship once the agent receives permanent employment. This cannot be part of the equilibrium in relational contracts. On the other hand, this ability to commit to such losses increases the ex ante payoff of the principal.

Infinitely-repeated principal agent problem as been studied by Spear and Srivastava (1987) (hereafter SS). SS implicitly assume that the optimal contract is renegotiation-proof, so the Pareto frontier is downward sloping. In addition, SS considers a more general environment in which the agent is risk averse and has a continuous level of efforts. The risk aversion of the agent creates a need for smoothing of wage across periods, and this makes the analysis significantly more difficult.

6 Conclusion and Discussion

This paper develops a tractable model of relational contracts of imperfect public monitoring with limited liability. The optimal relational contract highlights key features of the efficient use of termination and bonus to induce effort when a job has rent: bonus payment should be postponed as much as possible and termination always occurs with positive probability. The optimal relational contract generates a number of patterns of employment dynamics. First, the worker sometimes starts the employment relationship in a probation period. Second, wage is more sensitive to performance over time and compensation is deferred. Third, turnover rate may be inverse-U shaped with employment duration. Fourth, earlier successes lead to more favorable wage and turnover dynamics for the agent.

The tractability of the model enables us to study how the welfare of the firm, the worker, and the overall relationship is affected by exogenous conditions. While most of the comparative statics results are intuitive, our technique of deriving these results are of independent interest. Moreover, some of the comparative statics results shed new light on important policy issues. For example, while most of the debate on minimum wages focus on the total number of employment, our analysis indicates that minimum wage may harm workers who are already employed through an increase in involuntary turnover.

The tractability of the model also implies that some interesting margins of adjustment are not explored in this model. For example, if the agent can put in multiple levels of effort, then one may study how the agent's effort choice is affected by the history of outputs. With multiple effort levels, the basic lesson that bonus should be postponed as much as possible and that termination will occur with positive probability still remains valid. But it appears difficult to state something general about how effort changes.[14]

Another obvious extension of the model is to allow for multiple projects. Two results are obtained here. First, as in the static framework, limited liability may induce the principal to assign "safer" but less efficient projects to the agent. Second, multiple projects can generate richer employment dynamics depending on the properties of available projects. For example, if there is a safe project which is more efficient than the outside option but less efficient than a risky project, then the punishment to the agent may not be to terminate the relationship, but instead to be assigned to the safe project forever. This helps explain, for example, why the promotion probability of a worker decreases the longer he's at a job, yet he is not necessarily fired or demoted; see for example, Baker, Gibbs, Holmstrom (1994).

Finally, it may be interesting to embed this model into a general equilibrium framework, so that the outside options of the workers and firms are endogenized. Different from existing

[14] In models where the principal can commit, effort level may be nonmonotone in the agent's continuation value; see for example Spear and Srivastava (1987) and Clementi and Hopenhayn (2006).

efficiency wage models, firms and workers can separate on the equilibrium path. It follows that the equilibrium unemployment rate will depend on the stochastic nature of the production function (and thus the rate of involuntary turnover). In a related paper, Fong and Li (2008) examine a model in which the principal is able to replace the current worker immediately after he is fired. It is shown that the possibility of replacement preserves the basic structure and intuition of this model.

References

[1] Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990) "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring", *Econometrica*, 58(5), pp. 1041-1063.

[2] Akerlof, George and Lawrence Katz (1989) "Workers' Trust Funds and the Logic of Wage Profiles," *Quarterly Journal of Economics*, 104 (3) pp. 525-536.

[3] Albuquerque, Rui and Hugo Hopenhayn (2004) "Optimal Lending Contracts and Firm Dynamics", *Review of Economic Studies*, 71 (2), pp. 285–315.

[4] Biais, Bruno, Thomas Mariotti, Guillaume Plantin, and Jean-Charles Rochet (2007) "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications," *Review of Economic Studies*, 74(2), pp. 345-390.

[5] Baker, George, Robert Gibbons, and Kevin J. Murphy (1994) "Subjective Performance Measures in Optimal Incentive Contract", *Quarterly Journal of Economics*, 109 (4); pp. 1125-1156.

[6] ———, ———, and ——— (2002), "Relational Contracts and the Theory of the Firm", *Quarterly Journal of Economics*, 117 (1); pp. 39-84.

[7] ———, ———, and ——— (2006), "Contracting for Control", mimeo.

[8] Baker, G., M. Gibbs, and B. Holmstrom, (1994a) "The Internal Economics of the Firm: Evidence From Personnel Data," *Quarterly Journal of Economics,* 109 , pp. 881-919.

[9] ———, ———, and ———, (1994b) "The Wage Policy of a Firm," *Quarterly Journal of Economics*, 109, pp. 921-955.

[10] Bull, Clive (1987), "The Existence of Self-Enforcing Implicit Contracts", *Quarterly Journal of Economics*, 102 (1); pp. 147-159.

[11] Bull, Clive and Piero Tedeschi, (1989), "Optimal Probation for New Hires," *Journal of Institutional and Theoretical Economics,* 145(4), pp. 627-642.

[12] Clementi, Gina, and Hugo A Hopenhayn, (2006). "A Theory of Financing Constraints and Firm Dynamics," *Quarterly Journal of Economics*, 121(1), pp. 229-265,

[13] Chassang, Sylvain (2009). "Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts." Mimeo.

[14] Coughlan, Anne, Sanjog Misra and Chakravarthi Narasimhan, (2005) "Salesforce Compensation: An Analytical and Empirical Examination of the Agency Theoretic Approach," *Quantitative Marketing and Economics*, Vol. 3, pp. 5-39.

[15] DeMarzo Peter, and Mike Fishman (2007), "Optimal Long-Term Financial Contracting with Privately Observed Cash Flows" *Review of Financial Studies*, 20 (6) pp. 2079-2128.

[16] DeMarzo Peter and Yuliy Sannikov (2006), "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *Journal of Finance*, 61, pp. 2681–2724.

[17] Doornik, Katherine (2006), "Relational Contracting in Partnerships," *Journal of Economics and Management Strategy*, 15(2); pp. 517-548.

[18] Fong and Li (2008) "A Theory of Endogenous Turnover," Mimeo.

[19] Halac, Marina (2008) "Relational Contracts and the Value of Relationships," Mimeo.

[20] Hall, Robert (1982), "The Importance of Lifetime Jobs in the U.S. Economy," *American Economic Review*, 72(4), pp. 716-724.

[21] Harris, Milton, and Bengt Holmstrom (1982), "A Theory of Wage Dynamics," *Review of Economic Studies*, 49 (3), pp. 315-333.

[22] Innes, Robert (1990), "Limited Liability and Incentive Contracting with Ex Ante Action Choices." Journal of Economic Theory, 52, pp. 45-67

[23] Jovanovic, Boyan (1979), "Firm-specific capital and Turnover." *Journal of Political Economy,* Vol 87 (6) pp. 1246-1260.

[24] Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian (1978), "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process", *The Journal of Law and Economics*, 21 (2), pp. 297-326.

[25] ——— and Keith B. Leffler (1981) "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89 (4); pp. 615–641.

[26] Lazear, Edward (1981), "Agency, Earning Profiles, Productivity and Hour Restrictions," *American Economic Review*, 71; pp. 606-620.

[27] Levin, Jonathan (2003), "Relational Incentive Contracts", *American Economic Review*, 93 (3); pp. 835-57.

[28] Lin, Ming-Jen (2005), "Opening the Black Box: the Internal Labor Markets of Company X", *Industrial Relations*, 44(4), pp. 659-705.

[29] Loh, Eng Seng, (1994a), "The Determinants of Employment Probation Lengths", *Industrial Relations*, 33(3), 386-406.

[30] ——— , (1994b), "Employment Probation as a Sorting Mechanism", *Industrial and Labor Relations Review,* 47(3), 471-486.

[31] Macaulay, Stewart (1963). "Non-contractual relations in business: A preliminary study." American Sociological Review 28(1): pp. 55-67.

[32] MacLeod W. Bentley and James M. Malcomson, (1988), "Reputation and Hierarchy in Dynamic Models of Employment," *Journal of Political Economy*, 96 (4), pp. 832-854.

[33] ——— and ——— , (1989), "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment", *Econometrica*, 57 (2); pp. 447-480.

[34] ——— and ——— ,(1998), "Motivation and Markets", *American Economic Review*; 88 (3); pp. 388-411.

[35] MacNeil, Ian (1978). "Contracts: Adjustments of long-term economic relations under classical, neoclassical, and relational contract law." Northwestern University Law Review 72: 854-906.

[36] Meyer, Magaret (1991), "Learning from Coarse Information: Biased Contests and Career Profiles" *Review of Economic Studies*, 58, pp.15-41.

[37] Prendergast, Canice (1993), "The Role of Promotion in Inducing Specific Human Capital Acquisition." *Quarterly Journal of Economics*, 108 (2), pp. 523-534

[38] Ray, Debraj (2002), "The time structure of self-enforcing agreements," *Econometrica;* 70(2): pp. 547-582.

[39] Rayo, Luis (2007), "Relational incentives and moral hazard in teams." *Review of Economic Studies*, 74(3): 937-963.

[40] Riphahn, Regina, and Anja Thalmaier (2001) "Behavioral Effects of Probation Periods: An Analysis of Worker Absenteeism", Jahrbücher für Nationalökonomie und Statistik (Journal of Economics and Statistics), 221(2), pp. 179-201.

[41] Rogerson, William (1985), "Repeated Moral Hazard," *Econometrica,* 53(1), pp. 69-76.

[42] Sappington, David (1983) "Limited Liability Contract Between the Principal and Agent," *Journal of Economic Theory*, 29(1) pp. 1-21.

[43] Shapiro,Carl and Joe Stiglitz (1984) "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74(3), pp. 433-444.

[44] Spear, Stephen, and Sanjay Srivastava, (1987) "On Repeated Moral Hazard with Discounting," *Review of Economic Studies*, 54(4), pp. 599-617.

[45] Stevens, Ann (2005) "“The More Things Change, the More they Stay the Same: Trends in Lifetime Employment 1969-1998." NBER working paper No. W11878.

[46] Thomas, Jonathan, and Tim Worrall (1988), "Self-enforcing Wage Contracts," *Review of Economic Studies*, 55(4): pp. 541-554.

[47] Thomas, Jonathan, and Tim Worrall (2007), "Dynamic Relational Contracts under Limited Liability," Mimeo.

[48] Ureta, Manuelita (1992), "The Importance of Lifetime Jobs in the U.S. Economy," *American Economic Review*, 82 (1), pp. 322-335.

[49] Wang, Ruqu and Andrew Weiss, (1998), "Probation, Layoffs, and Wage-Tenure Profiles: A Sorting Explanation", *Labour Economics*, 5(3), pp. 359-383.

[50] Weiss, Andrew and Ruqu Wang, (1990), "A Sorting Model of Labor Contracts: Implications for Layoffs and Wage-Tenure Profiles," NBER Working Paper No. 3448.

[51] Williamson, Oliver E. (1975), *Markets and Hierarchies*, Free Press, New York.

[52] Yang, Huanxin (2005), "Nonstationary Relational Contracts", mimeo.

7 Appendix

The appendix has two parts. The first part characterizes the Pareto frontier. The second part collects the proofs for comparative statics.

7.1 Characterize the Pareto Frontier

We follow six steps. First, we show that there exists a threshold u_e such that f has a slope of -1 to the right of u_e (Lemma 1). Second, we determine the exact value of u_e and $f(u_e)$ (Lemma 2). Third, we show that there exists a threshold u_0 (to be defined below) such that to the left of u_0, f is a straight line between $(\underline{u}, \underline{v})$ and $(u_0, f(u_0))$ (Lemma 4). Fourth, we determine the value of u_0 (Lemma 5). Fifth, we show that the value of f between u_0 and u_e is the fixed point of a contraction mapping indexed by the value of $f(u_0)$ (Lemma 6). Finally, we show that $f(u_0)$ can be found as a fixed point to a monotone contraction mapping.

Denote u_e as the smallest PPE payoff of the agent that maximizes the sum of the payoff of the principal and agent. First, we show that the Pareto frontier of the PPE past u_e has a slope of -1.

Lemma 1: *For* $u \in [u_e, u_e + f(u_e) - \underline{v}]$,

$$f(u) = f(u_e) + u_e - u.$$

Proof. Suppose in the PPE that achieves $(u_e, f(u_e))$, the principal offers w_1 in period 1. Now consider a different strategy that follows the previous equilibrium except in period 1 the principal offers $w_1 + u - u_e$. If the agent rejects the offer in period 1 (or the principal fails to make this offer in period 1), the principal and the agent will believe that and no effort will be put in. It is then easy to check that this strategy profile is a PPE. Moreover, this PPE achieves a payoff of $(u, f(u_e) + u_e - u)$. Therefore, $f(u) \geq f(u_e) + u_e - u$.

On the other hand, if $f(u) > f(u_e) + u_e - u$, this violates the the definition that u_e maximizes the sum of the principal and the agent's payoff. ■

We next determine the exact value of $(u_e, f(u_e))$. In the efficient outcome of the game, effort is put in each period. For the agent to have incentive to put in effort in any period, we must have

$$(1-\delta)(-c) + \delta((1-p)u_L + pu_H) \geq \delta((1-q)u_L + qu_H),$$

where u_H corresponds to the agent's continuation payoff after a good outcome and u_L corresponds to the agent's continuation payoff after a bad outcome.

The next lemma shows that the smallest agent's payoff in an efficient equilibrium satisfies

$$u_e = L(u_e).$$

Lemma 2:

$$u_e = (\underline{w} - c) + \frac{\delta pk}{(1-\delta)} = L(u_e).$$

Proof. Consider the following strategy: on the equilibrium path: the principal offers the agent $\underline{w}$ in period 1. The agent accepts and puts in effort. In all future periods, the principal offers the agent $\underline{w} + k$ if the previous outcome is $Y = y$ and offers $\underline{w}$ otherwise. Off the equilibrium path: the agent never puts in effort and the agent never offers a contract to the agent. By Assumption1, this strategy can be shown to be a PPE, and it achieves a payoff of $((\underline{w} - c) + \frac{\delta pk}{(1-\delta)}, py - c - ((\underline{w} - c) + \frac{\delta pk}{(1-\delta)}))$. Therefore, $u_e \leq (\underline{w} - c) + \frac{\delta pk}{(1-\delta)}$.

Now if $u_e < (\underline{w} - c) + \frac{\delta pk}{(1-\delta)}$, this implies that $L(u_e) < u_e$. Since $f(u_e) + u_e = py - c$, this implies that $L(u_e) + f(L(u_e) = py - c$ as well. But by the definition of u_e, we know that $L(u_e) + f(L(u_e) < u_e + f(u_e) = py - c$, so this is a contradiction. ■

Next, we characterize f to the left of u_e. The key observation is that the PPE payoff is convex, so f must be concave. This leads to the next lemma, which shows that in any PPE payoff that obtains the Pareto frontier with $u \in (\underline{u}, u_e)$, the first period wage must be the minimum wage.

Lemma 3: *For any PPE that reaches $(u, f(u))$ with $u \in (\underline{u}, u_e)$, the first period wage of the agent must be $w_1 = \underline{w}$.*

Proof. If $\underline{u} = u_e$, then the only equilibrium is that both parties take their outside options and there is nothing to prove.

Now let $\underline{u} < u_e$. Take $u \in (\underline{u}, u_e)$. Suppose in a PPE profile that obtains $(u, f(u))$, the first period wage $w_1 > \underline{w}$. Now adapt the equilibrium by lowering the first period wage to $\underline{w}$. It remains an equilibrium that the agent accepts the period 1 wage and continue the play with the previous equilibrium. (If the agent rejects the period 1 wage, the principal never offers the contract again). This new PPE achieves a payoff of $(u-(1-\delta)(w_1-\underline{w}), f(u)+(1-\delta)(w_1-\underline{w}))$.

Therefore, we have

$$f(u-(1-\delta)(w_1-\underline{w})) \geq f(u)+(1-\delta)(w_1-\underline{w}).$$

But this says that the slope of f at u is weakly smaller than -1, yet the slope of f at u_e is -1. Now if the slope of f at u is strictly smaller than -1, this violates the concavity of f. If the slope of f at u is equal to -1, this implies that $u + f(u) = u_e + f(u_e)$, and since $u < u_e$, this contradicts the definition of u_e. ■

Now define u_0 as the smallest u in which $(u, f(u))$ is obtained by requiring the agent to put in effort in period 1. The next lemma shows that $f(\underline{u}) = \underline{v}$ and f is a straight line between $(\underline{u}, \underline{v})$ and $(u_0, f(u_0))$.

Lemma 4*: For* $u \in [\underline{u}, u_0]$,

$$f(u) = \underline{v} + \frac{u-\underline{u}}{u_0-\underline{u}}(f(u_0)-\underline{v}).$$

Proof. It is clear that $(\underline{u}, \underline{v})$ is a PPE payoff, where on the equilibrium path the principal never offers the agent a contract, and off the equilibrium path, the agent never puts in effort. By Assumption 2, we have $\underline{w} \geq \underline{u}$, and the assumption that $qy < \underline{v}$ implies that we must have $\underline{v} = f(\underline{u})$.

The convexity of PPE payoff immediately imply that $f(u) \geq \underline{v} + \frac{u-\underline{u}}{u_0-\underline{u}} f(u_0)$. Now suppose the inequality is strict, there are two cases to consider. First, suppose the equilibrium payoff $(u, f(u))$ is reached by a combination of $(\underline{u}, \underline{v})$ and $(u', f(u'))$ for some $u' \geq u_0$. In this case, the weak concavity of f implies that $(u, f(u))$ cannot lie strictly above the segment formed by $(\underline{u}, \underline{v})$ and $(u_0, f(u_0))$.

In the second case, the equilibrium payoff is reached by a pure play. Now let u be the largest payoff for the agent such that $f(u) > \underline{v} + \frac{u-\underline{u}}{u_0-\underline{u}} f(u_0)$.[15] From Lemma 3 and the definition of u_0, we know that the first period play payoff is given by either $(\underline{u}, \underline{v})$ or $(\underline{w}, qy)$. In either case, it is clear that $(u_0, f(u_0))$ lies strictly below the linear combination of $(u, f(u))$ and its continuation payoff. This is a contradiction. ■

[15] If no such points exist, take one close enough to the limsup.

The next lemma gives the exactly value of u_0. In particular, we have

$$L(u_0) = \underline{u}.$$

Lemma 5:

$$u_0 = (1-\delta)(\underline{w} - c) + \delta(\underline{u} + pk)$$

Proof. It is clear $L(u_0) \geq \underline{u}$, which is the agent's maxmin payoff. Now if $L(u_0) > \underline{u}$, we argue that there exists a PPE payoff that gives the agent the payoff of $u_0 - \varepsilon$ and the principal a payoff that lies on the line segment between $(\underline{u}, \underline{v})$ and $(u_0, f(u_0))$, and this violates the definition of u_0. In particular, let s be the slope between $(\underline{u}, \underline{v})$ and $(u_0, f(u_0))$. Then by the weak concavity of f, we know that both $(L(u_0) - \varepsilon, f(L(u_0)) - s\varepsilon)$ and $(L(u_0) + k - \varepsilon, f(L(u_0) + k) - s\varepsilon)$ are PPE payoffs. And strategy profile that pays the agent $w_1 = \underline{w}$, requires the agent to put in effort in period 1, and promise the agent with the above two continuation payoffs (given the output as y or 0), will be an equilibrium that gives payoff of $(u_0 - \varepsilon, f(u_0) - s\varepsilon)$. This proves that $L(u_0) = \underline{u}$. ■

Note that we have determined the shape of f to the left of u_0 and to the right of u_0, we are ready to determine the value of f between these two points. The next lemma gives such a functional equation.

Lemma 6: *For* $u \in [u_0, u_e]$,

$$f(u) = (1-\delta)(py - \underline{w}) + \delta[pf(L(u) + k) + (1-p)f(L(u))]. \tag{12}$$

Proof. This follows in two steps. The first step shows that for $u \in [u_0, u_e]$, $(u, f(u))$ can be obtained by an equilibrium profile in which the first period play requires effort. Now suppose the contrary. Let u_l be the largest point such that $(u_l, f(u_l))$ lies on the line given by $(\underline{u}, \underline{v})$ and $(u_0, f(u_0))$. Suppose $u \in [u_l, u_e]$. There are two cases to consider.

In the first case, $(u, f(u))$ can be reached by a pure play in period 1. In this case, by Lemma 3, the first period play payoff is given by either $(\underline{u}, \underline{v})$ or $(\underline{w}, qy)$. In either case, the slope between $(u, f(u))$ and its continuation payoff will exceed the slope between $(\underline{u}, \underline{v})$ and $(u_0, f(u_0))$. This violates the concavity of f.

In the second case, the $(u, f(u))$ is reached through a mix. Since $(u, f(u))$ lives in a two dimension space and f is concave, we may assume

$$(u, f(u)) = p_1((u_1, f(u_1)) + (1-p_1)((u_2, f(u_2))$$

for some p_1, u_1,and u_2, where $(u_1, f(u_1))$ and $(u_2, f(u_2))$ are reached through pure play in period 1. If $(u_i, f(u_i))$, $i = 1, 2$ has first period play that doesn't require effort, the previous paragraph

implies that $u_i \notin [u_l, u_e]$. But then the concavity of f implies that $(u, f(u))$ can be obtained by linear combination of points with the agent's payoff belonging to $[u_l, u_e]$.

To finish the first step, if $u \in (u_0, u_l)$, then $(u, f(u))$ can be achieved by mixing $(u_0, f(u_0))$ and $(u_l, f(u_l))$. And since both $(u_0, f(u_0))$ and $(u_l, f(u_l))$ can be achieved by requiring effort in period 1, so can $(u, f(u))$.

In the second step, we note that the equation follows because to achieve the maximum payoff for the principal, a) the continuation payoff must lie on the Pareto frontier, and b) the distance of payoff between the good and bad outcomes for the agent needs not to exceed k by the concavity of f. ■

Lemma 6 motivates us to define the following operator. Let g be a bounded function on $[u_0, u_e]$. We define T_Z as follows:

$$\begin{aligned} T_Z g(u) = (1-\delta)(py - \underline{w}) \\ + \delta p(1_{\{L(u)+k<u_e\}} g(L(u)+k) + 1_{\{L(u)+k\geq u_e\}}(py - c - (L(u)+k)) \\ + \delta(1-p)(1_{\{L(u)<u_0\}}(\underline{v} + \frac{L(u)-\underline{u}}{u_0 - \underline{u}}(Z - \underline{v}) + 1_{\{L(u)\geq u_0\}} g(L(u))). \end{aligned}$$

Lemma 7:

$$0 \leq \frac{dg_Z(u_0)}{dZ} < 1.$$

Proof. There are two cases to consider. First, suppose $L(u_0) + k \geq u_e$. In this case, the result is immediate because

$$g_Z(u_0) = (1-\delta)(py - \underline{w}) + \delta((1-p)\underline{v} + pf(L(u_0)+k))$$

And since when $L(u_0) + k \geq u_e$, $f(L(u_0)+k)$ is independent of Z, we have

$$\frac{dg_Z(u_0)}{dZ} = 0.$$

Now suppose $L(u_0) + k \geq u_e$. Take the unique fixed point of Tz, g_Z, and note that

$$T_{Z+\varepsilon}(g_Z + \varepsilon) \leq T_Z g_Z + \delta\varepsilon \leq T_Z g_Z + \varepsilon = g_Z + \varepsilon.$$

Since T is monotone (i.e. $Tg_1 \leq Tg_2$ if $g_1 \leq g_2$ where $g_1 \leq g_2$ iff $g_1(u) \leq g_2(u)$ for all u) , we have

$$T^2_{Z+\varepsilon}(g_Z + \varepsilon) \leq T^1_{Z+\varepsilon}(g_Z + \varepsilon),$$

and more generally

$$T_{Z+\varepsilon}^{n}(g_Z+\varepsilon) \leq T_{Z+\varepsilon}^{n-1}(g_Z+\varepsilon) \leq ... \leq g_Z+\varepsilon.$$

Since $g_{Z+\varepsilon} = \lim_{n\to\infty} T_{Z+\varepsilon}^{n}(g_Z+\varepsilon)$, we have

$$g_{Z+\varepsilon} \leq g_Z+\varepsilon.$$

In particular, this implies that

$$g_{Z+\varepsilon}(L(u_0)+k) \leq g_Z(L(u_0)+k)+\varepsilon.$$

Finally, we have

$$g_Z(u_0) = (1-\delta)py + \delta((1-p)x + pg_Z(L(u_0)+k).$$

$$g_{Z+\varepsilon}(u_0) = (1-\delta)py + \delta((1-p)x + pg_{Z+\varepsilon}(L(u_0)+k).$$

Therefore,

$$g_{Z+\varepsilon}(u_0) - g_Z(u_0) = \delta p\left[g_{Z+\varepsilon}(L(u_0)+k) - g_Z(L(u_0)+k)\right] \leq \delta p\varepsilon$$

and thus

$$\frac{dg_Z(u_0)}{dZ} = \lim_{\varepsilon\to\infty} \frac{g_{Z+\varepsilon}(u_0) - g_Z(u_0)}{\varepsilon} \leq \delta p < 1.$$

On the other hand, it can be checked that

$$T_{Z+\varepsilon}(g_Z) \geq g_Z,$$

so a similar reasoning as above gives that

$$\frac{dg_Z(u_0)}{dZ} \geq 0.$$

■

Lemma 7 immediately implies that there will be at most one value $f(u_0)$ that satisfies (7). The existence of such value is guaranteed by the fact that f is properly defined and it must satisfy (7).

An more direct approach to prove the existence is as follows. In the proof of Lemma 7, we know that g_Z is weakly increasing in Z. In addition, it can be shown that the g_Z induced by T_Z is nonexpansive in the sense that

$$||g_{Z+\varepsilon} - g_Z|| \leq \varepsilon.$$

It can also be shown that a weakly increasing non-expansive map has a unique fixed point

(see the proof of Lemma 9 for a formal proof), so there is a unique point Z^* such that

$$Z^* = g_{Z^*}(u_0).$$

And we have

$$f(u_0) = Z^*.$$

Corollary 2: *If* $\underline{u} + \frac{(1-\delta-\delta q)}{\delta(p-q)}c \geq \underline{w}$, *then*

$$f(u) = \begin{cases} \underline{v} + s_0(u - \underline{u}) & \text{if } u \in [\underline{u}, u_0] \\ \underline{v} + \frac{(u_0-\underline{u})}{1-\delta(1-p)}(s_0 - \frac{p\delta(1-\delta^{n+1})}{1-\delta} - \delta^{n+1}s_n(1-p)) + s_{n+1}(u - u_n) & \text{if } u \in [u_n, u_{n+1}] \\ f(u_e) + u_e - u & \text{if } u \in [u_e, u_e + f(u_e) - \underline{v}], \end{cases}$$

where $u_0 = (1-\delta)(\underline{w} - c) + \delta(\underline{u} + pk)$, $s_0 = \frac{(1-\delta)(py-\underline{w})+\delta((1-p)\underline{v}+p(py-c-(\underline{u}+k)))-\underline{v}}{(1-\delta)(\underline{w}-c)+\delta(\underline{u}+pk)-\underline{u}}$, $u_n = u_0 + \frac{\delta(1-\delta^n)}{1-\delta}(u_0 - \underline{u})$, $s_n = s_0 - (1 + s_0)(p + (1 - (1-p)^{n+1}))$, $u_e = (\underline{w} - c) + \frac{\delta pk}{(1-\delta)}$, $f(u_e) = py - c - ((\underline{w} - c) + \frac{\delta pk}{(1-\delta)})$.

Proof. First, define

$$s_0 = \frac{f(u_0) - \underline{v}}{u_0 - \underline{u}}.$$

Then by the equation on the derivative of f, we have

$$f'(u) = -p + (1-p)s_0 \quad \text{for } u \in (u_0, u_1).$$

Define s_{n+1} as the slope of f in (u_n, u_{n+1}) : then we have

$$\begin{aligned} s_1 &= -p + (1-p)s_0; \\ s_{n+1} &= -p + (1-p)s_n. \end{aligned}$$

Now note that

$$s_{n+1} - s_n = (1-p)(s_{n+1} - s_n),$$

so

$$s_n = s_0 - (1 + s_0)(p + (1 - (1-p)^{n+1})).$$

Now note that for $u \in [u_n, u_{n+1})$,

$$\begin{aligned} f(u) &= f(u_0) + \sum_{j=1}^{n} (u_j - u_{j-1}) s_j + s_{n+1}(u - u_n) \\ &= \underline{v} + s_0(u_0 - \underline{u}) + \sum_{j=1}^{n} \delta^j (u_0 - \underline{u}) s_j + s_{n+1}(u - u_n) \\ &= \underline{v} + (u_0 - \underline{u}) \sum_{j=0}^{n} \delta^j s_j + s_{n+1}(u - u_n). \end{aligned}$$

The summation $\sum_{j=0}^{n} \delta^j s_j$ can be evaluated explicitly. Define

$$S(n) = \sum_{j=0}^{n} \delta^j s_j.$$

Then

$$\begin{aligned} (1 - \delta(1-p)) S(n) &= s_0 + \sum_{j=1}^{n} \delta^j (s_j - (1-p) s_{j-1}) - \delta^{n+1} s_n (1-p) \\ &= s_0 - \frac{p\delta(1 - \delta^{n+1})}{1 - \delta} - \delta^{n+1} s_n (1-p). \end{aligned}$$

Substituting this back into the expression for $f(u)$, we have

$$\begin{aligned} f(u) &= \underline{v} + (u_0 - \underline{u}) \sum_{j=0}^{n} \delta^j s_j + s_{n+1}(u - u_n). \\ &= \underline{v} + \frac{(u_0 - \underline{u})}{1 - \delta(1-p)} (s_0 - \frac{p\delta(1 - \delta^{n+1})}{1 - \delta} - \delta^{n+1} s_n (1-p)) + s_{n+1}(u - u_n). \end{aligned}$$

Since $s_n = s_0 - (1 + s_0)(p + (1 - (1-p)^{n+1}))$, the expression above determines f completely as long as we know what s_0 is. Note that

$$\begin{aligned} \underline{v} + s_0(u_0 - \underline{u}) &= f(u_0) \\ &= (1 - \delta)(py - \underline{w}) + \delta((1-p)v + p(py - c - (\underline{u} + k))), \end{aligned}$$

where the second inequality uses Lemma 6. This implies that[16]

$$s_0 = \frac{(1-\delta)(py-\underline{w}) + \delta((1-p)\underline{v} + p(py - c - (\underline{u}+k))) - \underline{v}}{(1-\delta)(\underline{w}-c) + \delta(\underline{u}+pk) - \underline{u}}.$$

■

7.2 Comparative Statics

In this subsection, we derive comparative statics by exploiting properties of functional operators. The first lemma shows that, in the range of agent's payoffs where efforts are called for, the PPE frontier shrinks continuously as the agent's outside option improves.

Lemma 8: *If* $\max_u\{F(u,\underline{u},\underline{v},\underline{w})\} > F(u_0(\underline{u}),\underline{u},\underline{v},\underline{w})$, *then all* $u \in [u_0(\underline{u}), u_e)$

$$-s < \frac{\partial F(u,\underline{u},\underline{v},\underline{w})}{\partial \underline{u}} < 0,$$

where s *is the left derivative (with respect to* u*) of* $F(u_0(\underline{u}),\underline{u},\underline{v},\underline{w})$.

Proof. Suppose we increase the agent's outside option from $\underline{u}$ to $\underline{u}+\varepsilon$. To simplify notation, suppose f_1 is the PPE Pareto frontier with respect to $\underline{u}$ defined on $[u_0(\underline{u}+\varepsilon),\ u_e]$. And let f_2 be the PPE Pareto frontier with respect to $\underline{u}+\varepsilon$ defined on $[u_0(\underline{u}+\varepsilon),\ u_e]$.

We know that $u_0(\underline{u}+\varepsilon) = u_0(\underline{u}) + \delta\varepsilon$. We also know that $u^A(\underline{u}+\varepsilon) \in (u_0(\underline{u}+\varepsilon), u_e)$. Now take a line segment between $(\underline{u}+\varepsilon, \underline{v})$ and $(u_0(\underline{u}+\varepsilon), f_1(u_0(\underline{u}+\varepsilon)))$. This line segment lies strictly below f_1 (except at the right end point where the two are equal.).

Now define an operator T_1 on functions on $[u_0(\underline{u}+\varepsilon), u_e]$ as follows.

$$\begin{aligned} T_1 g(u) \quad = \quad & (1-\delta)(py-\underline{w}) \\ & +\delta p(1_{\{L(u)+k<u_e\}} g(L(u)+k) + 1_{\{L(u)+k\geq u_e\}}(py - c - (L(u)+k)) \\ & +\delta(1-p)(1_{\{L(u)\geq u_0(\underline{u}+\varepsilon)\}} g(L(u) + \\ & 1_{\{L(u)<u_0(\underline{u}+\varepsilon)\}}(\underline{v} + \frac{L(u) - (\underline{u}+\varepsilon)}{u_0(\underline{u}+\varepsilon)) - (\underline{u}+\varepsilon)}(f_1(u_0(\underline{u}+\varepsilon) - \underline{v})) \end{aligned}$$

[16] An alternative method of calculating s_0 is to note that

$$f(u_e) = f(u_0) + \sum_{n=1}^{\infty} s_n(u_n - u_{n-1}) = \underline{v} + s_0(u_0 - \underline{u}) + \frac{\delta u_0(-p + (1-p)s) - \frac{p\delta^2 u_0}{1-\delta}}{1-(1-p)\delta}.$$

And in addition,

$$f(u_e) = py - c - ((\underline{w} - c) + \frac{\delta pk}{(1-\delta)}).$$

Some algebra shows that this gives the same s_0 as before.

Note that this operator is very similar to the operator T in Lemma 7, except the "straight line on the left" is smaller than that in T. It follows that

$$T_1 g \leq Tg \text{ for all } g.$$

In addition, T_1 can be checked to be monotone.

Define g_1 on $[u_0(\underline{u}+\varepsilon), u_e]$ by $g_1(u) = f_1(u)$. Then it is clear that

$$T_1 g_1(u) \leq Tg_1(u) \leq g_1(u).$$

Since T is monotone, it follows that

$$g_1(u) \geq g_{T_1}(u),$$

where g_{T_1} is the fixed point of T_1. From Lemma 7, this implies that $f_2(u_0(\underline{u}+\varepsilon)) = g_{T_1}(u_0(\underline{u}+\varepsilon)) < g_1(u_0(\underline{u}+\varepsilon)) = f_1(u_0(\underline{u}+\varepsilon))$.

Now define the operator T_2 on functions on $[u_0(\underline{u}+\varepsilon), u_e]$ as follows.

$$\begin{aligned} T_2 g(u) &= (1-\delta)(py - \underline{w}) \\ &\quad +\delta p(1_{\{L(u)+k<u_e\}} g(L(u)+k) + 1_{\{L(u)+k\geq u_e\}}(py - c - (L(u)+k)) \\ &\quad +\delta(1-p)(1_{\{L(u)\geq u_0(\underline{u}+\varepsilon)\}} g(L(u) \\ &\quad +1_{\{L(u)<u_0(\underline{u}+\varepsilon)\}}(\underline{v} + \frac{L(u)-(\underline{u}+\varepsilon)}{u_0(\underline{u}+\varepsilon))-(\underline{u}+\varepsilon)}(f_2(u_0(\underline{u}+\varepsilon)-\underline{v})) \end{aligned}$$

Define g_2 on $[u_0(\underline{u}+\varepsilon), u_e]$ by $g_2(u) = f_2(u)$. We know that g_2 is a fixed point of T_2. On the other hand, we can check that $T_2 g_1(u) \geq g_1(u)$. It immediately follows (as in Lemma 7) that $g_1(u) \geq g_2(u)$ for all u. Moreover, we note that $g_1(u) > g_2(u)$ for some neighborhood of $u_0(\underline{u}+\varepsilon)$. And since all of $u < u_e$ reaches the the neighborhood of with positive probability, so we have

$$f_2(u) = g_2(u) < g_1(u) = f_1(u) \text{ for all } u \in [u_0(\underline{u}+\varepsilon), u_e).$$

And $f_2(u_e) = f_1(u_e)$. This proves the first part of the lemma.

For the second part, suppose the slope between $(\underline{u}, \underline{v})$ and $(u_0(\underline{u}), f_1(u_0(\underline{u}))$ is s, where $f_1(u_0(\underline{u}))$ is the maximal PPE payoff of the principal at $u_0(\underline{u})$ assuming that the outside option is $\underline{u}$. Take a small ε, construct a line through $(\underline{u}+\varepsilon, \underline{v})$ with slope s. For small enough ε, we can show that this line lies strictly below f_1 from $\underline{u}$ till $u_0(\underline{u}+\varepsilon)$. This is because $\max_u\{F(u, \underline{u}, \underline{v}, \underline{w})\} > F(u_0(\underline{u}), \underline{u}, \underline{v}, \underline{w})$, the right derivative of f_1 at $u_0(\underline{u})$ is strictly positive, so there exists a small ε_1 such that $f_1(u_0(\underline{u})+\varepsilon_1) \geq f_1(u_0(\underline{u}))$. Now take $\varepsilon = \varepsilon_1/\delta$ will work.)

Let $d = s\varepsilon$. Define an operator T_3 on functions defined on $[u_0(\underline{u}+\varepsilon), u_e]$, such that the "straight line on the left" is given by the line through $(\underline{u}+\varepsilon, \underline{v})$ with slope s. Define g_1 on

$[u_0(\underline{u}+\varepsilon), u_e]$ by $g_1(u) = f_1(u)$.

$$\begin{aligned} & T_3(g_1(u)-d) \\ = \; & (1-\delta)py \\ & +\delta p(1_{\{L(u)+k<u_e\}}(g_1(L(u)+k)-d) + 1_{\{L(u)+k\geq u_e\}}(py-c-(L(u)+k)-d)) \\ & +\delta(1-p)(1_{\{L(u)\geq u_0(\underline{u}+\varepsilon)\}}(g_1(L(u)-d) + 1_{\{L(u)<u_0(\underline{u}+\varepsilon)\}}(\underline{v}+s(L(u)-\underline{u}-\varepsilon)). \end{aligned}$$

For small enough ε, we see that

$$T_3(g_1(u)-d) \simeq g_1(u) - \delta d \geq g_1(u) - d.$$

Now by the uniqueness proof (and that the slope of f_1 is smaller than s for $u > u_0(\underline{u}+\varepsilon)$), we see that

$$f_2(u_0(\underline{u}+\varepsilon)) > f_1(u_0(\underline{u}+\varepsilon)) - d.$$

Finally, follow the proof procedure in the first part of the theorem, we see immediately that

$$f_2(u) > f_1(u) - d \text{ for all } u \in [u_0(\underline{u}+\varepsilon), u_e].$$

■

The next lemma shows that, while the value of the PPE set shrinks for $u \in [u_0(\underline{u}), u_e]$, its slope increases.

Lemma 9: *If* $\max_u\{F(u,\underline{u},\underline{v},\underline{w})\} > F(u_0(\underline{u}),\underline{u},\underline{v},\underline{w})$, *then for almost all* $u \in [u_0(\underline{u}), u_e]$

$$\frac{\partial^2 F(u,\underline{u},\underline{v},\underline{w})}{\partial \underline{u}\partial u} > 0.$$

Proof. Again to simplify notation, suppose f_1 is the PPE Pareto frontier with respect to $\underline{u}$ defined on $[u_0(\underline{u}+\varepsilon),\ u_e]$. And let f_2 be the PPE Pareto frontier with respect to $\underline{u}+\varepsilon$ defined on $[u_0(\underline{u}+\varepsilon),\ u_e]$. Now define $h_1 = f_1'$; and $h_2 = f_2'$. When f_1' and f_2' are not properly defined, we use the right limit.

Lemma 8 implies that

$$h_2 > h_1 \quad \text{for } u < u_0(\underline{u}+\varepsilon).$$

Now define

$$\begin{aligned} \widetilde{T}(h) \; = \; & p(1_{\{L(u)+k<u_e\}}(h) - 1_{\{L(u)+k\geq u_e\}}) \\ & +(1-p)(1_{\{L(u)\geq u_0(\underline{u}+\varepsilon)\}}h(L(u)) + s_1 1_{\{L(u)<u_0(\underline{u}+\varepsilon)\}}), \end{aligned}$$

where s_1 is the slope between $(\underline{u}+\varepsilon, \max\{x-(\underline{u}+\varepsilon),0\})$ and $(u_0(\underline{u}+\varepsilon), f_2(u_0(\underline{u}+\varepsilon))$. It is easy to see that

$$\widetilde{T}(h_1) \geq h_1.$$

Let us define that

$$h^* = \lim_n \widetilde{T}^n(h_1).$$

Since the operator is monotone,

$$\widetilde{T}(h^*) = \widetilde{T}(\lim_n \widetilde{T}^n(h_1)) \geq \widetilde{T}(\widetilde{T}^n(h_1)) \quad \text{for all } n,$$

so we have $\widetilde{T}(h^*) \geq n$.

On the other hand, take any u, we know that

$$\begin{aligned}
\widetilde{T}(h^*(u)) &= p(1_{\{L(u)+k<u_e\}}(h^*(u)) - 1_{\{L(u)+k\geq u_e\}}) \\
&\quad +(1-p)(1_{\{L(u)\geq u_0(\underline{u}+\varepsilon)\}}h^*(L(u)) + s_1 1_{\{L(u)<u_0(\underline{u}+\varepsilon)\}}) \\
&\leq p(1_{\{L(u)+k<u_e\}}(\widetilde{T}^n(h_1(u)) + \varepsilon) - 1_{\{L(u)+k\geq u_e\}}) \\
&\quad +(1-p)(1_{\{L(u)\geq u_0(\underline{u}+\varepsilon)\}}\widetilde{T}^n(h_1 L(u)) + \varepsilon) + s_1 1_{\{L(u)<u_0(\underline{u}+\varepsilon)\}}) \\
&\leq \widetilde{T}(h^*(u)) + \varepsilon.
\end{aligned}$$

And therefore,

$$\widetilde{T}h^* = h^*.$$

Moreover, while $\widetilde{T}$ isn't a contraction mapping, it is nevertheless non-expansive, in the sense that

$$||\widetilde{T}h_1 - \widetilde{T}h_2|| \leq ||h_1 - h_2||$$

and it can be checked that in this case, it has a unique fixed point.

To see this, suppose h_a and h_b are two fixed points of $\widetilde{T}$. And let $M = ||h_a - h_b||$. Take the smallest u^* such that $||h_a(u^*) - h_b(u^*)|| = M$[17]. Then we see that $||\widetilde{T}h_a(u^*) - \widetilde{T}h_b(u^*)|| < M$ by the definition of $\widetilde{T}$ unless $M = 0$.

Now, this implies that

$$h^*(u) = h_2(u) = f_2'(u).$$

And therefore, we have

$$f_1'(u) \leq f_2'(u)$$

for all $u \in [u_0(\underline{u}+\varepsilon), u_e]$. ■

With these two lemmas, we are ready to study the effect of the agent's outside option on the

[17] Note that such V^* may not exist both because of the achievability of the sup norm and because of the liminf of such V^*. In this case, we can take the appropriate approximation.

optimal relational contract.

Define $u^P(\underline{u}, \underline{v}, \underline{w})$ as the principal's maximum payoff is the agent's outside option is $\underline{u}$. Define $u^A(\underline{u}, \underline{v}, \underline{w})$ as the associated payoff of the agent. We have the following results.

Proposition 3: *If* $\max_u\{F(u, \underline{u}, \underline{v}, \underline{w})\} > F(u_0(\underline{u}), \underline{u}, \underline{v}, \underline{w})$, *then*

$$\begin{aligned} \frac{\partial u^P(\underline{u}, \underline{v}, \underline{w})}{\partial \underline{u}} &< 0; \\ \frac{\partial u^A(\underline{u}, \underline{v}, \underline{w})}{\partial \underline{u}} &> 0; \\ \frac{\partial (u^P(\underline{u}, \underline{v}, \underline{w}) + u^A(\underline{u}, \underline{v}, \underline{w}))}{\partial \underline{u}} &> 0. \end{aligned}$$

Proof. The first inequality follows directly from Lemma 8.

Lemma 9 implies that $\frac{\partial u^A(\underline{u}, \underline{v}, \underline{w})}{\partial \underline{u}} \geq 0$. (It does not imply a strict inequality directly because f may have kinks at u^A.).

Moreover, Berger's continuity theorem implies that $u^A(\underline{u}, \underline{v}, \underline{w})$ is continuous with respect to $\underline{u}$. Again to simplify notation, suppose f_1 is the PPE Pareto frontier with respect to $\underline{u}$ defined on $[u_0(\underline{u}) + \varepsilon, u_e]$. And let f_2 be the PPE Pareto frontier with respect to $\underline{u} + \varepsilon$ defined on $[u_0(\underline{u}) + \varepsilon, u_e]$.It then suffices to show that, for small enough ε, $f_2(u^A(\underline{u}) + \varepsilon) \geq f_1(u^A(\underline{u})) - \varepsilon$.

To this end, define

$$g_1(u) = f_1(u - \varepsilon) - \varepsilon, \quad \text{for } u \in [u_0(\underline{u}) + \varepsilon, u_e).$$

Now link the line segment between $(\underline{u} + \varepsilon/\delta, \underline{v})$ and $(u_0(\underline{u}) + \varepsilon, f_1(u_0(\underline{u}))$, and define an operator T on bounded functions between $[u_0(\underline{u}) + \varepsilon, u_e]$ as

$$\begin{aligned} T(g(u)) = \; & (1-\delta)(py - \underline{w}) \\ & + \delta p(1_{\{L(u)+k<u_e\}}(g(L(u)+k)) + 1_{\{L(u)+k\geq u_e\}}(py - c - (L(u)+k)) \\ & + \delta(1-p)(1_{\{L(u)\geq u_0(\underline{u})+\varepsilon\}} g(L(u)) + \\ & 1_{\{L(u)<u_0(\underline{u})+\varepsilon\}}(\underline{v} + \frac{L(u) - (\underline{u} - \varepsilon/\delta)}{u_0(\underline{u}) + \varepsilon - (\underline{u} - \varepsilon/\delta)}(f_1(u_0(\underline{u})) - \underline{v})), \end{aligned}$$

With this operator, we can show that

$$T(f_1(u)) > f_1(u), \text{ for } u \in [u_0(\underline{u}) + \varepsilon, u_e).$$

Now notice that, for $u \in [u_0(\underline{u}) + \varepsilon, u_e)$,

$$\begin{aligned}
T(g_1(u)) &= (1-\delta)(py - \underline{w}) \\
&\quad +\delta p(1_{\{L(u)+k<u_e\}}(g_1(L(u)+k)) + 1_{\{L(u)+k\geq u_e\}}(py - c - (L(u)+k)) \\
&\quad +\delta(1-p)(1_{\{L(u)\geq u_0(\underline{u})+\varepsilon\}} g_1(L(u)) \\
&\quad +1_{\{L(u)<u_0(\underline{u})+\varepsilon\}}(\underline{v} + \frac{L(u) - (\underline{u} - \varepsilon/\delta)}{u_0(\underline{u}) + \varepsilon - (\underline{u} - \varepsilon/\delta)}(f_1(u_0(\underline{u})) - \underline{v})) \\
&= (1-\delta)(py - \underline{w}) \\
&\quad +\delta p(1_{\{L(u)+k<u_e\}}(f_1(L(u-\delta\varepsilon)+k) - \varepsilon) + 1_{\{L(u)+k\geq u_e\}}(py - c - (L(u)+k)) \\
&\quad +\delta(1-p)(1_{\{L(u)\geq u_0(\underline{u})+\varepsilon(\}} f_1(L(u-\delta\varepsilon)) - \varepsilon) \\
&\quad +1_{\{L(u)<u_0(\underline{u})+\varepsilon\}}(\underline{v} + \frac{L(u) - (\underline{u} - \varepsilon/\delta)}{u_0(\underline{u}) + \varepsilon - (\underline{u} - \varepsilon/\delta)}(f_1(u_0(\underline{u})) - \underline{v})) \\
&\geq T(f_1(u-\delta\varepsilon)) - \delta\varepsilon \\
&> f_1(u-\delta\varepsilon)) - \delta\varepsilon \\
&\approx f_1(u) - (f_1'(u)+1)\delta\varepsilon \\
&= f_1(u) - (f_1'(u)+1)\varepsilon + (f_1'(u)+1)(1-\delta)\varepsilon \\
&\approx f_1(u-\varepsilon) - \varepsilon + (f_1'(u)+1)(1-\delta)\varepsilon \\
&= g_1(u) + (f_1'(u)+1)(1-\delta)\varepsilon.
\end{aligned}$$

The last inequality follows from the fact that $f_1'(u) > -1$ for $u \in [u_0(\underline{u}) + \varepsilon, u_e)$, which is a consequence of the concavity of f and the definition of u_e.

By a similar proof method as in Lemma 8, this implies that for $u \in [u_0(\underline{u}) + \varepsilon, u_e)$,

$$\begin{aligned}
f_2(u) &> g_1(u) + (f_1'(u)+1)(1-\delta)\varepsilon \\
&= f_1(u-\varepsilon) - \varepsilon + (f_1'(u)+1)(1-\delta)\varepsilon
\end{aligned}$$

This implies that the surplus is strictly increasing.

Finally, from this result and that $\frac{\partial u^P(\underline{u},\underline{v},\underline{w})}{\partial \underline{u}} < 0$, we have that $\frac{\partial u^A(\underline{u},\underline{v},\underline{w})}{\partial \underline{u}} > 0$. ■

Similar method as above can establish the following result.

Proposition 4: *If* $\max_u\{F(u,\underline{u},\underline{v},\underline{w})\} > F(u_0(\underline{u}),\underline{u},\underline{v},\underline{w})$, *then*

$$\begin{aligned}
\frac{\partial u^P(\underline{u},\underline{v},\underline{w})}{\partial \underline{v}} &> 0; \\
\frac{\partial u^A(\underline{u},\underline{v},\underline{w})}{\partial \underline{v}} &< 0; \\
\frac{\partial (u^P(\underline{u},\underline{v},\underline{w}) + u^A(\underline{u},\underline{v},\underline{w}))}{\partial \underline{v}} &< 0.
\end{aligned}$$

Proof. The proof parallels the arguments in Proposition 3, and it is easier. so we omit it here. Neither u_0 and u_e changes with $\underline{v}$, but the value of f to the left of u_0 increases because of the increase in $\underline{v}$. It follows in an argument like Lemma 8 that the value of f in $[u_0, u_e]$ must increase so that the principal's payoff must increase. Moreover, an argument similar to Lemma 9 shows that the derivative of f in (u_0, u_e) must decrease, so the expected payoff of the agent decreases. Finally, following Proposition, we can define a function $g(u) = f(u + \varepsilon) + \varepsilon$, and a similar argument can show that the overall efficiency decreases. ■

Finally, we examine the effect of the minimum wage. We show that the minimum wage decreases the principal's payoff and the overall efficiency.

Proposition 5: *If* $\max_u\{F(u, \underline{u}, \underline{v}, \underline{w})\} > F(u_0(\underline{u}), \underline{u}, \underline{v}, \underline{w})$, *then*

$$\begin{aligned} \frac{\partial u^P(\underline{u}, \underline{v}, \underline{w})}{\partial \underline{w}} &< 0 \\ \frac{\partial (u^P(\underline{u}, \underline{v}, \underline{w}) + u^A(\underline{u}, \underline{v}, \underline{w}))}{\partial \underline{w}} &< 0. \end{aligned}$$

Proof. Suppose the minimum wage increases from $\underline{w}$ to $\underline{w} + \varepsilon$. In this case, both u_0 and u_e increases by $\frac{1-\delta}{\delta}\varepsilon$. Then the first inequality can be established by using a method similar to Lemma 8 and noting that $f' \geq -1$. The second inequality follows Proposition 3 closely. Again to simplify notation, suppose f_1 is the PPE Pareto frontier with respect to $\underline{w}$ defined on $[u_0(\underline{u}) + \frac{1-\delta}{\delta}\varepsilon,\ u_e + \frac{1-\delta}{\delta}\varepsilon]$. And let f_2 be the PPE Pareto frontier with respect to $\underline{w} + \varepsilon$ defined on $[u_0(\underline{u}) + \frac{1-\delta}{\delta}\varepsilon,\ u_e + \frac{1-\delta}{\delta}\varepsilon]$.It then suffices to show that, for small enough ε, $f_2(u^A(\underline{u}) + \varepsilon) \geq f_1(u^A(\underline{u})) - \varepsilon$. (If u^A decreases, we immediately have the decrease in efficiency because both the principal and the agent's payoff decrease.)

To this end, define

$$g_1(u) = f_1(u - \varepsilon) - \varepsilon, \quad \text{for } u \in [u_0(\underline{u}) + \frac{1-\delta}{\delta}\varepsilon, u_e + \frac{1-\delta}{\delta}\varepsilon].$$

Now link the line segment between $(\underline{u}, \underline{v})$ and $(u^A(\underline{u}) + \frac{1-\delta}{\delta}\varepsilon, f_1(u_0(\underline{u}) + \frac{1-\delta}{\delta}\varepsilon)$, and define an

operator T on bounded functions between $[u_0(\underline{u}) + \frac{1-\delta}{\delta}\varepsilon, u_e + \frac{1-\delta}{\delta}\varepsilon]$ as

$$\begin{aligned} T(g(u)) &= (1-\delta)(py - \underline{w} - \varepsilon) \\ &\quad +\delta p(1_{\{L(u)+k-\frac{1-\delta}{\delta}\varepsilon<u_e\}}(g(L(u)+k-\frac{1-\delta}{\delta}\varepsilon) + \\ &\quad 1_{\{L(u)+k-\frac{1-\delta}{\delta}\varepsilon\geq u_e\}}(py - c - (L(u)+k-\frac{1-\delta}{\delta}\varepsilon)) \\ &\quad +\delta(1-p)(1_{\{L(u)-\frac{1-\delta}{\delta}\varepsilon\geq u_0(\underline{u})+\frac{1-\delta}{\delta}\varepsilon\}}g(L(u)-\frac{1-\delta}{\delta}\varepsilon) + \\ &\quad 1_{\{L(u)-\frac{1-\delta}{\delta}\varepsilon\geq u_0(\underline{u})+\frac{1-\delta}{\delta}\varepsilon\}}(\underline{v} + \frac{L(u)-\frac{1-\delta}{\delta}\varepsilon - \underline{u}}{u_0(\underline{u})+\frac{1-\delta}{\delta}\varepsilon - \underline{u}}(f_1(u_0(\underline{u}) + \frac{1-\delta}{\delta}\varepsilon) - \underline{v})), \end{aligned}$$

With this operator, we can show that

$$T(f_1(u)) > f_1(u), \text{ for } u \in [u_0(\underline{u}) + \frac{1-\delta}{\delta}\varepsilon, u_e + \frac{1-\delta}{\delta}\varepsilon).$$

Finally, as in Proposition 3, we have the following:

$$\begin{aligned} T(g_1(u)) &\leq T(f_1(u-\delta\varepsilon)) - \delta\varepsilon - (1-\delta)\varepsilon \\ &\quad -(1-\delta)\varepsilon(pf_1'(u-\delta\varepsilon+k) + (1-p)f_1'(u-\delta\varepsilon)) \\ &< f_1(u-\delta\varepsilon)) - \delta\varepsilon - (1-\delta)\varepsilon - (1-\delta)\varepsilon f_1'(u) \\ &\approx f_1(u) - (f_1'(u)+1)\delta\varepsilon - (1-\delta)\varepsilon - (1-\delta)\varepsilon f_1'(u) \\ &= g_1(u). \end{aligned}$$

■