On the Performance of Quota-based Compensation Scheme for Heterogeneous Salesforces

Masahiro Watabe
Rissho University
Department of Economics

May 7, 2018
On the Performance of Quota-based Compensation Scheme for Heterogeneous Salesforces

Masahiro Watabe*
Rissho University
Department of Economics

May 7, 2018

Abstract

This paper examines the performance of quota-based compensation scheme relative to the menu of linear contracts in a principal-agent problem under moral hazard and adverse selection. A firm compensates heterogeneous workers based on uncertain outcomes of the worker's private skill and effort. I derive an analytical expression for the performance measure of simple contract and provide explicit upper and lower bounds on the relative performance. A primary finding is that a properly designed simple piece-wise-linear-threshold contract always captures more than 73 percent of the incremental gain secured under the optimal menu of linear contracts on the entire region of the relevant parameter values.

Keywords: piece-rate contract, quota-based contract, complexity of contracts, private information, worker heterogeneity

JEL Classification Numbers: D82; M52; J33

*Department of Economics, Rissho University, 4-2-16 Osaki, Shinagawa-ku, Tokyo, 141-8602 Japan. Email: mwatabe@ris.ac.jp. I thank for helpful comments to Akifumi Ishiara and participants at 21st Annual Conference of the Society for Institutional & Organizational Economics (SIOE2017), Hitotsubashi University and Japanese Economic Association 2017 Autumn Meeting. All errors are my own. This work was supported by JSPS KAKENHI Grant Number 15K03363, and Economic Research Institute (Rissho University).
1 Introduction

The paper analyzes a simple quota-based compensation scheme between a risk-neutral principal and a risk-averse agent when an adverse selection problem and a moral hazard problem occur. A quota-based contract is a piece-wise linear performance-based compensation scheme that can be regarded as a menu of two simple contracts such as fixed wage and piece-rate contracts. The paper shows that a simple quota-based contract can secure a substantial gain that the firm secures with a theoretically optimal menu of linear contracts, although a single linear contract does not perform well in general. The identified strong performance of quota-based piece-wise linear contracts may help to explain their widespread use in practice.

How do firms compensate heterogeneous workers by a simple compensation scheme? Fixed-payment and piece-rate contracts are common in practice. The literature on pay for performance has argued the prevalence of simple compensation plans rather than complex and fine-tuned nonlinear contracts. Simple plans are easy to understand by the salespersons, and the cost of enforcement of contracts is low.

Most incentive schemes observed in practice are piece-wise linear. My paper is close in spirit to Schmalensee (1989) and Larkin (2014). Regarding the issue of pay for performance, as Larkin (2014) noted, quotas are commonly used as an incentive-free threshold that salespeople must clear in order to start earning commissions, or sometimes to earn a bonus. However, Gibbs (2013, pp.411-414) argued a couple of negative effects under piece-wise linear-threshold compensation contracts. Oyer (1998) also discussed the problem of timing gaming caused by the use of sales quotas. Tsuru (2008) empirically examined the impact of of a performance-based pay scheme introduced by a large Japanese auto sales firm. A linear compensation system

---

1Eggleston et al (2000, p.91) point out that recent work in the law and economics of contracts suggests that contracts ought to be highly complex and fine-tuned. They argue a number of reasons why simple contracts prevail.
was replaced with a quota-based contract. Tsuru (2008) found undesirable gaming behavior by examining the data supplied by the dealership personnel for the period before and after the pay scheme change. Larkin (2014) empirically examined the gaming cost of a nonlinear incentive scheme in terms of forgone revenue for the firm. As mentioned in Gibbs (2013, p.413), a natural question raised can be now stated: given the negative effects of thresholds, why do firms use quota-based contracts? The paper provides theoretical justification for the use of a simple quota-based compensation plan and the intuition for why a single piece-wise linear contract is a good approximation of more complex contracts.

A series of studies have utilized the framework of agency theory to examine the relative performance of different contracts. Basu et al (1985) (henceforth called BLSS in the literature) and subsequent studies have derived the compensation scheme for risk-averse and homogeneous workers, not as a screening device. Several works have discussed the structure of a piece-wise-linear-threshold contract. Bose et al (2011) examined the ability of linear contracts with a threshold to replicate the performance of optimal contracts in the canonical moral hazard setting in which a risk-neutral principal contracts with a single risk-averse agent. Somewhat confusingly, Bose et al (2011, p.104) refer to a quota-based piece-wise linear contract as a linear contract. Raju and Srinivasan (1996) showed that a quota plan with a common salary and commissions rate across salespersons, but with quotas varying across the salesperson-territories to the optimal agency-theoretical contract proposed in BLSS. Their numerical experiments indicate that the loss in profits is likely to be small, only about 1% for the parametric scenarios studied. Chen and Miller (2009) employed numerical simulations and showed that the optimal piece-wise-linear-threshold contract is significantly superior to the best linear contract when the agent has a power utility function in a multi-period model. Daljord et al (2016) paid attention to the constraint faced by firms in fine-tuned contracts to the full distribution of heterogeneity of their employees, and examine the profitability under partially heterogeneous contracts by means of numerical experiments. These papers except Bose et al (2011) conducted numerical experiments to show how a simple piece-wise linear contract performs compared to the agency-theoretic compensation plan. In contrast to numerical experiments, I shall derive an analytical expression for the performance measure of simple piece-wise-linear contracts regarded as screening devices, and provide explicit upper and lower bounds on the entire region of the relevant parameter values. I will treat the menu of linear contracts, in which each personalized linear contract is linear in a performance measure but the two components such as fixed salary and commission rate can be non-linear functions of the reported type, as a theoretical benchmark.

Several empirical works have examined the role of compensation policy in influencing worker performance. Lazear (2000) studied a firm that switched its compensation policy from hourly wages to piece-rate. He estimated that the average worker’s output increased following the change in compensation policy. Dohmen and Falk (2011) conducted a further examination of output differences between different incentive schemes. They studied the impact of multidimensional characteristics such as productivity, risk attitudes, social preferences, and gender on the choice between a fixed-payment contract and a variable-payment contract such as piece-rate, tournament, and revenue-sharing contracts. They found that output in all variable-payment contracts is higher than output under the fixed-payment contract. More recently, Larkin and Leider (2012) examined subjects’ choices over the linear and convex pay schemes in a laboratory experiment. They found that overconfident subjects facing a choice between a linear contract and a convex one are more likely to choose the convex scheme, and suggested that unmotivated employees who wish to exert lower effort may prefer a linear incentive scheme rather than a convex one. Furthermore, several empirical studies also examined the change in productivity caused by changes in incentive schemes, for example, Paarsch and

---

2See Albers (1996) for intensive reviews about the design of salesforce compensation schemes mainly in the 1980s and 1990s.
INTRODUCTION

Shearer (1999, 2000) and Franceschelli et al (2010). The present paper examines the difference in productivity under alternative contracts, and argues the relationship between productivity and profitability.

As mentioned earlier, I consider the optimal menu of linear contracts in which workers announce their types as a theoretical benchmark in this paper. Of course, the menu of linear contracts may not be fully optimal for the firm in general. For risk-neutral and heterogeneous workers, Rao (1990) examined a sufficient condition for the implementability of the optimal compensation scheme by a menu of linear contracts. For risk-averse and heterogeneous agents, Bernardo et al (2001) examined optimal capital allocation and managerial compensation policies in a decentralized firm in the presence of both moral hazard and adverse selection. Bernardo et al (2001, Proposition 3, p.322) began by considering managerial compensation contracts in the form of a menu of linear contracts, and then showed that there does not exist a general mechanism that improves upon the menu of linear contracts. I am not sure whether the same result would hold in my framework, but the structure of the fully optimal contract is beyond the scope of this paper.

Combining the results regarding the relative profitability and productivity of simple contracts, the paper sheds light on the question of why firms use piece-wise-linear-threshold contracts in practice even though several negative effects has been discussed. The paper examines the performances of simple piece-wise-linear contracts with and without a quota. The paper provides an explanation for the use of piece-wise-linear-threshold contract from the view point of economics. The optimal piece-rate contract (as an indirect mechanism) provides incentives enough to achieve the average productivity of the optimal menu of linear contracts (as a direct revelation mechanism), and can secure a large fraction of the gain that the menu of linear contracts captures on the region of parameter values over which there is no bunching in the menu of linear contracts. However, such good performance no longer persists for a broad parameterizations in which there is a bunching in the menu of linear contracts. Besides, the paper shows that it is not difficult to construct a simple piece-wise-linear-threshold contract can secure a substantial gain of the optimal menu of linear contracts over the whole range of admissible parameter values, although such quota-based contract would have weaker incentives than the piece-rate contract on average.

An important aspect of the paper is the complexity of contracts. The paper also argues the shapes of compensation schemes defined as direct revelation mechanisms based on the announcements by heterogeneous workers. The theoretical benchmark I consider here is the optimal menu of linear contracts, a continuum of type-dependent linear contracts. Each contract is linear in the performance measure, but it could be non-linear in the reported type. Needless to say, the implementational complexity of the fully optimal compensation plan is prohibitive in practice as well as a menu of linear contracts that depends on announcements of private information. By contrast, piece-wise-linear compensation schemes are functions of the performance measure, including a fixed-payment contract and a piece-rate contract with and without thresholds as typical examples. These much simpler contracts are indirect mechanisms from the point of view of mechanism design. It is worth comparing all of compensation schemes studied in the paper in the same domain because the optimal menu of linear contracts depends on an announcement of type or private information rather than on observed output. I provide a procedure transform the optimal menu of linear contracts as a direct revelation mechanism into an anonymous nonlinear compensation scheme as an indirect mechanism. This kind of argument is referred to as the “taxation principle”, the reverse of the “revelation principle” in the literature on mechanism design. For any menu of linear contracts satisfying incentive compatibility and participation constraints, a proposition in the paper provides a constructive procedure to obtain an indirect mechanism which duplicates the same outcome as the given direct revelation mechanism. The advantage of the constructive procedure in the paper is that the first-and second-derivatives of the im-
plied nonlinear compensation scheme are so tractable that I am able to determine the shape of that scheme. It is shown that the optimal menu of linear contracts can be replaced by a single convex compensation scheme.

The rest of the paper is organized as follows: In the next section I describe an analytical model. Following this, I discuss the structure of the compensation plan for risk averse and heterogeneous workers. I pay attention to two alternative contracts: theory-based menu of linear contracts (direct revelation mechanism) in Section 3, and quota-based contract (indirect mechanism) in Section 4. Also, a single piece-rate contract is obtained as a corollary. In Section 5, I analyze the relative performance of simple contracts relative to the menu of linear contracts. I shall conclude that the quota-based contract is superior to the piece-rate contract within the entire region of the relevant parameter values. The purpose of Section 6 is twofold. First, I compare incentive effects among the three forms of compensation schemes in Section 6.1. Secondly, I provide a constructive way of transforming any feasible menu of linear contracts into an anonymous nonlinear compensation scheme in Section 6.2. Section 7 concludes.

2 The Model

This paper examines the performance of simple incentive schemes relative to the menu of linear contracts in a principal-agent problem under moral hazard and adverse selection. I consider a one-period principal-agent relationship between a risk-neutral firm (principal) and a risk-averse worker (agent) of the firm. A firm compensates heterogeneous workers based on uncertain outcomes generated by salespersons’ effort and skill. The salespersons or the workers are at different skill levels. A given amount of effort does not always result in the same amount of sales. The firm compensates workers based on sales achieved, and realizes the profit from the sales net of compensation. Not surprisingly, salesforce compensation has received considerable attention from marketing academics on theoretical and empirical fronts (see Basu and Kalyanaram, 1990; Basu et al, 1985; Coughlan and Narasimhan, 1992; Misra et al, 2005; Raju and Srinivasan, 1996).

An important aspect of the selling environment is that a certain amount of effort does not always result in the same amount of sales because of the existence of uncertainty. Denote a skill level by \( q \) and an effort level by \( e \). A response function specifies the expected sales or output \( y \) that depends on the levels of effort and skill. In order to evaluate the performances of simple contracts, I assume a linear response function with normal errors as is standard in the literature (see Dutta, 2008; Hölmstrom and Milgrom, 1987; Misra et al, 2005; Picard, 1987). Output \( y \) is determined by the effort the worker expends \( e \) and the productivity \( q \) of him, according to \( y = \gamma \theta + e + \varepsilon \), where \( \varepsilon \) is the normally distributed noise or measurement error in output with mean zero and variance \( \sigma^2 \). The term \( \theta \) can be interpreted as the base level of sales in the absence of salesforce effort. The marginal product of effort is the same across the workers. The parameter \( \gamma \) represents the complementarity between the firm and a worker in the sense that high \( \gamma \) firm produces more

---

\(^{3}\)Dutta (2008) employed the response function of the form \( y = \gamma \theta + e + \varepsilon \). He mainly discussed a possibility of bunching in the menu of linear contracts. My model differs from Dutta (2008) in that I assume that the reservation utility is constant in order to restrict a type of bunching. Several papers also employed additively linear response functions in effort and private information term, for instance, Gibbons (1987) and Picard (1987). Gibbons (1987) considered type \( \theta \) that appears in \( y = \theta + e \) as the difficulty of the job. Only workers know the difficulty of their jobs. In contrast, Goldmanis and Ray (2015) assumed the multiplicative response function of the form \( y = \gamma \theta e + \varepsilon \) whose marginal product of effort is type-dependent, namely, \( \gamma \theta \). Baker and Jorgensen (2003) assumed that the output \( y = \theta e + \varepsilon \) depends on two independent random variables, one of which \( \theta \) affects the agent’s marginal product of effort, and the other \( \varepsilon \) does not.
output with high \( \theta \) types than with low \( \theta \) types. The constant \( \gamma > 0 \) is a known parameter to both sides.\(^4\) The expression \( \gamma \theta \) can be interpreted as effective skill. The interpretation is that the parameter \( \gamma \) reflects the relative importance of the innate skill in determining the output.

The agent has private information about his skill \( \theta \) that is a degree of heterogeneity in productivity. There is a continuum of types. The firm’s prior beliefs regarding \( \theta \) are represented by a distribution function \( F(\theta) \) defined on the interval \([\theta_{\text{min}}, \theta_{\text{max}}]\). It is not necessarily distributed uniformly:

\[
F(\theta) \triangleq 1 - \left( \frac{\theta_{\text{max}} - \theta}{\Delta} \right)^\eta = 1 - \left( 1 - \frac{\theta - \theta_{\text{min}}}{\Delta} \right)^\eta,
\]

where \( \Delta = \theta_{\text{max}} - \theta_{\text{min}} \) and \( \eta \in (0, \infty) \).\(^5\) In what follows, the salesperson’s participation is considered as exogenous under any types of contracts in the sense that all of types are employed and exclusion is not allowed.\(^6\) Notice that the distribution coincides with the uniform distribution when \( \eta = 1 \). Roughly speaking, the distribution function has a value near the lower bound \( \theta_{\text{min}} \) when \( \eta \in (0, 1) \), whereas the distribution has a value near the upper bound \( \theta_{\text{max}} \) when \( \eta \in (1, \infty) \). The distribution becomes more favorable when \( \eta \) increases, because a higher value of \( \eta \) means that there is a larger proportion of high skill salesforces. In fact, the parameter \( \eta \) is interpreted as a first-order stochastic dominance shift parameter.\(^7\) The distribution function is flexible, and can accommodate J- and inverted-J-shapes.\(^8\) Figure 1 illustrates the shape of the distribution for some alternative values of \( \eta \) ranging between 0.5 and 4.

\[
\begin{align*}
  \theta_{\text{max}} &= 5, \\
  \theta_{\text{min}} &= 1
\end{align*}
\]

Figure 1: Illustration of the distribution \((\theta_{\text{max}} = 5, \theta_{\text{min}} = 1)\)

---

4 If \( \gamma = 0 \) then the response function is independent of the private information, and salesforces are heterogeneous.

5 This kind of a distribution is referred to as a Burr type XII distribution in Miravete (2004). He employed this type of distribution function in the context of the second-degree price discrimination.

6 In this regard, Rao (1990) also assumed that all of salespersons are active in the salesforce.

7 \( dF(\theta; \eta) / d\eta = (\theta_{\text{max}} - \theta)^{\frac{1}{\eta}} \ln \left( ((\theta_{\text{max}} - \theta) / \Delta) / \eta^2 \Delta^2 \right) \leq 0 \) with equality only if \( \theta = \theta_{\text{min}} \).

8 A couple of previous works consider a deviation from the uniform distribution when adverse selection is a potential concern. For example, Reichelstein (1992) and Chu and Sappington (2007) in the procurement problem. Also, Rao (1990) in the context of salesforce compensation.
distribution in Eq.(1), the inverse hazard rate is linear in the private information. This is calculated explicitly as in Eq.(2). The monotonicity of the inverse hazard rate is automatically satisfied.

\[
\frac{1 - F(\theta)}{f(\theta)} = \eta(\theta_{\text{max}} - \theta). \tag{2}
\]

Some useful properties of the above family of distributions are summarized by the following lemma. The proof is omitted.

**Lemma 1 (expectation and variance).** When \( \bar{\theta} = \theta_{\text{max}} - \varepsilon \) for some \( \varepsilon \geq 0 \),

\[
\int_{\bar{\theta}_{\text{min}}}^{\bar{\theta}} \theta f(\theta)d\theta = \mathbb{E}(\theta) - \text{Prob}(\theta > \bar{\theta}) \left( \mathbb{E}(\theta) + \frac{\Delta - \varepsilon}{\eta + 1} \right) \tag{3}
\]

and

\[
\int_{\bar{\theta}_{\text{min}}}^{\bar{\theta}} \theta^2 f(\theta)d\theta = \text{Var}(\theta) + \mathbb{E}(\theta)^2 - \text{Prob}(\theta > \bar{\theta}) \left( \mathbb{E}(\theta) + \frac{\Delta - 2\varepsilon}{\eta + 1} \right) + \frac{\varepsilon^2}{2(\eta + 1)}, \tag{4}
\]

where

\[
\mathbb{E}(\theta) = \frac{\eta \theta_{\text{max}} + \theta_{\text{min}}}{\eta + 1} \quad \text{and} \quad \text{Var}(\theta) = \frac{\Delta^2 \eta^2}{(\eta + 1)^2(2\eta + 1)}. \]

By Lemma 1, the average over the interval \([\bar{\theta}, \theta_{\text{max}}]\) can be calculated as

\[
\mathbb{E}[\theta | \theta \geq \bar{\theta}] = \frac{\int_{\bar{\theta}}^{\theta_{\text{max}}} \theta f(\theta)d\theta}{1 - F(\theta)} = \mathbb{E}(\theta) + \frac{\hat{\theta} - \theta_{\text{min}}}{\eta + 1} = \frac{\eta \theta_{\text{max}} + \bar{\theta}}{\eta + 1}, \tag{5}
\]

where the second equality follows because \( \Delta - \varepsilon = (\theta_{\text{max}} - \theta_{\text{min}}) - (\theta_{\text{max}} - \hat{\theta}) = \hat{\theta} - \theta_{\text{min}} \) in Eq.(3).

I now describe the salesperson’s optimization program that is similar to the most of the literature. The salesperson is assumed to be risk-averse. The worker’s risk averse preferences are represented by a negative exponential utility function of the form \( u(w) = -\exp(-\rho w) \) as a function of his end-of-period compensation \( w \), where \( \rho > 0 \) is his constant absolute risk aversion coefficient. The disutility function to the salesperson of exerting effort level \( \varepsilon \) is denoted by \( C(\varepsilon) \). For tractability, I assume a specific functional form for the cost of effort, and it is captured by a quadratic function \( C(\varepsilon) = \frac{e^2}{2k} \) for some \( k > 0 \). Here, \( \frac{1}{k} \) is the marginal disutility of effort. The assumption of a quadratic disutility function is made for expositional convenience. As the standard argument, his certainty equivalent is approximately given by \( CE \approx \mathbb{E}(z) - \frac{k}{2} \text{Var}(z) \), where \( z = w - C(\varepsilon) \) is the worker’s net payoff.\(^9\) The salesperson then maximizes his certainty equivalent. Under a linear contract \( w(y) = \alpha + \beta y \), his certainty equivalent can be written as \( CE = \alpha + \beta(\gamma \varepsilon + e) - C(e) - \frac{e^2}{2k} \beta^2 \). The risk averse agent bears a disutility from uncertainty captured by the risk premium term \( \rho e^2 \beta^2 / 2 \). In addition, the reservation wage or the minimum certainty equivalent of the salesperson is denoted by \( \bar{w} \geq 0 \). I assume that the reservation wage does not depend on the private information.\(^{10}\)

\(^9\)The certainty equivalent can be approximately by \( CE \approx \bar{z} - \frac{1}{2} \rho(\bar{z}) \text{Var}(z) \), where \( \rho(\bar{z}) = -u''(\bar{z})/u'(\bar{z}) \) and \( \bar{z} = \mathbb{E}(z) \). When the compensation scheme is given by \( w(y) = \alpha + \beta y = \alpha + \beta g(\varepsilon, \theta) + e \), the expectation of the net utility with respect to \( \varepsilon \) is given by \( \mathbb{E}(z) = \alpha + \beta g(\varepsilon, \theta) - C(e) \), and hence \( z - \mathbb{E}(z) = \beta e \). Moreover, the variance of the net utility is given by \( \text{Var}(z) = \mathbb{E}((z - \mathbb{E}(z))^2) = \mathbb{E}((\beta e)^2) = \beta^2 \mathbb{E}(e^2) = \beta^2 \text{Var}(e) = \beta^2 e^2 \). See Milgrom and Roberts (1992, p.247) for more details.

\(^{10}\)In contrast, Dutta (2008) assumed that the salesperson’s reservation wage has the form of \( w(\theta) = \bar{w} + \lambda \gamma \theta \) for some \( \lambda \geq 0 \).
Finally, define $\delta = \frac{k}{\eta}$ and $\lambda = \frac{\xi}{\lambda}$ for convenience. The parameter $\lambda$ can be regarded as a measure of the amount of asymmetric information and the degree of worker heterogeneity in the response function. It becomes a negligible factor when either the difference $\Delta = \theta_{\text{max}} - \theta_{\text{min}}$ or the marginal product $\gamma$ of worker heterogeneity $\theta$ in the response function tends to infinity.

3 Menu of Linear Contracts

To start this section, I shall consider the fixed wage contract as a benchmark. If the firm sets a fixed wage, then it is not difficult to see that the worker does not exert any effort, and thus, the optimal fixed wage will equal to the reservation wage based on the binding participation constraint. The firm’s maximized expected profit will be written as

$$
\int_{\theta_{\text{min}}}^{\theta_{\text{max}}} \left[ y - \bar{w} \right] f(\theta) d\theta = \gamma \int_{\theta_{\text{min}}}^{\theta_{\text{max}}} \theta f(\theta) d\theta - \bar{w} = \gamma \mathbb{E}(\theta) - \bar{w} = \frac{\gamma(\eta \theta_{\text{max}} + \theta_{\text{min}})}{\eta + 1} - \bar{w}.
$$

As a consequence, the firm’s expected profit under the optimal fixed wage is given by

$$
\mathbb{E}\pi_{fw}(\gamma, \eta) \triangleq \frac{\gamma(\eta \theta_{\text{max}} + \theta_{\text{min}})}{\eta + 1} - \bar{w}.
$$

The firm can choose more complex wage schedule to maximize his expected profit subject to the worker’s optimizing behavior. The effort of the salesperson is not observable and hence not contractible. I restrict focus to a menu of linear compensation schemes $\{a_{m}(\cdot), \beta_{m}(\cdot)\}$ of the form $w_{m}(\theta, y) = a_{m}(\theta) + \beta_{m}(\theta) y$ with the pay-performance sensitivity $\beta_{m}(\theta)$ restricted to be non-negative as in the literature (see Bernardo et al, 2001; Dutta, 2008; Dutta and Fan, 2014; Reichelstein, 1992). The contact is linear in $y$ but the two components $a_{m}(\theta)$ and $\beta_{m}(\theta)$ can be non-linear functions of the reported type $\theta$. A menu of linear contracts are designed to elicit each worker’s type and to motivate each worker to take a desired action for the firm.

To express the incentive compatibility constraint, define $CE_{m}(\hat{\theta}, \theta)$ as the certainty equivalent of the worker of type $\theta$ who reports type $\hat{\theta}$:

$$
CE_{m}(\hat{\theta}, \theta) \triangleq a_{m}(\hat{\theta}) + \beta_{m}(\hat{\theta})(\gamma \theta + c_{m}(\hat{\theta})) - C(c_{m}(\hat{\theta})) - \frac{\rho \sigma^{2}}{2} \beta_{m}(\hat{\theta})^{2},
$$

where

$$
c_{m}(\hat{\theta}) \in \text{argmax} \left[ a_{m}(\hat{\theta}) + \beta_{m}(\hat{\theta})(\gamma \theta + c) - C(c) - \frac{\rho \sigma^{2}}{2} \beta_{m}(\hat{\theta})^{2} | c \geq 0 \right].
$$

Also, define $CE_{m}(\theta) = CE_{m}(\hat{\theta}, \theta)$ as the utility from reporting the true type $\theta$. Then, the incentive compatibility constraint is $CE_{m}(\theta) \geq CE_{m}(\hat{\theta}, \theta)$ for every pair $(\theta, \hat{\theta})$, and the participation constraint is given by $CE_{m}(\theta) \geq \bar{w}$ for every $\theta$. In terms of the information rent in excess of the reservation wage $\bar{w}$, define $R_{m}(\hat{\theta}, \theta) \triangleq CE_{m}(\hat{\theta}, \theta) - \bar{w}$ and $R_{m}(\theta) \triangleq CE_{m}(\theta) - \bar{w}$. The incentive compatibility and participation constraints are both taken into account by the firm in its profit maximization problem. Any form of direct revelation mechanism is said to be feasible if it satisfies both types of constraints. Similar to Dutta (2008), I have the following result. The proof is omitted.
Lemma 2 (feasible contracts). A direct revelation mechanism \( (\alpha_{mt}(\cdot), \beta_{mt}(\cdot)) \), where \( w_{mt}(\theta, y) = \alpha_{mt}(\theta) + \beta_{mt}(\theta)y \), satisfies (IC) \( R_{mt}(\theta) \geq R_{mt}(\hat{\theta}, \theta) \) and (IR) \( R_{mt}(\theta) \geq 0 \) for every \( \theta \) if and only if \( \hat{R}_{mt}(\theta) = \beta_{mt}(\theta)\gamma \), \( R_{mt}(\theta_{\text{min}}) \geq 0 \), and \( \beta_{mt}(\cdot) \) is non-decreasing.

The standard arguments yield the following pay-performance sensitivity and the effort function as a direct revelation mechanism.

Proposition 1 (incentive-intensity under the menu of linear contracts). The optimal pay-performance sensitivity is given by

\[
\beta_{mt}(\theta) = \max \left\{ \frac{k}{k + \rho \sigma^2} \left( 1 - \frac{\gamma}{k} \frac{1 - F(\theta)}{f(\theta)} \right), 0 \right\}.
\]

Moreover, The optimal effort function is the following:

\[
e_{mt}(\theta) = \max \{ \beta_{mt}(\theta) k, 0 \}
\]

Denote the corresponding production rule by \( y_{mt}(\theta) = \gamma \theta + e_{mt}(\theta) \) under the menu of linear contracts.

Recall that the incentive compatibility constraint requires a weak monotonicity of the pay-performance sensitivity. Substituting the inverse hazard rate \( (1 - F(\theta))/f(\theta) = \eta (\theta_{\text{max}} - \theta) \) into Eq.(7), the pay-performance sensitivity is written as

\[
\beta_{mt}(\theta) = \frac{k}{k + \rho \sigma^2} \left( 1 - \frac{\gamma \eta (\theta_{\text{max}} - \theta)}{k} \right) = \frac{k - \gamma \eta (\theta_{\text{max}} - \theta)}{k + \rho \sigma^2}.
\]

This implies that the weak monotonicity is guaranteed without any additional assumption because \( \beta_{mt}(\theta) = \max \{ \frac{\gamma \eta}{k + \rho \sigma^2}, 0 \} \geq 0 \) over the interval \([\theta_{\text{min}}, \theta_{\text{max}}]\). However, the non-negativity of the pay-performance sensitivity itself requires additional restrictions on the parameters. Notice that \( \beta_{mt}(\theta) \geq 0 \) holds when the expression inside round brackets in Eq.(7) is positive:

\[
1 - \frac{\gamma}{k} \frac{1 - F(\theta)}{f(\theta)} = 1 - \frac{\gamma \eta (\theta_{\text{max}} - \theta)}{k} > 0 \iff \theta > \theta_{\text{max}} - \frac{k}{\gamma \eta}.
\]

Indeed, \( e_{mt}(\theta) > 0 \) for every \( \theta > \theta_{\text{mt}}^{\text{min}} \), where the marginal type is given by \( \theta_{\text{mt}}^{\text{min}} \triangleq \theta_{\text{max}} - k / \gamma \eta \). In other words, a sufficient condition for that all of types exert positive effort is that \( \theta_{\text{mt}}^{\text{min}} < \theta_{\text{min}} \). This condition can be written as \( \theta_{\text{mt}}^{\text{min}} < \theta_{\text{min}} \iff \theta_{\text{max}} - \theta_{\text{min}} < k / \gamma \eta \iff \Delta < k / \gamma \eta \iff k > \gamma \eta \Delta \iff \lambda = k / \gamma \eta \Delta > 1 \).

In other words, the existence of a bunching in the menu of linear contracts is determined whether the parameter \( \lambda \) is bounded from below or the degree of asymmetric information \( \Delta \) is bounded from above.\(^{11}\)

It is convenient to characterize the relevant parameters in terms of \((\eta, \lambda)\) instead of \((\gamma, \eta)\). I use them interchangeably.

Proposition 2 (bunching under the menu of linear contracts). There is no bunching (resp. there is a bunching) in the menu of linear contracts in Proposition 1 if \( \lambda > 1 \) (resp. \( \lambda \leq 1 \)).

\(^{11}\)If \( k - \gamma \eta \Delta > 0 \) then \( 0 < k - \gamma \eta (\theta_{\text{max}} - \theta_{\text{min}}) \leq k - \gamma \eta (\theta_{\text{max}} - \theta) \) for every \( \theta \geq \theta_{\text{min}} \).
I shall derive the maximized expected profit under the menu of linear contracts derived in Proposition 1. The firm’s objective function to be maximized is given by

\[ \int_{\theta_{\min}}^{\theta_{\max}} \left[ y_{m\ell}(\theta) - \{\alpha_{m\ell}(\theta) + \beta_{m\ell}(\theta)y_{m\ell}(\theta)\}\right] f(\theta)d\theta. \]

I can use the fact that \( R_{m\ell}(\theta_{\min}) = 0 \) to obtain the total payment to type \( \theta \):

\[
\alpha_{m\ell}(\theta) + \beta_{m\ell}(\theta)y_{m\ell}(\theta) = R_{m\ell}(\theta) + C(e_{m\ell}(\theta)) + \frac{\rho\sigma^2}{2} \beta_{m\ell}(\theta)^2 + \bar{\omega}
\]

\[
= \int_{\theta_{\min}}^{\theta} \beta_{m\ell}(s)\gamma ds + C(e_{m\ell}(\theta)) + \frac{\rho\sigma^2}{2} \beta_{m\ell}(\theta)^2 + \bar{\omega}.
\]

Get back into the firm’s objective to obtain

\[
\int_{\theta_{\min}}^{\theta_{\max}} \left[ y_{m\ell}(\theta) - \int_{\theta_{\min}}^{\theta} \beta_{m\ell}(s)\gamma ds - C(e_{m\ell}(\theta)) - \frac{\rho\sigma^2}{2} \beta_{m\ell}(\theta)^2 \right] f(\theta)d\theta - \bar{\omega}.
\]

The expression for the double integral can be written as

\[
\int_{\theta_{\min}}^{\theta_{\max}} \left( \int_{\theta_{\min}}^{\theta} \beta_{m\ell}(s)\gamma ds \right) f(\theta)d\theta
\]

\[
= \int_{\theta_{\min}}^{\theta_{\max}} \left( \int_{\theta_{\min}}^{\theta} \beta_{m\ell}(s)\gamma ds \right) \frac{d}{d\theta}(F(\theta) - 1)) d\theta
\]

\[
= \left( \int_{\theta_{\min}}^{\theta} \beta_{m\ell}(s)\gamma ds \right) (F(\theta) - 1)) \bigg|_{\theta_{\min}}^{\theta_{\max}} - \int_{\theta_{\min}}^{\theta_{\max}} \beta_{m\ell}(\theta)\gamma (F(\theta) - 1)) d\theta
\]

\[
= \int_{\theta_{\min}}^{\theta_{\max}} (1 - F(\theta)) \beta_{m\ell}(\theta)\gamma d\theta.
\]

Therefore, the firm’s objective function is summarized as

\[
\int_{\theta_{\min}}^{\theta_{\max}} \pi_{m\ell}(\theta)f(\theta)d\theta - \bar{\omega},
\]

where the profit contribution of type \( \theta \) is represented as

\[
\pi_{m\ell}(\theta) \equiv \gamma \theta + e_{m\ell}(\theta) - \frac{1 - F(\theta)}{f(\theta)} \beta_{m\ell}(\theta)\gamma - C(e_{m\ell}(\theta)) - \frac{\rho\sigma^2}{2} \beta_{m\ell}(\theta)^2.
\]  

Substituting \( e_{m\ell}(\theta) = \beta_{m\ell}(\theta)k \) into Eq.(8), I conclude that \( \pi_{m\ell}(\theta) = \gamma \theta \) for every \( \theta \leq \theta_{m\ell}^{\min} \), whereas for every \( \theta > \theta_{m\ell}^{\min} \).
\[ \pi_{ml}(\theta) = \gamma \theta + \beta_{ml}(\theta)k - \frac{1 - F(\theta)}{f(\theta)} \beta_{ml}(\theta)\gamma - \frac{\beta_{ml}(\theta)^2 k}{2} - \frac{\rho \sigma^2}{2} \beta_{ml}(\theta)^2 \]

\[ = \gamma \theta + \beta_{ml}(\theta)k \left(1 - \frac{\gamma - 1}{k} F(\theta)\right) - \frac{k + \rho \sigma^2}{2} \beta_{ml}(\theta)^2 \]

\[ = \gamma \theta + \frac{k^2}{k + \rho \sigma^2} \left(1 - \frac{\gamma - 1}{k} F(\theta)\right) - \frac{k + \rho \sigma^2}{2} \left(\frac{k}{k + \rho \sigma^2}\right)^2 \left(1 - \frac{\gamma - 1}{k} F(\theta)\right)^2 \]

\[ = \gamma \theta + \frac{k^2}{2(k + \rho \sigma^2)} \left(1 - \frac{\gamma - 1}{k} F(\theta)\right)^2. \quad (9) \]

Substituting the expression of the inverse hazard rate \((1 - F(\theta))/f(\theta) = \eta(\theta_{\text{max}} - \theta)\) into Eq.(9), I obtain

\[ \pi_{ml}(\theta) = \gamma \theta + \frac{k^2}{2(k + \rho \sigma^2)} \left(1 - \frac{\gamma \eta(\theta_{\text{max}} - \theta)}{k}\right)^2 = \gamma \theta + \frac{(k - \gamma \eta(\theta_{\text{max}} - \theta))^2}{2(k + \rho \sigma^2)}. \quad (10) \]

I am ready to find the maximized firm’s expected profit under a menu of linear compensation schemes. In the following computation, I use the following facts from Lemma 1:

\[ \mathbb{E}(\theta) = \frac{\eta \theta_{\text{max}} + \theta_{\text{min}}}{\eta + 1} \quad \text{and} \quad \text{Var}(\theta) = \frac{\eta^2 \Delta^2}{(\eta + 1)^2 (2\eta + 1)}. \]

Firstly, if there is no bunching in the menu of linear contracts \((\lambda > 1)\), then the firm’s expected profit can be written as

\[ \mathbb{E} \pi_{ml}(\gamma, \eta, \lambda) \mid_{\lambda > 1} \]

\[ = \int_{\theta_{\text{min}}}^{\theta_{\text{max}}} \pi_{ml}(\theta)f(\theta)d\theta - \bar{w} \]

\[ = \gamma \mathbb{E}(\theta) - \bar{w} + \frac{1}{2(k + \rho \sigma^2)} \int_{\theta_{\text{min}}}^{\theta_{\text{max}}} (k - \gamma \eta(\theta_{\text{max}} - \theta))^2 f(\theta)d\theta \]

\[ = \mathbb{E}\pi_{fw}(\gamma, \eta) + \frac{1}{2(k + \rho \sigma^2)} \left((k - \gamma \eta \theta_{\text{max}})^2 + 2\gamma \eta (k - \gamma \eta \theta_{\text{max}})\mathbb{E}(\theta) + \gamma^2 \eta^2 \mathbb{E}(\theta^2)\right) \]

\[ = \mathbb{E}\pi_{fw}(\gamma, \eta) + \frac{(k - \gamma \eta \theta_{\text{max}} + \gamma \eta \mathbb{E}(\theta))^2 + \gamma^2 \eta^2 \text{Var}(\theta)}{2(k + \rho \sigma^2)} \]

\[ = \mathbb{E}\pi_{fw}(\gamma, \eta) + \frac{(k - \gamma \eta (\theta_{\text{max}} - \mathbb{E}(\theta)))^2 + \gamma^2 \eta^2 \text{Var}(\theta)}{2(k + \rho \sigma^2)}. \quad (11) \]

The second term on the right-hand side in Eq.(11) can be written as a form of the variance of \(\theta\). Since \(\theta_{\text{max}} - \mathbb{E}(\theta) = \theta_{\text{max}} - \frac{\theta_{\text{max}} + \theta_{\text{min}}}{\eta + 1} = \frac{\Delta}{\eta + 1}\), it follows that
\[(k - \gamma(\theta_{\text{max}} - \mathbb{E}(\theta)))^2 = \left(k - \frac{\gamma \Delta}{\eta + 1}\right)^2\]

\[= (\gamma \eta \Delta)^2 \left(\frac{k}{\gamma \eta \Delta} - \frac{1}{\eta + 1}\right)^2\]

\[= (\gamma \eta \Delta)^2 \left(\frac{\lambda}{\eta + 1}\right)^2\]

\[= (\gamma \eta \Delta)^2 \left(\frac{(\eta + 1)\lambda - 1}{\eta + 1}\right)^2.\]

Therefore, the maximized expected profit under the menu of linear contracts in which there is no bunching \((\lambda > 1)\) is given by

\[
\mathbb{E}_{\pi_{ml}}(\gamma, \eta, \lambda) \mid \lambda > 1 = \mathbb{E}_{\pi_{lw}}(\gamma, \eta) + \frac{(\gamma \eta \Delta)^2}{2(k + \rho \sigma^2)} \left(\frac{(\eta + 1)\lambda - 1}{\eta + 1}\right)^2 + \frac{\gamma^2 \eta^2 \text{Var}(\theta)}{2(k + \rho \sigma^2)}
\]

\[
= \mathbb{E}_{\pi_{lw}}(\gamma, \eta) + \left(\frac{2(\eta + 1)\gamma^2((\eta + 1)\lambda - 1)^2}{2(k + \rho \sigma^2)}\right) \left(\frac{\eta^2 \Delta^2}{(\eta + 1)^2(2\eta + 1)}\right) + \frac{\gamma^2 \eta^2 \text{Var}(\theta)}{2(k + \rho \sigma^2)}
\]

\[
= \mathbb{E}_{\pi_{lw}}(\gamma, \eta) + \frac{(2\eta + 1)\gamma^2((\eta + 1)\lambda - 1)^2}{2(k + \rho \sigma^2)} \text{Var}(\theta) + \frac{\gamma^2 \eta^2 \text{Var}(\theta)}{2(k + \rho \sigma^2)}.
\]

Finally, I conclude that when there is no bunching in the menu of linear contracts \((\lambda > 1)\), the maximized expected profit is given by Eq.(12).

Proposition 3 (maximized profit under the menu of linear contracts if there is no bunching). If \(\lambda > 1\) then there is no bunching in the menu of linear contracts, and the firm earns the expected profit given in Eq.(12):

\[
\mathbb{E}_{\pi_{ml}}(\gamma, \eta, \lambda) \mid \lambda > 1 = \mathbb{E}_{\pi_{lw}}(\gamma, \eta) + \frac{\gamma^2 \text{Var}(\theta)}{2(k + \rho \sigma^2)} \left((2\eta + 1)((\eta + 1)\lambda - 1)^2 + \eta^2\right).
\]

(12)

It remains to compute the maximized expected profit in which there is a bunching \((\lambda \leq 1)\). If \(\lambda \leq 1\) then there is a bunching in \(\beta_{ml}(\cdot)\), and hence in \(e_{ml}(\cdot)\). That is, less productive workers will not exert any effort. The profit contribution of type \(\theta\) in Eq.(10) is given by

\[
\pi_{ml}(\theta) = \gamma \theta + \beta_{ml}(\theta) k \left(1 - \frac{1 - F(\theta)}{k} f(\theta)\right) k + \frac{\rho \sigma^2}{2} \beta_{ml}(\theta)^2 - \bar{w}
\]

\[
= \begin{cases} 
\gamma \theta - \bar{w} & \text{for } \theta \leq \theta_{\text{ml}}^{\text{min}}, \\
\gamma \theta - \bar{w} + \frac{(k - \gamma \eta (\theta_{\text{max}} - \theta))^2}{2(k + \rho \sigma^2)} & \text{for } \theta > \theta_{\text{ml}}^{\text{min}}.
\end{cases}
\]
Therefore, the principal’s expected profit is as follows:

$$E_{\pi_{MT}}(\gamma, \eta, \lambda) \mid \lambda \leq 1$$

$$= \int_{\theta_{\min}}^{\theta_{\max}} \gamma \theta f(\theta) d\theta + \int_{\theta_{\min}}^{\theta_{\max}} \left[ \gamma \theta + \frac{(k - \gamma \eta(\theta_{\max} - \theta))^2}{2(k + \rho \sigma^2)} \right] f(\theta) d\theta - \bar{w}$$

$$= \int_{\theta_{\min}}^{\theta_{\max}} \gamma \theta f(\theta) d\theta + \int_{\theta_{\min}}^{\theta_{\max}} \frac{(k - \gamma \eta(\theta_{\max} - \theta))^2}{2(k + \rho \sigma^2)} f(\theta) d\theta - \bar{w}$$

$$= \gamma E(\theta) - \bar{w} + \frac{(k - \gamma \eta(\theta_{\max} - E(\theta)))^2 + \gamma^2 \eta^2 \text{Var}(\theta)}{2(k + \rho \sigma^2)}$$

$$- \frac{1}{2(k + \rho \sigma^2)} \int_{\theta_{\min}}^{\theta_{\max}} (k - \gamma \eta(\theta_{\max} - \theta))^2 f(\theta) d\theta$$

$$= \frac{1}{2(k + \rho \sigma^2)} \int_{\theta_{\min}}^{\theta_{\max}} (k - \gamma \eta(\theta_{\max} - \theta))^2 f(\theta) d\theta.$$

Eventually, the right-hand side in Eq.(13) can be simplified as in Proposition 4.

Proposition 4 (maximized profit under the menu of linear contracts if there is a bunching). If $\lambda \leq 1$ then there is a bunching in the menu of linear contracts, and the firm earns the expected profit given in Eq.(14):

$$E_{\pi_{MT}}(\gamma, \eta) \mid \lambda > 1 = \frac{1}{2(k + \rho \sigma^2)} \int_{\theta_{\min}}^{\theta_{\max}} (k - \gamma \eta(\theta_{\max} - \theta))^2 f(\theta) d\theta.$$

Proof. See the Appendix. ■

In this section, I have focused on the menu of linear contracts that is a direct revelation mechanism as a function of private information. In Section 6.2, an economic interpretation of such menu of linear contracts will be discussed. More precisely, I shall investigate the implementability of the menu of linear contracts through a single and anonymous compensation plan as a function of performance.

4 Quota-based Compensation Scheme

In this section, I shall construct a particular piece-wise linear compensation scheme referred to as a quota-based contract. The purpose of the paper is not to derive the fully optimal quota-based compensation scheme. I assume that the firm intends to set a quota-based compensation scheme given in Eq.(15) whose the fixed salary component is predetermined:

$$w(y) = \begin{cases} 
    a & \text{if } y \leq q \\
    a + b(y - q) & \text{if } q < y \leq c \\
    a + b(c - q) & \text{if } y > c.
\end{cases}$$

12A class of quota-based contracts may allow for more than two kinks. For instance, a quota-based contract with bounded payments is considered in Chen and Miller (2009). They employ numerical simulations to compare the relative performance of linear contracts with piece-wise-linear-threshold contracts in the case where the worker chooses actions over time. Their contract can be regarded as a menu of two fixed salaries and a piece-rate contract, described as
\[ w_{qb}(y) = \bar{w} + \beta_{qb} \max\{0, y - y_{qb}\} \quad \text{for some } \beta_{qb} > 0 \text{ and } y_{qb} \geq 0. \] (15)

The coefficient \( \beta_{qb} \) represents the piece-rate paid for each unit of performance beyond the quota \( y_{qb} \). The quota-based contract defined in Eq.(15) is a piece-wise linear compensation plan, so that productive workers benefit from a higher once their output surpass the certain threshold.

It must be emphasized that such contract is not a direct revelation mechanism as a function of type of the worker. The firm cannot condition the contract on the worker’s type. Since the quota \( y_{qb} \) is assumed to be positive, the quota-based compensation scheme can be considered as a “menu” of the fixed-payment contract \( \bar{w} \) and a corresponding piece-rate compensation scheme \( \bar{w}_{pr}(y) = \alpha_{qb} + \beta_{qb}y \) as illustrated in Figure 2. In other words, the original quota-based compensation scheme is the upper boundary of the two simple contracts, \( w_{qb}(y) = \max\{\bar{w}, \bar{w}_{pr}(y)\} \), as depicted in Figure 2.\(^{13}\) Here, the vertical intercept \( \alpha_{qb} \) must satisfy \( \bar{w} - \alpha_{qb} = \beta_{qb}y_{qb} \).

![Figure 2: Quota-based contract as a menu of simple contracts](image)

If the worker of type \( \theta \) chooses the fixed wage contract then the resulting payoff will be \( \text{CE} = \bar{w} \), and so his information rent is zero. On the other hand, the information rent under a piece-rate compensation scheme \( \bar{w}_{qb}(y) \) is given by

\[ R_{pr}(\theta) \triangleq \alpha_{qb} + \beta_{qb}(\gamma\theta + e) - C(e) - \frac{\rho\sigma^2}{2} \beta_{qb}^2 - \bar{w}. \] (16)

Because of the quadratic form of the disutility function \( C(e) = \frac{e^2}{2k} \), the implied effort level is \( e = \beta_{qb}k \). Substituting this effort level into Eq.(16), I have

\[ R_{pr}(\theta) = \alpha_{qb} + \beta_{qb}(\gamma\theta + \beta_{qb}k) - \frac{\beta_{qb}^2k}{2} - \frac{\rho\sigma^2}{2} \beta_{qb}^2 - \bar{w} = \alpha_{qb} + \beta_{qb}\gamma\theta + \frac{k-\rho\sigma^2}{2} \beta_{qb}^2 - \bar{w}. \] (17)

\(^{13}\)If the base component \( \alpha_{qb} \) is strictly higher than the reservation wage \( \bar{w} \), then the quota-level \( y_{qb} \) of output is strictly negative. However, it is not optimal for the firm to set \( \alpha_{qb} > \bar{w} \), and hence I shall pay attention to the case that \( \bar{w} \geq \alpha_{qb} \) or equivalently \( y_{qb} \geq 0 \).
The worker of type $\theta$ will choose the piece-rate incentive scheme only if the expression in Eq.(17) is strictly positive because the information rent under the fixed-wage contract is zero. The information rent of the marginal type $\theta^{q_{\min}}_{\text{min}}$ exactly equals zero, as well as each type $\theta \leq \theta^{q_{\min}}_{\text{min}}$. Every type $\theta > \theta^{q_{\min}}_{\text{min}}$ enjoys a positive information rent. Therefore, the participation constraint under the quota-based contract is described by

$$R_{q_{\min}}(\theta) \triangleq \max \{0, R_{\text{pr}}(\theta)\} \geq 0. \quad (18)$$

Thus, contracts will serve the dual purpose of sorting workers and providing incentives.

Since the information rent is strictly increasing ($\dot{R}_{q_{\min}}(\theta) = b_{q_{\min}}g > 0$) as long as the worker chose the piece-rate contract, there must be a unique value $\theta^{q_{\min}}_{\text{min}}$ such that $R_{q_{\min}}(\theta^{q_{\min}}_{\text{min}}) = 0$. At this moment, the marginal worker type $\theta^{q_{\min}}_{\text{min}}$ does not necessarily belong to the type space. There are several possible cases to be considered. If $\theta^{q_{\min}}_{\text{min}} < \theta_{\text{min}}$, then it must be the case that $R_{\text{pr}}(\theta_{\text{min}}) > 0$, which implies that the firm can improve his expected profit by choosing a lower fixed component $a_{q_{\min}}$ to all of types without altering their actions. This case is ruled out, and thus I pay attention to the case of $\theta_{\text{min}} \leq \theta^{q_{\min}}_{\text{min}}$ or equivalently $R_{\text{pr}}(\theta_{\text{min}}) \leq 0$ with equality only if $\theta_{\text{min}} = \theta^{q_{\min}}_{\text{min}}$. Firstly, no type would choose the fixed wage contract when $\theta_{\text{min}} = \theta^{q_{\min}}_{\text{min}}$. Secondly, there are two subcases to be considered when $\theta_{\text{min}} \leq \theta^{q_{\min}}_{\text{min}}$ or $R_{\text{pr}}(\theta_{\text{min}}) < 0$. If $R_{\text{pr}}(\theta_{\text{min}}) > 0$ then it must be the case that $\theta^{q_{\min}}_{\text{min}}$ belongs to the type space so that $\theta_{\text{min}} < \theta^{q_{\min}}_{\text{min}} < \theta_{\text{max}}$. On the other hand, if $R_{\text{pr}}(\theta_{\text{max}}) \leq 0$ then it must be the case that $\theta^{q_{\min}}_{\text{min}} \geq \theta_{\text{max}}$. In the latter situation, the quota-based contract is dominated by the sole fixed wage contract, however, this case never happens. In what follows, I may restrict my attention to the case that $\theta_{\text{min}} \leq \theta^{q_{\min}}_{\text{min}} < \theta_{\text{max}}$.

I assume that the marginal worker type $\theta^{q_{\min}}_{\text{min}}$ prefers the fixed wage contract $\bar{w}$ to the piece-rate contract $\tilde{w}_{q_{\min}}(y)$, although he is indifferent between the two simple contracts. Eventually, the effort level under the menu of simple contracts is the following:

$$e_{q_{\min}}(\theta) = \begin{cases} \beta_{q_{\min}}g & \text{if } R_{\text{pr}}(\theta) > 0, \\ 0 & \text{if } R_{\text{pr}}(\theta) \leq 0. \end{cases}$$

The corresponding output plan is given by $y_{q_{\min}}(\theta) = \gamma \theta + e_{q_{\min}}(\theta)$.

Suppose now that $\theta_{\text{min}} < \theta^{q_{\min}}_{\text{min}}$ or equivalently $R_{\text{pr}}(\theta_{\text{min}}) < 0$. From the binding participation constraint $R_{\text{pr}}(\theta^{q_{\min}}_{\text{min}}) = 0$ for the marginal worker type $\theta^{q_{\min}}_{\text{min}}$, the fixed component of the quota-based compensation scheme is determined as

$$a_{q_{\min}} = \bar{w} - \left(\beta_{q_{\min}}g^{q_{\min}}_{q_{\min}} + \frac{k - \rho v^2}{2} \theta^{q_{\min}}_{q_{\min}} \right). \quad (19)$$

Therefore, the profit contribution of each type $\theta > \theta^{q_{\min}}_{\text{min}}$ of the worker whose effort level is $e_{q_{\min}}(\theta) = \beta_{q_{\min}}g$ (who chose the piece-rate contract) becomes:

$${}^{14}$See Proposition 6. The quota-based contract is strictly superior to the fixed-wage contract.
\[ \pi_{qb}(\theta) = y_{qb}(\theta) - w_{pr}(y_{qb}(\theta)) = \gamma \theta + e_{qb}(\theta) - (\alpha_{qb} + \beta_{qb}(\gamma \theta + e_{qb}(\theta))) \]

\[ = (1 - \beta_{qb})(\gamma \theta + \beta_{qb} k) - \delta_{qb} \]

\[ = (1 - \beta_{qb})(\gamma \theta + \beta_{qb} k) + \beta_{qb} \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}^2 - \bar{w}. \]

On the other hand, the profit contribution of each type \( \theta < \theta_{\min}^{qb} \) of the worker whose effort level is \( e_{qb}(\theta) = 0 \) (who chose the fixed wage scheme) becomes:

\[ \pi_{qb}(\theta) = \gamma \theta - \bar{w}. \]

Now, I rewrite the firm’s expected profit as a function of the pay-performance sensitivity \( \beta_{qb} \) and the marginal worker type \( \theta_{\min}^{qb} \):

\[ \mathbb{E}\pi_{qb}(\beta_{qb}, \theta_{\min}^{qb}) = \int_{\theta_{\min}^{qb}}^{\theta_{\max}^{qb}} [\gamma \theta - \bar{w}] f(\theta) d\theta \]

\[ + \int_{\theta_{\min}^{qb}}^{\theta_{\max}^{qb}} \left[ (1 - \beta_{qb})(\gamma \theta + \beta_{qb} k) + \beta_{qb} \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}^2 - \bar{w} \right] f(\theta) d\theta \]

\[ = \gamma \int_{\theta_{\min}^{qb}}^{\theta_{\max}^{qb}} \theta f(\theta) d\theta + (1 - \beta_{qb}) \gamma \int_{\theta_{\min}^{qb}}^{\theta_{\max}^{qb}} \theta f(\theta) d\theta \]

\[ + \left( (1 - \beta_{qb}) \beta_{qb} k + \beta_{qb} \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}^2 \right) \int_{\theta_{\min}^{qb}}^{\theta_{\max}^{qb}} f(\theta) d\theta - \bar{w} \]

\[ = \gamma \mathbb{E}(\theta) - \bar{w} - \beta_{qb} \gamma \int_{\theta_{\min}^{qb}}^{\theta_{\max}^{qb}} \theta f(\theta) d\theta \]

\[ + \left( (1 - \beta_{qb}) \beta_{qb} k + \beta_{qb} \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}^2 \right) (1 - F(\theta_{\min}^{qb})). \] (20)

The first-order condition that characterizes the optimal component of the piece-rate incentive scheme are the following. The first-order conditions with respect to \( \beta_{qb} \) is as follows:

\[ 0 = \frac{\partial \mathbb{E}\pi_{qb}(\beta_{qb}, \theta_{\min}^{qb})}{\partial \beta_{qb}} \]

\[ = -\gamma \int_{\theta_{\min}^{qb}}^{\theta_{\max}^{qb}} \theta f(\theta) d\theta + \left( k - 2 \beta_{qb} k + \gamma \theta_{\min}^{qb} + (k - \rho \sigma^2) \beta_{qb} \right) (1 - F(\theta_{\min}^{qb})) \]

\[ = -\gamma \int_{\theta_{\min}^{qb}}^{\theta_{\max}^{qb}} \theta f(\theta) d\theta + \left( k + \gamma \theta_{\min}^{qb} - (k + \rho \sigma^2) \beta_{qb} \right) (1 - F(\theta_{\min}^{qb})). \]

This yields that
The parameter $\theta_{\min}^b$ and hence a sufficient condition for that the incentive sensitivity $S$ at

\[
\frac{\gamma f_{\theta_{\min}^b}^{\theta_{\max}^b} \theta f(\theta) d\theta}{1 - F(\theta_{\min}^b)} = k + \gamma \theta_{\min}^{\theta_{\min}^b} - (k + \rho \sigma^2) \beta_{\min}^b.
\]

(22)

Solving Eq.(22) for $\beta_{\min}^b$ to get

\[
\beta_{\min}^b = \frac{k + \gamma \theta_{\min}^{\theta_{\min}^b} - \gamma f_{\theta_{\min}^b}^{\theta_{\max}^b} \theta f(\theta) d\theta / (1 - F(\theta_{\min}^b))}{k + \rho \sigma^2}
\]

\[
= \frac{k - \gamma \left( f_{\theta_{\min}^b}^{\theta_{\max}^b} \theta f(\theta) d\theta / (1 - F(\theta_{\min}^b)) - \theta_{\min}^{\theta_{\min}^b} \right)}{k + \rho \sigma^2}
\]

\[
= \frac{k}{k + \rho \sigma^2} \left( 1 - \frac{\gamma}{k} \left( \mathbb{E}[\theta | \theta \geq \theta_{\min}^b] - \theta_{\min}^{\theta_{\min}^b} \right) \right).
\]

(23)

(24)

Here, I would like to check under what condition the incentive sensitivity $\beta_{\min}^b$ is positive indeed. Since $\mathbb{E}[\theta | \theta \geq \hat{\theta}] - \hat{\theta} = \frac{\eta \theta_{\max}^b + \hat{\theta}}{\eta + 1} - \hat{\theta} = \frac{\eta (\theta_{\max}^b - \hat{\theta})}{\eta + 1}$ by Eq.(5), it follows that $1 - \frac{\gamma}{k} \left( \mathbb{E}[\theta | \theta \geq \hat{\theta}] - \hat{\theta} \right)$ is minimized at $\hat{\theta} = \theta_{\min}$. I see that the expression in the right-hand side in Eq.(24) evaluated at $\theta_{\min}^{\theta_{\min}^b} = \theta_{\min}$ is positive if $k > \gamma (\mathbb{E}[\theta] - \theta_{\min})$ holds. Notice that

\[
\gamma (\mathbb{E}[\theta] - \theta_{\min}) = \gamma \left( \frac{\eta \theta_{\max}^b + \theta_{\min}}{\eta + 1} - \theta_{\min} \right) = \gamma \left( \frac{\eta \theta_{\max}^b + \theta_{\min} - (\eta + 1) \theta_{\min}}{\eta + 1} \right) = \gamma \eta \Delta
\]

and hence a sufficient condition for that the incentive sensitivity $\beta_{\min}^b$ is positive restricts an upper bound for the parameter $\lambda$:

\[
k > \gamma (\mathbb{E}[\theta] - \theta_{\min}) \iff \lambda = \frac{k}{\gamma \eta \Delta} > \frac{1}{\eta + 1}.
\]

(25)

Next, the first-order condition with respect to $\theta_{\min}^{\theta_{\min}^b}$ is the following:

\[
0 = \frac{\partial \pi_{\min}^b(\beta_{\min}^b, \theta_{\min}^{\theta_{\min}^b})}{\partial \theta_{\min}^{\theta_{\min}^b}}
\]

\[
= \beta_{\min}^b \gamma \theta_{\min}^{\theta_{\min}^b} f(\theta_{\min}^{\theta_{\min}^b}) + \beta_{\min}^b \gamma (1 - F(\theta_{\min}^{\theta_{\min}^b})) - \left( 1 - \beta_{\min}^b \right) \beta_{\min}^b + \beta_{\min}^b \gamma \theta_{\min}^{\theta_{\min}^b} + \frac{k - \rho \sigma^2}{2} \beta_{\min}^b^2 \theta_{\min}^{\theta_{\min}^b} f(\theta_{\min}^{\theta_{\min}^b})
\]

\[
= \beta_{\min}^b \gamma (1 - F(\theta_{\min}^{\theta_{\min}^b})) - \left( 1 - \beta_{\min}^b \right) \beta_{\min}^b + \frac{k - \rho \sigma^2}{2} \beta_{\min}^2 \theta_{\min}^{\theta_{\min}^b} f(\theta_{\min}^{\theta_{\min}^b}).
\]

(26)

(27)

Dividing the both sides of Eq.(26) by $\beta_{\min}^b f(\theta_{\min}^{\theta_{\min}^b})$, I have

\[
\gamma \left( \frac{1 - F(\theta_{\min}^{\theta_{\min}^b})}{f(\theta_{\min}^{\theta_{\min}^b})} \right) = (1 - \beta_{\min}^b) k + \frac{k - \rho \sigma^2}{2} \beta_{\min}^b = \frac{2k - 2 \beta_{\min}^b k + \beta_{\min}^b k - \rho \sigma^2 \beta_{\min}^b}{2} = \frac{2k - (k + \rho \sigma^2) \beta_{\min}^b}{2}.
\]
and then,

\[ 2\gamma\eta(\theta_{\max} - \theta_{\min}^{gb}) = 2k - (k + \rho\sigma^2)\beta_{gb}. \] (28)

Substituting Eq.(23) into Eq.(28) to get

\[ 2\gamma\eta(\theta_{\max} - \theta_{\min}^{gb}) = 2k - \left( k - \gamma \left( \int_{\theta_{\min}^{gb}}^{\theta_{\max}^{gb}} \theta f(\theta)d\theta/(1 - F(\theta_{\min}^{gb})) - \theta_{\min}^{gb} \right) \right) \]

\[ = k + \gamma \left( \int_{\theta_{\min}^{gb}}^{\theta_{\max}^{gb}} \theta f(\theta)d\theta/(1 - F(\theta_{\min}^{gb}))) - \theta_{\min}^{gb} \right), \]

and then,

\[ \frac{2\gamma\eta(\theta_{\max} - \theta_{\min}^{gb}) - k}{\gamma} = \int_{\theta_{\min}^{gb}}^{\theta_{\max}^{gb}} \theta f(\theta)d\theta/(1 - F(\theta_{\min}^{gb}))) - \theta_{\min}^{gb}. \]

Here, by Lemma 1, the truncated mean over the interval \([\theta_{\min}^{gb}, \theta_{\max}^{gb}]\) is expressed as

\[ \int_{\theta_{\min}^{gb}}^{\theta_{\max}^{gb}} \theta f(\theta)d\theta/(1 - F(\theta_{\min}^{gb}))) = \frac{\eta\theta_{\max} + \theta_{\min}^{gb}}{\eta + 1}. \]

Then,

\[ \frac{2\gamma\eta(\theta_{\max} - \theta_{\min}^{gb}) - k}{\gamma} = \frac{\eta\theta_{\max} + \theta_{\min}^{gb}}{\eta + 1} - \theta_{\min}^{gb} = \frac{\eta(\theta_{\max} - \theta_{\min}^{gb})}{\eta + 1}. \] (29)

Solving Eq.(29) for the marginal type \(\theta_{\min}^{gb}\) to get\(^{15}\)

\[ \theta_{\min}^{gb} = \theta_{\max} - \frac{k(\eta + 1)}{\gamma\eta(2\eta + 1)}. \]

As I assumed, the marginal type \(\theta_{\min}^{gb}\) cannot be lower than the lowest type \(\theta_{\min}\). The requirement \(\theta_{\min}^{gb} \geq \theta_{\min}\) is guaranteed if \(\frac{2\eta + 1}{\eta + 1} \geq \lambda\) holds.\(^{16}\) That is, the ratio \(\lambda = \frac{k}{\gamma\eta}\) is supposed to be bounded from

\(^{15}\)Eq.(29) yields that \((\theta_{\max} - \theta_{\min}^{gb}) \left( 2\eta - \frac{\eta}{\eta + 1} \right) = (\theta_{\max} - \theta_{\min}^{gb})\frac{2\eta + 1}{\eta + 1} - (\theta_{\max} - \theta_{\min}^{gb})\frac{\eta}{\eta + 1} = \frac{k}{\gamma\eta}.\) Solving this for \(\theta_{\max} - \theta_{\min}^{gb}\) to get \(\theta_{\max} - \theta_{\min}^{gb} = \frac{k}{\gamma\eta} + \frac{1}{\eta + 1} = \frac{k\eta + 1}{\eta(\eta + 1)}\). Here, \(\theta_{\min}^{gb} > \theta_{\min}^{gb} = \theta_{\max} - \frac{k}{\gamma\eta}\) for any \(\eta > 0\) because \(\theta_{\min}^{gb} = \theta_{\max} - \frac{k\eta + 1}{\eta(\eta + 1)} > \theta_{\max} - \frac{k}{\gamma\eta} = \theta_{\min}^{gb}\). Therefore, the sorting effect under the quota-based contract is stronger than that under the menu of linear contracts.

\(^{16}\)Notice that \(\theta_{\min}^{gb} \geq \theta_{\min} \iff \Delta = \theta_{\max} - \theta_{\min} \geq \theta_{\max} - \theta_{\min}^{gb} = \frac{k\eta + 1}{\eta(\eta + 1)}\) is guaranteed if \(\frac{2\eta + 1}{\eta + 1} \geq \lambda\) holds.
above, as is depicted in Figure 3. Recall that if the condition \( \lambda > \frac{1}{\eta + 1} \) does not hold, the problem would be uninteresting because it is optimal for the firm to set the fixed-payment contract only. Putting together the two restrictions on the set of parameters, the relevant set of parameters in terms of \((\eta, \lambda)\) so that the quota-based contract is well-defined is denoted by

\[
A_{qb} = \left\{ (\eta, \lambda) \mid \frac{2\eta + 1}{\eta + 1} \geq \lambda > \frac{1}{\eta + 1} \right\}.
\]

![Figure 3: Admissible parameters for quota-based compensation scheme](image)

It remains to calculate the expression of the pay-performance sensitivity \( \beta_{qb} \). Substituting \( \theta_{\text{max}} - \theta_{\text{min}}^{qb} = \frac{k(\eta + 1)}{\gamma(2\eta + 1)} \) into Eq. (23) to get

\[
\beta_{qb} = \frac{k - \gamma \left( \int_{\theta_{\text{min}}^{qb}}^{\theta_{\text{max}}^{qb}} \theta f(\theta)d\theta / (1 - F(\theta_{\text{min}}^{qb})) - \theta_{\text{min}}^{qb} \right)}{k + \rho \sigma^2}
\]

(30)

because the expression in the round brackets in Eq. (30) is written as

\[
\int_{\theta_{\text{min}}^{qb}}^{\theta_{\text{max}}^{qb}} \theta f(\theta)d\theta / (1 - F(\theta_{\text{min}}^{qb})) - \theta_{\text{min}}^{qb} = \frac{\eta(\theta_{\text{max}}^{qb} - \theta_{\text{min}}^{qb})}{\eta + 1} = \frac{k(\eta + 1)}{\gamma(2\eta + 1)}.\]

It turns out that the pay-performance sensitivity \( \beta_{qb} \) of the quota-based contract is positive indeed. In addition, the marginal worker type \( \theta_{\text{min}}^{qb} \) would produce the amount \( y_{qb}(\theta_{\text{min}}^{qb}) = \gamma \theta_{\text{min}}^{qb} \) because he chose the fixed-wage contract. It turns out that any type \( \theta > \theta_{\text{min}}^{qb} \) has an incentive to produce more than the quota \( y_{qb} \). Recall that \( \bar{w} - \alpha_{qb} = \beta_{qb} y_{qb} \), as shown in Figure 2. Also, \( \bar{w} - \alpha_{qb} = \beta_{qb} \gamma \theta_{\text{min}}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}^2 \) from Eq. (19). Combining these equations, I have
\[ y_{qb} = \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}. \]

Finally, I conclude that any type \( \theta > \theta_{\min}^{qb} \) (who prefers the piece-rate contract to the fixed-wage contract) produces the output \( y_{qb}(\theta) = \gamma \theta + \beta_{qb} > \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb} = y_{qb} \) because \( \frac{k - \rho \sigma^2}{2} \beta_{qb} < \frac{k}{2} \beta_{qb} < \beta_{qb} \) holds.\(^\text{17}\)

Proposition 5 (incentive-intensity under the quota-based contract). Suppose that \( \frac{2 \eta + 1}{\eta + 1} \geq \lambda \). The quota-based contract considered as the menu of fixed wage and piece rate contracts is given by

\[ w_{qb}(y) = \bar{w} + \beta_{qb} \max \{0, y - y_{qb}\}, \quad \text{where} \]

\[ \beta_{qb} = \frac{2 \eta k}{(2 \eta + 1)(k + \rho \sigma^2)} \quad \text{and} \quad y_{qb} = \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}. \]

The corresponding marginal type is given by \( \theta_{\min}^{qb} = \theta_{\max} - \frac{k(\eta + 1)}{\eta(2 \eta + 1)} \theta_{\min} \) with equality only if \( \frac{2 \eta + 1}{\eta + 1} = \lambda \).\(^\text{18}\)

From Eq.(20) and Eq.(21), the maximized profit under the quota-based contract as a function of the incentive sensitivity \( \beta_{qb} \) and the marginal worker type \( \theta_{\min}^{qb} \) can be written as

\[
\mathbb{E} \pi_{qb}(\beta_{qb}, \theta_{\min}^{qb}) \\
= \gamma \mathbb{E}(\theta) - \bar{w} \\
- \beta_{qb} \gamma \int_{\theta_{\min}^{qb}}^{\theta_{\max}^{qb}} \theta f(\theta) d\theta + \left( (1 - \beta_{qb}) \beta_{qb} k + \beta_{qb} \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}^2 \right) (1 - F(\theta_{\min}^{qb})) \\
= \mathbb{E} \pi_{fw}(\gamma, \eta) - \beta_{qb} \gamma (1 - F(\theta_{\min}^{qb})) \left( \frac{\eta \theta_{\max}^{qb} + \theta_{\min}^{qb}}{\eta + 1} \right) \\
+ \left( (1 - \beta_{qb}) \beta_{qb} k + \beta_{qb} \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}^2 \right) (1 - F(\theta_{\min}^{qb})) \\
= \mathbb{E} \pi_{fw}(\gamma, \eta) + \left( -\beta_{qb} \gamma \frac{\eta \theta_{\max}^{qb} + \theta_{\min}^{qb}}{\eta + 1} + (1 - \beta_{qb}) \beta_{qb} k + \beta_{qb} \gamma \theta_{\min}^{qb} + \frac{k - \rho \sigma^2}{2} \beta_{qb}^2 \right) (1 - F(\theta_{\min}^{qb})). \quad (31) \]

The expression of the inside the round brackets in Eq.(31) is written as

\(^\text{17}\)For \( \theta < \theta_{\min}^{qb} \), the output \( y(\theta) = \gamma \theta \) could be higher or lower than the quota \( y_{qb} \) because the sign of \( k - \rho \sigma^2 \) is indeterminate.\(^\text{18}\) If \( \frac{2 \eta + 1}{\eta + 1} = \lambda \) then the marginal type \( \theta_{\min}^{qb} \) coincides with the lowest type \( \theta_{\min} \): if \( \frac{2 \eta + 1}{\eta + 1} = \lambda \) then \( \theta_{\min}^{qb} = \theta_{\max} - \frac{k(\eta + 1)}{\eta(2 \eta + 1)} = \theta_{\max} - \frac{\lambda}{\frac{\eta + 1}{\eta}} = \theta_{\min} \).
Therefore, the maximized expected profit under the quota-based contract will be written as

\[
\mathbb{E}\pi_{qb}(\theta_{\text{max}}) = \mathbb{E}\pi_{fb}(\gamma, \eta) + \frac{2\eta^2 k^2}{(2\eta + 1)^2(k + \rho \sigma^2)} \left( \frac{\theta_{\text{max}} - \theta_{\text{min}}^b}{\Delta} \right)^{\frac{1}{\eta}}
\]

\[
= \mathbb{E}\pi_{fb}(\gamma, \eta) + \left( \frac{2(\eta + 1)^{\frac{2\eta + 1}{\eta}}}{(2\eta + 1)^{\frac{2\eta + 1}{\eta}}} \left( \frac{\Delta^2 \eta^2}{(\eta + 1)^2(2\eta + 1)} \right) \right)
\]

\[
= \mathbb{E}\pi_{fb}(\gamma, \eta) + \frac{2k}{(2\eta + 1)^{\frac{2\eta + 1}{\eta}}} \left( k + \frac{\rho \sigma^2}{1} \right)^{\frac{1}{\eta}} \text{Var}(\theta)
\]

\[
= \mathbb{E}\pi_{fb}(\gamma, \eta) + \frac{2\Delta^2}{(2\eta + 1)^{\frac{2\eta + 1}{\eta}}} \left( \frac{\Delta^2 \eta^2}{(\eta + 1)^2(2\eta + 1)} \right)
\]

Proposition 6 (maximized profit under the quota-based contract). If \(\frac{2\eta + 1}{\eta + 1} \geq \lambda\) then the maximized expected
profit under the quota-based contract is given by

$$E \pi_{qb}(\gamma, \eta, \lambda) = E \pi_{fw}(\gamma, \eta) + \frac{2\lambda^{2\eta+1}}{(k + \rho^2)(2\eta + 1)} \gamma^2 \eta^2 \text{Var}(\theta).$$

(32)

In particular, if the parameter $\lambda$ attains the upper bound $2^{2\eta+1} = \lambda$, then the maximized expected profit can be simplified to the following:

$$E \pi_{qb}(\gamma, \eta, \frac{2\eta+1}{2\eta+1}) = E \pi_{fw}(\gamma, \eta) + \frac{2(2\eta + 1)\gamma^2 \eta^2}{k + \rho^2} \text{Var}(\theta).$$

Notice that I have obtained the maximized expected profit under the "restricted" quota-based contract with the predetermined fixed-wage component given in Eq.(15). An "unrestricted" quota-based contract may earn more than the maximized expected profit in Eq.(32), however, it is shown that such restricted quota-based contract is enough to secure a substantial gain relative to the menu of linear contracts in Section 5.1.

5 Comparison of Performances of Simple Contracts

I begin by analyzing the performance of the quota-based contract constructed in Section 4. I focus on the incremental gain achieved under the quota-based contract beyond the fixed-wage contract. The performance measure I employ is the ratio of the incremental gain under the quota-based contract divided by the incremental gain under the menu of linear contracts. I shall derive an analytical expression for such relative performance measure in order to provide explicit upper and lower bounds on the relative performance. I adopt the same strategy to examine the relative performance of the optimal piece-rate contract. My findings below may provide an explanation for the use of a piece-wise-linear-threshold contract in practice although several researchers have pointed out negative effects of quota(s).

5.1 Performance of quota-based compensation scheme

The subsection analyzes the performance measurement of the quota-based contract as a menu of simple contracts derived in the previous section. Tractable and analytical upper and lower bounds on such measurement can be explicitly derived under the specific distribution function of private information I consider. This is the incremental gain relative to the optimal fixed wage defined by $[E \pi_{qb} - E \pi_{fw}] / [E \pi_{ml} - E \pi_{fw}]$, rather than the ratio of $E \pi_{qb}$ to $E \pi_{ml}$.

19In the literature on the procurement contract, Rogerson (2003) and Chu and Sappington (2007) demonstrate that simple binary menus secure a substantial share of the gains achieved by the fully optimal contract. Rogerson (2003) uses a menu of a fixed price contract and a cost-reimbursement contract (FPCR), whereas Chu and Sappington (2007) uses (LCSCR), a menu of a linear cost-sharing contract (i.e., two-part tariff) and cost-reimbursement contract.
\[
\phi_{qb}(\eta, \lambda) = \frac{E\pi_{qb}(\gamma, \eta, \lambda) - E\pi_{fw}(\gamma, \eta)}{E\pi_{me}(\gamma, \eta, \lambda) - E\pi_{fw}(\gamma, \eta)}
\]

\[
\phi_{qb}(\eta, \lambda) = \begin{cases} 
\frac{4\eta^2(\eta + 1) - \frac{2\eta + 1}{\eta + 1} \lambda - \frac{2\eta + 1}{\eta + 1}}{(2\eta + 1)((\eta + 1)(\lambda - 1)^2 + \eta^2)} & \text{if } \frac{2\eta + 1}{\eta + 1} \geq \lambda > 1, \\
2 \left( \frac{\eta + 1}{2\eta + 1} \right)^\frac{2\eta + 1}{\eta + 1} & \text{if } \frac{1}{\eta + 1} < \lambda \leq 1.
\end{cases}
\]

The function \( \phi_{qb}(\eta, \lambda) \) is continuous in \( \lambda \). The relevant set of parameters \( A_{qb} \) can be decomposed into two regions, as depicted in Figure 4.

Firstly, I consider the case in which there is no-bunching in the menu of linear contracts \( \lambda > 1 \). The relative incremental gain will be calculated as

\[
\phi_{qb}(\eta, \lambda) |_{\lambda > 1} = \frac{4\eta^2(\eta + 1) - \frac{2\eta + 1}{\eta + 1} \lambda - \frac{2\eta + 1}{\eta + 1}}{(2\eta + 1)((\eta + 1)(\lambda - 1)^2 + \eta^2)}.
\]

I shall examine whether \( \phi_{qb}(\eta, \lambda) |_{\lambda > 1} \) is increasing or decreasing in \( \lambda \) for each \( \eta \). The derivative of the relevant part of the function \( \phi_{qb}(\eta, \lambda) |_{\lambda > 1} \) is the following:

---

\(^{20}\) Notice that when \( \frac{2\eta + 1}{\eta + 1} \geq \lambda > 1 \),

\[
\lim_{\lambda \to 1} \phi_{qb}(\eta, \lambda) = \frac{4\eta^2(\eta + 1) - \frac{2\eta + 1}{\eta + 1} \lambda - \frac{2\eta + 1}{\eta + 1}}{2(2\eta + 1)((\eta + 1)(\lambda - 1)^2 + \eta^2)} = 2 \left( \frac{\eta + 1}{2\eta + 1} \right)^\frac{2\eta + 1}{\eta + 1}.
\]
\[
\frac{\partial}{\partial \lambda} \left( \frac{\Lambda^{2n+1}}{(2\eta + 1)((\eta + 1)\Lambda - 1)^2 + \eta^2} \right) \\
= \frac{(2\eta + 1)\Lambda^{n+1}}{\eta((2\eta + 1)((\eta + 1)\Lambda - 1)^2 + \eta^2)} - \frac{2(\eta + 1)(2\eta + 1)((\eta + 1)\Lambda - 1)\Lambda^{2n+1}}{((2\eta + 1)((\eta + 1)\Lambda - 1)^2 + \eta^2)^2} \\
= \frac{(2\eta + 1)\Lambda^{n+1}}{\eta((2\eta + 1)((\eta + 1)\Lambda - 1)^2 + \eta^2)} \left( 1 - \frac{2(\eta + 1)((\eta + 1)\Lambda - 1)\Lambda}{(2\eta + 1)((\eta + 1)\Lambda - 1)^2 + \eta^2} \right). 
\]  

(36)

The expression inside the round brackets in Eq.(36) is strictly positive in \( \mathcal{A}_{qB} \mid \lambda > 1 \) because \((2\eta + 1)((\eta + 1)\Lambda - 1)^2 + \eta^2 = (\eta + 1)^2(\Lambda - 1)^2 > 0 \). Recall that \((\eta, \lambda)\) is restricted to \( \frac{2\eta+1}{\eta+1} \geq \Lambda > 1 \) now. Therefore, given \( \eta \), the upper bound within the relevant region \( \mathcal{A}_{qB} \mid \lambda > 1 \) of parameter values is attained as \( \lambda \to \frac{2\eta+1}{\eta+1} \), while the lower bound as \( \lambda \to 1 \). The value of the upper bound on \( \phi_{qB}(\eta, \lambda) \mid \lambda > 1 \) is written as

\[
\phi_{qB} \left( \eta, \frac{2\eta + 1}{\eta + 1} \right) \mid \frac{2\eta+1}{\eta+1} \geq \lambda > 1 = \frac{4\eta^2(\eta + 1)\frac{2\eta+1}{\eta+1} \left( \frac{2\eta+1}{\eta+1} \right)^\frac{2\eta+1}{\eta+1}}{(2\eta + 1)\frac{2\eta+1}{\eta+1} \left( (2\eta + 1)(2\eta + 1 - 1)^2 + \eta^2 \right)} \\
= \frac{4\eta^2(\eta + 1)\frac{2\eta+1}{\eta+1} \left( \frac{2\eta+1}{\eta+1} \right)^\frac{2\eta+1}{\eta+1}}{\eta^2(4(2\eta + 1) + 1)} \\
= \frac{4(2\eta + 1)}{4(2\eta + 1) + 1} = \frac{8\eta + 4}{8\eta + 5} \in (0.8, 1).
\]

On the other hand, the value of the lower bound on \( \phi_{qB}(\eta, \lambda) \mid \lambda > 1 \) is written as

\[
\phi_{qB}(\eta, 1) \mid \frac{2\eta+1}{\eta+1} \geq \lambda > 1 = \frac{4\eta^2(\eta + 1)\frac{2\eta+1}{\eta+1} \left( \frac{2\eta+1}{\eta+1} \right)^\frac{2\eta+1}{\eta+1}}{(2\eta + 1)\frac{2\eta+1}{\eta+1} \left( (2\eta + 1)\eta^2 + \eta^2 \right)} = \frac{4\eta^2(\eta + 1)\frac{2\eta+1}{\eta+1} \left( \frac{2\eta+1}{\eta+1} \right)^\frac{2\eta+1}{\eta+1}}{2\eta^2(\eta + 1)(2\eta + 1) \frac{2\eta+1}{\eta+1}^\frac{2\eta+1}{\eta+1}} = 2 \left( \frac{\eta + 1}{2\eta + 1} \right)^\frac{2\eta+1}{\eta+1} \\
= \frac{2}{e} \approx 0.73 \text{ as } \eta \to 0,
\]

since \( \lim_{\eta \to 0} \left( \frac{\eta + 1}{2\eta + 1} \right)^\frac{2\eta+1}{\eta+1} = 1/e \approx 0.367879 \). These observations are summarized as follows:

\[
\begin{align*}
\sup \left[ \phi_{qB}(\eta, \lambda) \mid \frac{2\eta+1}{\eta+1} \geq \lambda > 1 \right] &= \phi_{qB} \left( \eta, \frac{2\eta+1}{\eta+1} \right) = \frac{8\eta+4}{8\eta+5}, \\
\inf \left[ \phi_{qB}(\eta, \lambda) \mid \frac{2\eta+1}{\eta+1} \geq \lambda > 1 \right] &= \phi_{qB}(\eta, 1) = 2 \left( \frac{\eta + 1}{2\eta + 1} \right)^\frac{2\eta+1}{\eta+1}.
\end{align*}
\]

(37)

It turns out that the quota-based contract performs well if there is no bunching in the menu of linear contracts, as shown in Figure 5.
Proposition 7 (performance of the quota-based contract). Suppose that there is no bunching in the menu of linear contracts \((\lambda > 1)\). The relative incremental gain varies from \(2 \left( \frac{\eta + 1}{2\eta + 1} \right)^{\frac{\eta + 1}{\eta}} \) to \(\frac{8\eta + 4}{8\eta + 5}\) for each first-order stochastic dominance shift parameter \(\eta\). Approximately, the lower bound is higher than \(\frac{2}{e} \approx 0.73\).

Table 1 summarizes the upper and the lower bounds on the relative incremental gain for several values of \(\eta\). The performance of the quota-based contract varies and increases as the distribution of the productivity of workers becomes more favorable to the firm. In particular, the quota-based contract can secure more than 88 percent of the gain secured under the menu of linear contracts when the worker’s type is distributed uniformly.

<table>
<thead>
<tr>
<th>(\eta)</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.4</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>upper bounds</td>
<td>0.8</td>
<td>0.827586</td>
<td>0.848485</td>
<td>0.878049</td>
<td>0.923077</td>
<td>0.952381</td>
<td>0.972973</td>
<td>0.988235</td>
</tr>
<tr>
<td>lower bounds</td>
<td>0.735759</td>
<td>0.76799</td>
<td>0.793139</td>
<td>0.829897</td>
<td>0.888889</td>
<td>0.929516</td>
<td>0.959267</td>
<td>0.982021</td>
</tr>
</tbody>
</table>

Table 1: The relative incremental gain of the quota-based contract \(\left( \frac{\mathbb{E}\pi_{qb} - \mathbb{E}\pi_{fw}}{\mathbb{E}\pi_{mt} - \mathbb{E}\pi_{fw}} \right) \) when there is no bunching in the menu of linear contracts

To end this subsection, it remains to consider the case in which there is a bunching in the menu of linear contracts \((\lambda \leq 1)\). The relative incremental gain will be calculated as

\[
\phi_{qb}(\eta, \lambda) \left| \frac{1}{\eta + 1} < \lambda \leq 1 \right. = 2 \left( \frac{\eta + 1}{2\eta + 1} \right)^{\frac{\eta + 1}{\eta}}.
\] (38)
Notice that this is equal to \( \phi_{qb}(\eta, 1) \). It turns out that the above gain is independent of \( \lambda \).

Proposition 8 (performance of the quota-based contract). When there is a bunching in the menu of linear contracts, the relative incremental gain is given by

\[
2 \left( \frac{\eta + 1}{2\eta + 1} \right)^{\frac{\eta + 1}{\eta}}
\]

for each \((\eta, l)\). The lowest relative incremental gain \(2/e \approx 0.73\) is achieved when \(\eta\) tends to zero.

From Propositions 7 and 8, I conclude that the quota-based contract can always secure a substantial portion (at least \(2/e \approx 73\) percent) of the incremental gain secured by the optimal menu of linear contracts over the entire region of the relevant parameter values.

5.2 Performance of piece-rate incentive scheme

For comparison, I shall consider the situation in which the firm can offer a single piece-rate contract. As a corollary of Proposition 6, the expected profit to be maximized under the piece-rate incentive scheme can be obtained by taking the limit of Eq.(20) and Eq.(21) as \(q_{\min}^{qb} \to \theta_{\min}^{pr}:

\[
E_{\pi_{pr}}(\beta_{pr}) = \gamma E(\theta) - \bar{w} - \beta_{pr} \gamma E(\theta) + (1 - \beta_{pr}) \beta_{pr} k + \beta_{pr} \gamma \theta_{\min} + \frac{k - \rho \sigma^2}{2} \beta_{pr}^2
\]

\[
= (1 - \beta_{pr}) \gamma E(\theta) + (1 - \beta_{pr}) \beta_{pr} k + \beta_{pr} \gamma \theta_{\min} + \frac{k - \rho \sigma^2}{2} \beta_{pr}^2 - \bar{w},
\]

and the corresponding incentive sensitivity is given by

\[
\beta_{pr} = \frac{k - \gamma (E(\theta) - \theta_{\min})}{k + \rho \sigma^2}.
\]
Eq. (40) can be obtained from the first-order condition for Eq. (39) as well. Then, the maximized expected profit can be written as

\[ \mathbb{E} \pi_{pr}(\beta_{pr}) = \gamma \mathbb{E}(\theta) - \omega - \beta_{pr} \gamma \mathbb{E}(\theta) + (1 - \beta_{pr}) \beta_{pr} k + \beta_{pr} \gamma \theta_{min} + \frac{k - \rho \sigma^2}{2} \beta_{pr}^2 \]

\[ = \mathbb{E} \pi_{fw}(\gamma, \eta) + \beta_{pr} k - \beta_{pr} \gamma \mathbb{E}(\theta) - \beta_{pr}^2 k + \beta_{pr} \gamma \theta_{min} + \frac{k - \rho \sigma^2}{2} \beta_{pr}^2 \]

\[ = \mathbb{E} \pi_{fw}(\gamma, \eta) + (k - \gamma (\mathbb{E}(\theta) - \theta_{min})) \beta_{pr} - \frac{k + \rho \sigma^2}{2} \beta_{pr}^2 \]

\[ = \mathbb{E} \pi_{fw}(\gamma, \eta) + \frac{(k - \gamma (\mathbb{E}(\theta) - \theta_{min}))^2}{2(k + \rho \sigma^2)} - \frac{k + \rho \sigma^2}{2} \frac{(k - \gamma (\mathbb{E}(\theta) - \theta_{min}))^2}{(k + \rho \sigma^2)^2} \]

\[ = \mathbb{E} \pi_{fw}(\gamma, \eta) + \frac{(k - \gamma (\mathbb{E}(\theta) - \theta_{min}))^2}{2(k + \rho \sigma^2)}. \]

The second term on the right-hand side can be written as a form of the variance of \( \theta \). Since \( \mathbb{E}(\theta) - \theta_{min} = \frac{\eta \theta_{max} + \theta_{min}}{\eta + 1} - \theta_{min} = \frac{\eta (\theta_{max} - \theta_{min})}{\eta + 1} = \frac{\eta \Delta}{\eta + 1} \), it follows that \( (k - \gamma (\mathbb{E}(\theta) - \theta_{min}))^2 = (\gamma \eta \Delta)^2 \left( \frac{(\eta + 1)\lambda - 1}{\eta + 1} \right)^2 \) as before. Therefore, the maximized expected profit under the optimal piece-rate incentive scheme is given by

\[ \mathbb{E} \pi_{pr}(\beta_{pr}) = \mathbb{E} \pi_{fw}(\gamma, \eta) + \frac{(\gamma \eta \Delta)^2}{2(k + \rho \sigma^2)} \left( \frac{(\eta + 1)\lambda - 1}{\eta + 1} \right)^2 \]

\[ = \mathbb{E} \pi_{fw}(\gamma, \eta) + \frac{(2\eta + 1) \gamma^2 ((\eta + 1)\lambda - 1)^2}{2(k + \rho \sigma^2)} - \frac{\eta^2 \Delta^2}{(\eta + 1)^2(2\eta + 1)} \]

\[ = \mathbb{E} \pi_{fw}(\gamma, \eta) + \frac{(2\eta + 1) \gamma^2 ((\eta + 1)\lambda - 1)^2}{2(k + \rho \sigma^2)} \text{Var}(\theta). \]

Proposition 9 (maximized profit under the piece-rate contract). The maximized expected profit under the optimal piece-rate incentive scheme is given by

\[ \mathbb{E} \pi_{pr}(\gamma, \eta, \lambda) = \mathbb{E} \pi_{fw}(\gamma, \eta) + \frac{(2\eta + 1) \gamma^2 ((\eta + 1)\lambda - 1)^2}{2(k + \rho \sigma^2)} \text{Var}(\theta). \]

The pay-performance sensitivity \( \beta_{pr} \) of the optimal piece-rate incentive scheme is supposed to be positive. This restricts the upper bound on the marginal productivity \( \gamma \) in the response function. By Eq. (25), the incentive intensity \( \beta_{pr} \) is strictly positive for every \( (\eta, \lambda) \in \mathcal{A}_{gb} \) indeed.

This subsection analyzes the following performance measurement of the optimal piece-rate incentive scheme. Now consider how the optimal piece-rate contract performs relative to the menu of linear contracts. The relevant incremental gain relative to the optimal fixed-wage contract is defined by \( \frac{\mathbb{E} \pi_{pr} - \mathbb{E} \pi_{fw}}{\mathbb{E} \pi_{fw} - \mathbb{E} \pi_{fw}} \). Rewrite the relative incremental gain as a function of \( \lambda \) and \( \eta \):

\[ \frac{\mathbb{E} \pi_{pr} - \mathbb{E} \pi_{fw}}{\mathbb{E} \pi_{fw} - \mathbb{E} \pi_{fw}} = \frac{k - \gamma \theta_{min} \gamma \mathbb{E}(\theta)}{k + \rho \sigma^2} \]

\[ = \frac{k - \gamma \theta_{min} \gamma \mathbb{E}(\theta)}{k + \rho \sigma^2}. \]
\[
\phi_{pr}(\eta, \lambda) = \frac{\mathbb{E}\pi_{pr}(\gamma, \eta, \lambda) - \mathbb{E}\pi_{w}(\gamma, \eta)}{\mathbb{E}\pi_{w}(\gamma, \eta) - \mathbb{E}\pi_{w}(\gamma, \eta)} = \begin{cases} 
\frac{(2\eta + 1)((\eta + 1)\lambda - 1)^2}{(2\eta + 1)((\eta + 1)\lambda - 1)^2 + \eta^2} & \text{if } \frac{2\eta + 1}{\eta + 1} \geq \lambda > 1, \\
\frac{(2\eta + 1)((\eta + 1)\lambda - 1)^2}{2(\eta + 1)\eta^2} & \text{if } \frac{1}{\eta + 1} < \lambda \leq 1.
\end{cases}
\]

Notice that the function \(\phi_{pr}(\eta, \lambda)\) is continuous.\(^{22}\) In order to find the upper and lower bounds on the relative incremental gain, I shall check the monotonicity of the function \(\phi_{pr}(\eta, \lambda)\) with respect to \(\lambda\) for each \(\eta\). The derivative of the function \(\phi_{pr}(\eta, \lambda)\) with respect to \(\lambda\) is the following:

\[
\frac{\partial \phi_{pr}(\eta, \lambda)}{\partial \lambda} = \begin{cases} 
\frac{2\eta^2(2\eta + 1)((\eta + 1)\lambda - 1)}{(\eta + 1)((1 - \lambda)^2 + 2\eta^2\lambda^2 + \eta(1 - \lambda)(1 - 3\lambda))} & \text{if } \lambda > 1, \\
\frac{(2\eta + 1)\lambda - \frac{3\eta + 1}{\eta} (2\eta + 1 - (\eta + 1)\lambda)((\eta + 1)\lambda - 1)}{2(\eta + 1)\eta^3} & \text{if } \frac{1}{\eta + 1} < \lambda \leq 1.
\end{cases}
\]

Firstly, suppose that there is no bunching in the menu of linear contracts \((\lambda > 1)\). Since \(\eta > 0\), it follows that \((\eta + 1)\lambda > 1\), and so \(\phi_{pr}(\eta, \lambda)\) is strictly increasing in \(\lambda\) as long as \(\lambda > 1\):

\[
\text{sgn} \frac{\partial \phi_{pr}(\eta, \lambda)}{\partial \lambda} \bigg|_{\lambda > 1} = \text{sgn}[(\eta + 1)\lambda - 1] = \text{plus}.
\]

Therefore, the upper bound is attained as \(\lambda \to \frac{2\eta + 1}{\eta + 1}\), while the lower bound as \(\lambda \to 1.\(^{23}\) In other words, \(\phi_{pr}(\eta, \lambda)\) varies from \(\frac{2\eta + 1}{\eta + 1}\) to \(\frac{8\eta + 4}{8\eta + 5}\) monotonically for each \(\eta\) when there is no bunching in the menu of linear contracts \((\lambda > 1)\):

\[
\left\{ \begin{array}{c}
\sup \left[ \phi_{pr}(\eta, \lambda) \mid \frac{2\eta + 1}{\eta + 1} \geq \lambda > 1 \right] = \phi_{pr} \left( \eta, \frac{2\eta + 1}{\eta + 1} \right) = \frac{8\eta + 4}{8\eta + 5}, \\
\inf \left[ \phi_{pr}(\eta, \lambda) \mid \frac{2\eta + 1}{\eta + 1} \geq \lambda > 1 \right] = \phi_{pr}(\eta, 1) = \frac{2\eta + 1}{2(\eta + 1)}.
\end{array} \right.
\]

\(^{22}\)When \(\lambda > 1\), \(\phi_{pr}(\eta, \lambda) \bigg|_{\lambda > 1} = \frac{(2\eta + 1)((\eta + 1)\lambda - 1)^2}{(2\eta + 1)(\eta + 1)\lambda^2 + \eta^2} \bigg|_{\lambda > 1} = \frac{(2\eta + 1)^2}{2(\eta + 1)^2 + \eta^2} = \frac{2\eta + 1}{2(\eta + 1)}.

\[^{23}\]If I drop the restriction \(\frac{2\eta + 1}{\eta + 1} \geq \lambda\) on the parameter \(\lambda\), I have \(\sup \phi_{pr}(\eta, \lambda) \mid \lambda > 1 = \lim_{\lambda \to 0} \phi_{pr}(\eta, \lambda) = 1.\)
I thus have proved the following proposition.

**Proposition 10 (performance of the piece-rate contract).** Suppose that there is no bunching in the menu of linear contracts ($\lambda > 1$). The relative incremental gain varies from $\frac{2\eta + 1}{2(\eta + 1)}$ to $\frac{8\eta + 4}{8\eta + 5}$ for each first-order stochastic dominance shift parameter $\eta$. In particular, if the worker heterogeneity is distributed uniformly, then the optimal piece-rate incentive scheme secures more than 75 percent of the gain under the menu of linear contracts.

The lower bound on the relative incremental gain is equal to 75% when the heterogeneity of the workers is distributed uniformly. However, in order to secure more substantial gain, the first-order stochastic dominance shift parameter $\eta$ must be much higher than the unity. For instance, the lower bound $\phi_{pr}(\eta, 1)$ is greater than 90% for any $\eta > 4$.

Next, I shall consider the case in which a single piece-rate contract is available for the firm and there is a bunching in the menu of linear contracts ($\lambda \leq 1$). There is a marked difference between the two simple contracts when there is a bunching in the menu of linear contracts. I need to verify that the function $\phi_{pr}(\eta, \lambda)$ is monotone with respect to $\lambda$ for each $\eta$ when $\lambda \leq 1$. Eq.(42) below tells us that when $\lambda$ varies from $\frac{1}{\eta + 1}$ to 1, the sign of the derivative of the relative incremental gain with respect to $\lambda$ is represented as the sign of the product of two terms: $2\eta + 1 - (\eta + 1)\lambda$ and $(\eta + 1)\lambda - 1$. Regarding the first term $2\eta + 1 - (\eta + 1)\lambda$, I see that $\lambda \leq 1$ implies $(\eta + 1)\lambda \leq \eta + 1$. Then, $(\eta + 1)\lambda - 1 - 2\eta \leq \eta - 2\eta = -\eta \leq 0$. I conclude that $2\eta + 1 - (\eta + 1)\lambda \geq 0$ with equality only if $\frac{2\eta + 1}{\eta + 1} = \lambda$ for every $(\eta, \lambda) \in \mathcal{A}_{qh}$.

On the other hand, the second term $(\eta + 1)\lambda - 1$ is strictly positive for every $(\eta, \lambda) \in \mathcal{A}_{pr}$. In summary, the sign of the partial derivative of the expression for $\phi(\eta, \lambda)$ with respect to $\eta$ is determined in the following way.

$$\text{sgn}\left(\frac{\partial \phi_{pr}(\eta, \lambda)}{\partial \lambda}\right)_{\frac{1}{\eta + 1} < \lambda \leq 1} = \text{sgn}\left(2\eta + 1 - (\eta + 1)\lambda)((\eta + 1)\lambda - 1)\right) = \text{plus.}$$  

**Figure 7:** $\phi_{pr}(\eta, \lambda)$ when there is no bunching in the menu of linear contracts

---

24 Notice that $(\eta, \lambda) = (0, 1)$ is ruled out by the restriction $\lambda > \frac{1}{\eta + 1}$.
According to Eq.(42), the relative incremental gain increases in $l$ for each $h$. Therefore, the upper bound is attained as $\lambda \to 1$, while the lower bound as $\lambda \to \frac{1}{\eta + 1}$. In other words, $\phi_{pr}(\eta, \lambda)$ varies from 0 to $\frac{2\eta + 1}{2(\eta + 1)}$ monotonically for each $\eta$ when there is a bunching in the menu of linear contracts ($\lambda \leq 1$):

$$\begin{align*}
\sup \left\{ \phi_{pr}(\eta, \lambda) \mid \frac{1}{\eta + 1} < \lambda \leq 1 \right\} &= \phi_{pr}(\eta, 1) = \frac{2\eta + 1}{2(\eta + 1)}, \\
\inf \left\{ \phi_{pr}(\eta, \lambda) \mid \frac{1}{\eta + 1} < \lambda \leq 1 \right\} &= \phi_{pr} \left( \eta, \frac{1}{\eta + 1} \right) = 0.
\end{align*}$$

(43)

These boundaries are depicted in Figure 8. The intuition behind the poor performance of the piece-rate contract when $\lambda = \frac{k}{\eta \Delta}$ converges to $\frac{1}{\eta + 1}$ stems from the fact that there is no incentive effect because the corresponding incentive intensity tends to zero.

![Figure 8: $\phi_{pr}(\eta, \lambda)$ when there is a bunching in the menu of linear contracts](image)

Table 2 summaries the upper and the lower bounds on the relative incremental gain of the piece-rate contract for several values of $\eta$. If there is a bunching in the menu of linear contracts and private information is distributed uniformly, the piece-rate incentive scheme can secure at most 75 percent of the gain secured by the optimal menu of linear contracts.

<table>
<thead>
<tr>
<th>$\eta$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.4</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>upper bounds</td>
<td>0.545455</td>
<td>0.583333</td>
<td>0.642857</td>
<td>0.75</td>
<td>0.833333</td>
<td>0.9</td>
<td>0.954545</td>
<td>0.97619</td>
</tr>
<tr>
<td>lower bounds</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2: The relative incremental gain of the piece-rate contract ($[E\pi_{pr} - E\pi_{fw}] / [E\pi^O - E\pi_{fw}]$) when there is a bunching in the menu of linear contracts.
5.3 Comparison of two simple contracts

In this subsection, I shall compare the relative performances analyzed in Sections 5.1 and 5.2. Suppose that there is no bunching in the menu of linear contracts \((\lambda > 1)\). According to Eq.(37) and Eq.(41), the remarkable fact is that both simple contracts, the quota-based contract and the piece-rate contract, attain the same upper bound when there is no bunching in the menu of linear contracts, that is, there is no difference between the two simple contracts for the "best-case" scenario:

\[
\sup \left[ \phi_{qb}(\eta, \lambda) \mid \frac{2\eta+1}{\eta+1} \geq\lambda > 1 \right] = \frac{8\eta + 4}{8\eta + 5} = \sup \left[ \phi_{pr}(\eta, \lambda) \mid \frac{2\eta+1}{\eta+1} \geq\lambda > 1 \right].
\]

For the "worse-case" scenario, the lower bound on the relative performance of the quota-based contract is superior to that of the piece-rate contract for the entire class of the distribution functions considered in the paper:

\[
\inf \left[ \phi_{qb}(\eta, \lambda) \mid \frac{2\eta+1}{\eta+1} \geq\lambda > 1 \right] = \frac{2\left( \frac{\eta+1}{2\eta+1} \right)^{\frac{2\eta+1}{\eta}}}{2\left( \frac{2\eta+1}{2\eta+1} \right)^{\frac{2\eta+1}{\eta}}} = 4 \left( \frac{\eta + 1}{2\eta + 1} \right)^{\frac{2\eta+1}{\eta}} \approx 1.47152 \quad \text{as} \quad \eta \to 0.
\]

Next, suppose that there is a bunching in the menu of linear contracts \((\lambda \leq 1)\). According to Eq.(38) and Eq.(43), needless to say, the quota-based contract is always superior to the piece-rate contract unless \(\eta\) tends to infinity, and the advantage of the quota-based contract is quite noticeable as \(\eta\) tends to zero:

\[
\frac{\phi_{qb}(\eta, \lambda) \mid \frac{1}{\eta+1} < \lambda \leq 1}{\sup \left[ \phi_{pr}(\eta, \lambda) \mid \frac{1}{\eta+1} < \lambda \leq 1 \right]} = 2 \left( \frac{\eta+1}{2\eta+1} \right)^{\frac{2\eta+1}{\eta}} = 4 \left( \frac{\eta + 1}{2\eta + 1} \right)^{\frac{2\eta+1}{\eta}}.
\]

6 Discussions

The purpose of this section is twofold. First, I shall compare the incentive sensitivities under alternative contracts. The literature on pay for performance has argued the effects of changing the compensation method on performance or productivity. I provide an answer to this question in Section 6.1.

It is important to understand the intuition behind the design of simple contracts. The theoretical benchmark I consider here is the optimal menu of linear contracts, a continuum of type-dependent linear contracts. Each contract is linear in the performance measure, but it could be non-linear in the reported type. It is worth comparing all of compensation schemes studied in the paper in the same domain because the optimal menu of linear contracts depends on an announcement of type or private information, although the simple contracts depend on output. In Section 6.2, I provide a procedure to transform the optimal menu of linear contracts as a direct revelation mechanism into an anonymous nonlinear compensation scheme as an indirect mechanism.
6.1 Incentive effects

I have derived the incentive sensitivities under alternative three contracts in Sections 3 and 4. The marginal types under the several optimal contracts are summarized in Table 3. Any type above the corresponding marginal type will exert positive effort. Less productive workers will choose the fixed wage contract if it is available. Since $\theta_{\text{min}}^\ell < \theta_{\text{min}}^{qb}$ always holds, I may conclude that the quota-based contract has a stronger sorting effect than the menu of linear contracts.

<table>
<thead>
<tr>
<th>Type of contract</th>
<th>Marginal type</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piece-rate incentive scheme or Linear contract</td>
<td>$\theta_{\text{min}}$</td>
<td></td>
</tr>
<tr>
<td>Quota-based compensation scheme</td>
<td>$\theta_{\text{min}}^{qb} = \theta_{\text{max}} - \frac{k(q_{\text{max}} + \frac{1}{2} + 1)}{q_{\text{max}}}$</td>
<td>$\theta_{\text{min}}^{qb} \geq \theta_{\text{min}} \iff \frac{2q_{\text{max}} + 1}{q_{\text{max}} + 1} \geq 1$</td>
</tr>
<tr>
<td>Menu of linear contracts</td>
<td>$\theta_{\text{min}}^\ell = \theta_{\text{max}} - \frac{k}{q_{\text{max}}}$</td>
<td>$\theta_{\text{min}}^\ell \geq \theta_{\text{min}} \iff \lambda \geq 1$</td>
</tr>
</tbody>
</table>

Table 3: Marginal types under alternative contracts

In order to evaluate incentive effects of the contracts discussed above, it suffices to pay attention to the performance pay-sensitivities because the forms of effort and sales response functions yield that production is strictly increasing in $\theta$:

$$y_j(\theta) = \gamma \theta + e_j(\theta) = \gamma \theta + \beta_j(\theta)k, \quad \text{where } j = m, qb, pr.$$  \hspace{2cm} (44)

Table 4 shows that all of incentive intensities measure the extent of unctrollable risk and incentives as in the early theoretical literature on incentives. All of them share the usual property such as it is smaller if the agent is more risk averse or there is more uncertainty or marginal disutility increases more quickly. These formulas show that how these contracts depend on the details of the distribution of the heterogeneity of workers.

Since both $\beta_{pr}$ and $\beta_{qb}$ are positive over the relevant set of parameters $A_{qb}$, it follows that the greater the measurement error $\sigma^2$, the weaker the strength of the incentive. Similar to the incentive intensity of the menu of linear contracts. Furthermore, it is obvious that $\beta_{qb}$ converges to $\beta_{pr}$ as $\theta_{\text{min}}^{qb} \rightarrow \theta_{\text{min}}$ or equivalently as $\frac{2q_{\text{max}} + 1}{q_{\text{max}} + 1} \rightarrow \lambda$.

<table>
<thead>
<tr>
<th>Type of contract</th>
<th>Pay-performance sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piece-rate incentive scheme or Linear contract</td>
<td>$\beta_{pr} = \frac{k(\gamma E(\theta) - \theta_{\text{min}})}{k + \rho r^2} = \frac{k}{k + \rho r^2} \left(1 - \frac{\gamma}{k}(E(\theta) - \theta_{\text{min}})\right)$</td>
</tr>
<tr>
<td>Quota-based compensation scheme</td>
<td>$\beta_{qb} = \frac{2\gamma k}{(2q_{\text{max}} + 1)(k + \rho r^2)} = \frac{k}{k + \rho r^2} \left(1 - \frac{\gamma}{k}(E(\theta) \geq \theta_{\text{min}}^{qb}) - \theta_{\text{min}}^{qb}\right)$</td>
</tr>
<tr>
<td>Menu of linear contracts</td>
<td>$\beta_{\text{min}}(\theta) = \max \left{ \frac{k}{k + \rho r^2} \left(1 - \frac{\gamma}{k}(1 - F(\theta))\right), 0 \right}$</td>
</tr>
</tbody>
</table>

Table 4: Incentive effects
Compared to the pay performance sensitivity of the complete contract \( \beta_c = \frac{k}{k + \rho \sigma^2} = \frac{1}{1 + \rho \sigma^2} \) with \( k = \frac{1}{c} \), the incentive sensitivity \( \beta_{ml}(\theta) \) and the corresponding effort \( \epsilon_{ml}(\theta) = \beta_{ml}(\theta)k \) of the menu of linear contracts are downward-distorted except at the top.\(^{25} \) That is, only the top type \( \theta_{\text{max}} \) puts forth the efficient level of effort:\(^{26} \)

\[
\beta_{ml}(\theta) = \frac{k}{k + \rho \sigma^2} \left( 1 - \frac{\gamma}{k} \frac{1 - F(\theta)}{f(\theta)} \right) \leq \beta_c \quad \text{with equality only if } \theta = \theta_{\text{max}}.
\]

It will be shown that both simple contracts have weaker incentive effects than the menu of linear contracts. More precisely, I shall show that the incentive sensitivities under those simple contracts are expressed in terms of the expectation of the incentive sensitivity of the menu of linear contracts.

For any interval \([\hat{\theta}, \theta_{\text{max}}]\), the expectation of the pay-performance sensitivity rule \( \beta_{ml}(\theta) \) of the menu of linear contracts is given by

\[
\int_{\hat{\theta}}^{\theta_{\text{max}}} \beta_{ml}(\theta) f(\theta) \, d\theta = \int_{\hat{\theta}}^{\theta_{\text{max}}} \frac{k}{k + \rho \sigma^2} \left( 1 - \frac{\gamma \eta (\theta_{\text{max}} - \theta)}{k} \right) f(\theta) \, d\theta
\]

\[
= \frac{k}{k + \rho \sigma^2} \int_{\hat{\theta}}^{\theta_{\text{max}}} \left( 1 - \frac{\gamma \eta \theta_{\text{max}}}{k} \right) f(\theta) \, d\theta + \frac{\gamma \eta}{k + \rho \sigma^2} \int_{\hat{\theta}}^{\theta_{\text{max}}} \theta f(\theta) \, d\theta
\]

\[
= \frac{k - \gamma \eta \theta_{\text{max}}}{k + \rho \sigma^2} (1 - F(\hat{\theta})) + \frac{\gamma \eta}{k + \rho \sigma^2} \mathbb{E} [\theta \mid \theta \geq \hat{\theta}] (1 - F(\hat{\theta}))
\]

\[
= \frac{1 - F(\hat{\theta})}{k + \rho \sigma^2} \left( k - \gamma \eta \theta_{\text{max}} + \frac{\gamma \eta (\theta_{\text{max}} + \hat{\theta})}{\eta + 1} \right)
\]

\[
= \frac{1 - F(\hat{\theta})}{k + \rho \sigma^2} \left( k - \gamma \eta \left( \theta_{\text{max}} - \hat{\theta} \right) \right)
\]

\[
= \left( \frac{1 - F(\hat{\theta})}{k + \rho \sigma^2} \right) \left( 1 - \frac{\gamma}{k} \left( \mathbb{E} [\theta \mid \theta \geq \hat{\theta}] - \hat{\theta} \right) \right)
\]

because by definition of the expectation of \( \theta \) over the interval \([\hat{\theta}, \theta_{\text{max}}]\), I have

\[
\frac{\gamma \eta (\theta_{\text{max}} - \hat{\theta})}{\eta + 1} = \gamma \left( \theta_{\text{max}} + \hat{\theta} - \hat{\theta} \right) = \gamma \left( \int_{\hat{\theta}}^{\theta_{\text{max}}} \theta f(\theta) \, d\theta \right) = \gamma \left( \mathbb{E} [\theta \mid \theta \geq \hat{\theta}] - \hat{\theta} \right).
\]

\(^{25} \) The contract under complete information can be derived as follows. The firm maximizes profit for type \( \theta \): \( y - w(y) = y - (a + \beta y) = (1 - \beta)y - a = (1 - \beta)(\gamma \theta + c + \epsilon) - a \). Since the expectation of the noise \( \epsilon \) is zero, the firm maximizes \( (1 - \beta(\theta))(\gamma \theta + c(\theta)) - a \) subject to \( a(\theta) + \beta(\theta)(\gamma \theta + c(\theta)) - C(c(\theta)) - \rho \sigma^2 \beta^2(\theta)^2/2 \geq 0 \) and \( c(\theta) \in \text{argmax} \left[ a(\theta) + \beta(\theta)(\gamma \theta + c) - C(c) - \rho \sigma^2 \beta^2(\theta)^2/2 \mid \epsilon \geq 0 \right] \). The first-order condition yields that \( \beta_c = \frac{k}{k + \rho \sigma^2} \). In particular, when the disutility function of the agent is given by \( C(c) = \frac{1}{2} \sigma^2 \), substituting \( k = 1/c \) yields the familiar expression for the optimal pay-performance sensitivity: \( \beta_c = \frac{1}{1 + \sigma^2} \). See Gibbons (1987, pp.2-3) for a general strictly convex disutility function.

\(^{26} \) Gibbons (1987) obtained the same no-distortion-at-the-top result without restricting the class of mechanisms when the agent is risk-neutral.
Therefore, the average of the incentive effects under the menu of linear contracts over the interval \([\hat{\theta}, \theta_{\text{max}}]\) is given by Eq.(47):

\[
\mathbb{E}[\beta_{\text{mt}}(\theta) \mid \theta \geq \hat{\theta}] \triangleq \int_{\hat{\theta}}^{\theta_{\text{max}}} \beta_{\text{mt}}(\theta) f(\theta) d\theta = \frac{k}{k + \rho \sigma^2} \left( 1 - \frac{\gamma}{k} \left( \mathbb{E}[\theta \mid \theta \geq \hat{\theta}] - \hat{\theta} \right) \right).
\]  

(47)

In particular, when \(\hat{\theta} = \theta^{qb}_{\text{min}}\),

\[
\mathbb{E}[\beta_{\text{mt}}(\theta) \mid \theta \geq \theta^{qb}_{\text{min}}] = \frac{k}{k + \rho \sigma^2} \left( 1 - \frac{\gamma}{k} \left( \mathbb{E}[\theta \mid \theta \geq \theta^{qb}_{\text{min}}] - \theta^{qb}_{\text{min}} \right) \right) = \beta_{q}.
\]

Similarly, when \(\hat{\theta} = \theta_{\text{min}}\),

\[
\mathbb{E}[\beta_{\text{mt}}(\theta)] = \mathbb{E}[\beta_{\text{mt}}(\theta) \mid \theta \geq \theta_{\text{min}}] = \frac{k}{k + \rho \sigma^2} \left( 1 - \frac{\gamma}{k} (\mathbb{E}(\theta) - \theta_{\text{min}}) \right) = \beta_{pr}.
\]

It remains to compare the two incentive sensitivities, \(\beta_{q} \) and \(\beta_{pr} \). Differentiating Eq.(47) with respect to the lower boundary \(\hat{\theta}\), I see that the sign of the derivative is strictly positive as long as \(\eta > 0\): since \(\mathbb{E}[\theta \mid \theta \geq \hat{\theta}] - \hat{\theta} = \frac{\eta(\theta_{\text{max}} - \hat{\theta})}{\eta + 1}\), it follows that the average of the personalized incentive intensities of the menu of linear contracts over the interval \([\hat{\theta}, \theta_{\text{max}}]\) increases with the marginal type:

\[
\frac{\partial}{\partial \hat{\theta}} \mathbb{E}[\beta_{\text{mt}}(\theta) \mid \theta \geq \hat{\theta}] = \frac{\partial}{\partial \hat{\theta}} \left( \frac{1}{k + \rho \sigma^2} \left( k - \frac{\gamma \eta (\theta_{\text{max}} - \hat{\theta})}{\eta + 1} \right) \right) = \frac{\gamma \eta}{(k + \rho \sigma^2)(\eta + 1)} > 0.
\]

Such monotonicity of \(\mathbb{E}[\theta \mid \theta \geq \hat{\theta}]\) in \(\hat{\theta}\), together with \(\theta_{\text{min}}^{qb} \geq \theta_{\text{min}}\), implies that \(\beta_{q} \geq \beta_{pr}\) with equality only if \(2\eta + 1 = \lambda\) or equivalently \(\theta_{\text{min}}^{qb} = \theta_{\text{min}}\). I have obtained the following result.

Proposition 11 (incentive effects). Compared with the menu of linear contracts, the two simple contracts, the quota-based contract and the piece-rate contract, are characterized as follows:

\[
\beta_{pr} = \mathbb{E}[\beta_{\text{mt}}(\theta)] \leq \mathbb{E}[\beta_{\text{mt}}(\theta) \mid \theta \geq \theta_{\text{min}}^{qb}] = \beta_{q} < \beta_{c},
\]

where the weak inequality is satisfied with equality only if \(2\eta + 1 = \lambda\).

Figure 9 illustrates how the simple contracts distort incentives, relative to the menu of linear contracts and the complete-information contract. Intuitively, under the quota-based contract, productive workers benefit from a higher rate.
As a corollary, it turns out that the simple linear contract and the menu of linear contracts yield the same average production:

\[
E[y_{pr}(q)] = \int_{q_{\min}}^{q_{\max}} [\gamma \theta + \beta_{pr}k] f(q)dq = \int_{q_{\min}}^{q_{\max}} \gamma \theta f(q)d\theta + E[\beta_{mt}(\theta)]k \\
= \int_{q_{\min}}^{q_{\max}} [\gamma \theta + \beta_{mt}(\theta)k] f(q)d\theta = E[y_{mt}(q)],
\]

because \( \beta_{mt}(\theta) = 0 \) for every \( \theta \leq \theta_{\text{mt}}^{\min} \). On the other hand, a switch to the quota-based contract reduces the average production:
\[
\mathbb{E}[y_{qb}(\theta)] = \int_{\theta_{\min}}^{\theta_{\min}^{q_{S}}} \gamma \theta f(\theta) d\theta + \int_{\theta_{\min}^{q_{S}}}^{\theta_{\max}^{q_{S}}} [\gamma \theta + \beta_{qb} k] f(\theta) d\theta
\]
\[
= \gamma \mathbb{E}(\theta) + \beta_{qb} k (1 - F(\theta_{\min}^{q_{S}}))
\]
\[
= \gamma \mathbb{E}(\theta) + \mathbb{E}[\beta_{ml}(\theta) | \theta \geq \theta_{\min}^{q_{S}}] k (1 - F(\theta_{\min}^{q_{S}}))
\]
\[
= \gamma \mathbb{E}(\theta) + \left( \int_{\theta_{\min}^{q_{S}}}^{\theta_{\max}^{q_{S}}} \beta_{ml}(\theta) f(\theta) d\theta \right) \frac{k (1 - F(\theta_{\min}^{q_{S}}))}{1 - F(\theta_{\min}^{q_{S}})}
\]
\[
= \gamma \mathbb{E}(\theta) + \int_{\theta_{\min}^{q_{S}}}^{\theta_{\max}^{q_{S}}} \beta_{ml}(\theta) k f(\theta) d\theta
\]
\[
< \gamma \mathbb{E}(\theta) + \int_{\theta_{\min}^{q_{S}}}^{\theta_{\max}^{q_{S}}} \beta_{ml}(\theta) k f(\theta) d\theta
\]
\[
= \int_{\theta_{\min}^{q_{S}}}^{\theta_{\max}^{q_{S}}} [\gamma \theta + \beta_{ml}(\theta) k] f(\theta) d\theta
\]
\[
= \mathbb{E}[y_{ml}(\theta)],
\]

where the last strict inequality follows from the fact that \(\theta_{\min}^{q_{S}} < \theta_{\min}^{q_{S}}\) and \(\beta_{ml}(\theta) > 0\) over the interval \((\theta_{\min}^{q_{S}}, \theta_{\min}^{q_{S}})\). Therefore, the average production under the quota-based contract is less than that of the linear contract and the menu of linear contracts. Following the results in the previous section, there is no clear relationship between productivity and profitability. Under the quota-based contract, workers are less motivated on average, however, it seem to be more profitable than the piece-rate contract.

### 6.2 The near optimality of quota-based compensation scheme

In this subsection, I discuss the implementability of menu of linear contracts via a single scheme. I have discussed the menu of linear contracts as a benchmark because most incentive schemes observed in practice are linear. The main purpose of this subsection is to show that such menu of linear contracts can be implemented via a single and interpretable compensation scheme. The result here provides an intuition for why a single quota-based contract performs well better than a single piece-rate contract. I would like to consider why a simple piece-wise linear contract sufficiently performs well instead of the menu of linear contracts. Denote the range of output under the menu of linear contracts by \(Y = [y_{ml}(\theta_{\min}), y_{ml}(\theta_{\max})]\). It will be shown that the worker’s optimization problem can be reformulated as a simple optimization problem in terms of the wage schedule \(t : Y \rightarrow \mathbb{R}\) in the sense that the production rule and the information rent under the menu of linear contracts are achieved under the unanimous wage schedule \(t : Y \rightarrow \mathbb{R}\) defined below.

Recall that the worker of type \(\theta\) faces the following certainty equivalent under the truth-telling:27

---

27Laffont and Martimort (2002, p.376) note that "To reconstruct the indirect mechanism \(T^{SB}(q)\) from the direct mechanism \(\{t^{SB}(\theta), q^{SB}(\theta)\}\) is rather easy. Indeed, we have \(T^{SB}(q) = t^{SB}(q^{SB}(q))\)." However, in my setting, it seems to be annoying to analyze the second derivative of such indirect mechanism.
\[
CE_{ml}(\theta) = \alpha_{ml}(\theta) + \beta_{ml}(\theta)(\gamma \theta + \epsilon_{ml}(\theta)) - \frac{\rho \sigma^2}{2} \beta_{ml}(\theta)^2 - C(\epsilon_{ml}(\theta))
\]
\[
= \left\{ \alpha_{ml}(\theta) - \frac{\rho \sigma^2}{2} \beta_{ml}(\theta)^2 + \beta_{ml}(\theta) y_{ml}(\theta) \right\} - C(y_{ml}(\theta) - \gamma \theta). \tag{48}
\]

The expression inside the brackets in Eq. (48) can be considered as the payment under the linear contract adjusted with risk premium. When the following two requirements in Eq. (49) are satisfied for a menu of linear contracts, I can say that the wage schedule \( t: \mathcal{Y} \to \mathbb{R} \) implements the menu of linear contracts. Here, for each type \( \theta \), the required effort \( \epsilon = y - \gamma \theta \) should be non-negative.

\[
\begin{dcases}
\text{Requirement 1: } y_{ml}(\theta) \in \arg\max [t(y) - C(y - \gamma \theta) \mid y \geq \gamma \theta], \\
\text{Requirement 2: } CE_{ml}(\theta) = \max [t(y) - C(y - \gamma \theta) \mid y \geq \gamma \theta].
\end{dcases} \tag{49}
\]

A version of the taxation principle in this context is the following. Proposition 12 holds not only for the optimal menu of linear contracts derived in Proposition 1.

Proposition 12 (Taxation Principle). For every feasible menu of linear contracts \( \langle \alpha_{ml}(\cdot), \beta_{ml}(\cdot) \rangle \), the corresponding certainty equivalent \( CE_{ml}(\cdot) \) with the boundary condition \( CE_{ml}(\theta_{\min}) = \bar{w} \) is achieved by the wage schedule \( t: \mathcal{Y} \to \mathbb{R} \) defined by Eq. (50):

\[
t(y) = \min \left[ \int_{\hat{\theta}_{\min}}^{\hat{\theta}} \beta_{ml}(s) \gamma ds + C(y - \gamma \hat{\theta}) \mid \hat{\theta} \in [\theta_{\min}, \theta_{\max}] \right] + \bar{w} \tag{50}
\]

for each \( y \in \mathcal{Y} \).

Proof of Proposition 12. It will be shown that \( y_{ml}(\theta) \in \arg\max [t(y) - C(y - \gamma \theta) \mid y \geq \gamma \theta] \) for every \( \theta \in [\theta_{\min}, \theta_{\max}] \). Consider any \( \theta \in [\theta_{\min}, \theta_{\max}] \). It suffices to show that \( \max [t(y) - C(y - \gamma \theta) \mid y \geq \gamma \theta] = t(y_{ml}(\theta)) - C(y_{ml}(\theta) - \gamma \theta) \). Let \( h(y, \hat{\theta}) = \int_{\hat{\theta}_{\min}}^{\hat{\theta}} \beta_{ml}(s) \gamma ds + C(y - \gamma \hat{\theta}) \). Notice that \( t(y) = \min [h(y, \hat{\theta}) \mid \hat{\theta} \in [\theta_{\min}, \theta_{\max}]] + \bar{w} \). The first-and second-order conditions with respect to \( \hat{\theta} \) are the following:

\[
h_{\hat{\theta}}(y, \hat{\theta}) = \beta_{ml}(\hat{\theta}) \gamma + C'(y - \gamma \hat{\theta})(-\gamma) = \beta_{ml}(\hat{\theta}) \gamma + \frac{\gamma}{k} (y - \gamma \hat{\theta}) = \frac{\gamma}{k} (\beta_{ml}(\hat{\theta}) k - y + \gamma \hat{\theta})
\]

and

\[
h_{\hat{\theta}\hat{\theta}}(y, \hat{\theta}) = \frac{\gamma}{k} (\beta_{ml}(\hat{\theta}) k + \gamma) = \gamma \beta_{ml}(\hat{\theta}) + \frac{\gamma^2}{k}.
\]

By the weak monotonicity of \( \beta_{ml}(\cdot) \) and \( \gamma > 0 \), the second-order condition is always satisfied; that is, \( h_{\hat{\theta}\hat{\theta}}(y, \hat{\theta}) \geq 0 \) holds. Thus, the first-order condition is sufficient. In particular, the pair \( (y(\theta), \theta) \) satisfies the first-order condition \( h_{\hat{\theta}}(y, \hat{\theta}) = 0 \) because \( \beta_{ml}(\theta) k - y(\theta) + \gamma \theta = (\epsilon_{ml}(\theta) + \gamma \theta) - y(\theta) = y(\theta) - y(\theta) = 0 \).

It turns out that \( h(y(\theta), \hat{\theta}) \) is minimized at \( \hat{\theta} = \theta \). Therefore, I have

\footnote{A similar method was used in Watabe (2016) in a context of second-degree price discrimination.}
I conclude that \( y \) is in the envelope condition. I call such selection \( y \) as a type-assignment function.

This yields that

\[
\int_{\theta_{\text{min}}}^{\theta} \beta_{m\ell}(s) \gamma ds = t(y_{m\ell}(\theta)) - C(y_{m\ell}(\theta) - \gamma) - \bar{\omega}.
\]

Recall the envelope condition \( C\text{E}_{m\ell}(\theta) = \beta_{m\ell}(\theta) \gamma \) in terms of the certainty equivalent rather than the information rent as a part of the incentive compatibility constraint. The certainty equivalent under the menu of linear contracts is written as

\[
C\text{E}_{m\ell}(\theta) = \bar{\omega} + \int_{\theta_{\text{min}}}^{\theta} \beta_{m\ell}(s) \gamma ds.
\]

This, together with the expression for the information rent, yields that the indirect utility under the wage schedule \( t(\cdot) \) achieves the same certainty equivalent under the menu of linear contracts:

\[
C\text{E}_{m\ell}(\theta) = \bar{\omega} + \int_{\theta_{\text{min}}}^{\theta} \beta_{m\ell}(s) \gamma ds = t(y_{m\ell}(\theta)) - C(y_{m\ell}(\theta) - \gamma) - \bar{\omega}.
\]

Next, I will show that \( \max \{t(y) - C(y - \gamma) \mid y \geq \gamma\} = C\text{E}_{m\ell}(\theta) \). Consider any \( y \geq 0 \). By definition of \( t(\cdot) \), I see that \( t(y) \leq \int_{\theta_{\text{min}}}^{\theta} \beta_{m\ell}(s) \gamma ds + C(y - \gamma) + \bar{\omega} \), which implies that \( t(y) - C(y - \gamma) \leq \int_{\theta_{\text{min}}}^{\theta} \beta_{m\ell}(s) \gamma ds + \bar{\omega} = C\text{E}_{m\ell}(\theta) \). Since \( y \) was arbitrary chosen, it follows that \( \max \{t(y) - C(y - \gamma) \mid y \geq \gamma\} \leq C\text{E}_{m\ell}(\theta) \). On the other hand, I see that \( \max \{t(y) - C(y - \gamma) \mid y \geq \gamma\} - C\text{E}_{m\ell}(\theta) \geq t(y_{m\ell}(\theta)) - C(y_{m\ell}(\theta) - \gamma) - C\text{E}_{m\ell}(\theta) = 0 \), where the last equality follows from Eq.(51). Hence, \( \max \{t(y) - C(y - \gamma) \mid y \geq \gamma\} \geq C\text{E}_{m\ell}(\theta) \). These two inequalities imply the following equivalence:

\[
\max \{t(y) - C(y - \gamma) \mid y \geq \gamma\} = C\text{E}_{m\ell}(\theta). \tag{52}
\]

Equations (51) and (52) yield that

\[
\max \{t(y) - C(y - \gamma) \mid y \geq \gamma\} = t(y_{m\ell}(\theta)) - C(y_{m\ell}(\theta) - \gamma).
\]

I conclude that \( y(\theta) \) solves the worker’s optimization problem under the wage schedule \( t : \mathcal{Y} \rightarrow \mathbb{R} \). This establishes the proposition.

The advantage of constructing the particular wage schedule is that its second-derivative is tractable for the analysis of the shape of it. Firstly, the wage schedule \( t(\cdot) \) defined in Eq.(50) satisfies the following envelope condition. I call such selection \( \psi(y) \) as a type-assignment function.\(^{29}\)

\(^{29}\)Nödeke and Samuelson (2007) formulate adverse selection principal-agent problems by restricting attention to envelope condition of nonlinear tariffs in the context of the second-degree price discrimination where the agent’s utility function satisfies the usual single-crossing property. They do not attempt to analyze the shape of nonlinear tariffs.
\[ f'(y) = C'(y - \gamma \psi(y)) \]

where \( \psi(y) \in \arg\min \left[ \int_{\theta_{\min}}^{\hat{\theta}} \beta(s) \gamma ds + C(y - \gamma \hat{\theta}) \mid \hat{\theta} \in [\theta_{\min}, \theta_{\max}] \right]. \]

It will be shown that it is strictly convex beyond \( \max \{ y_{ml}(\theta_{\min}), y_{ml}(\theta_{\max}) \} \), but it is constant over the interval \([y_{ml}(\theta_{\min}), y_{ml}(\theta_{\max})]\) if there is a bunching in the menu of linear contracts. I need to identify the sign of the second derivative of the wage schedule. Differentiating the envelope condition in Eq.(53) with respect to \( y \) to get,

\[ f''(y) = C''(y - \gamma \psi(y)) \times (1 - \gamma \psi'(y)). \]  

If the wage schedule has a piece-wise linear part, then the second derivative of the optimal wage schedule must be zero. The sign of the second derivative of the wage schedule is determined by the sign of the second term on the right-hand side in Eq.(54) because the disutility function of effort is assumed to be strictly convex. Here, it must be emphasized that the type-assignment function \( \psi(y) \) has an easy economic interpretation. Actually, its derivative and the derivative of the production rule are related inversely.

Lemma 3. For every feasible menu of linear contracts \( \langle s_{ml}(\cdot), \beta_{ml}(\cdot) \rangle \) and the corresponding certainty equivalent \( CE_{ml}(\cdot) \) with the boundary condition \( CE_{ml}(\theta_{\min}) = \hat{w} \), the type-assignment correspondence \( \Gamma(y) = \arg\min \left[ \int_{\theta_{\min}}^{\hat{\theta}} \beta_{ml}(s) \gamma ds + C(y - \gamma \hat{\theta}) \mid \hat{\theta} \in [\theta_{\min}, \theta_{\max}] \right] \) has the following properties:

1. \( \Gamma(y) \) is a compact subset of \([\theta_{\min}, \theta_{\max}]\) for each \( y \in Y \).
2. The composite \( \Gamma \circ y : [\theta_{\min}, \theta_{\max}] \to [\theta_{\min}, \theta_{\max}] \) is self-belonging in the sense that \( \theta \in \Gamma(y_{ml}(\theta)) \) for every \( \theta \in [\theta_{\min}, \theta_{\max}] \).

Proof of Lemma 3. The first assertion is immediate from the Berge’s maximum theorem. Suppose, by way of contradiction, that \( \theta \notin \Gamma(y_{ml}(\theta)) \). By definition, \( \psi(y_{ml}(\theta)) \in \Gamma(y_{ml}(\theta)) \). Since \( \theta \in [\theta_{\min}, \theta_{\max}] \), it must be the case that

\[ \int_{\theta_{\min}}^{\psi(y_{ml}(\theta))} \beta_{ml}(s) \gamma ds + C(y_{ml}(\theta) - \gamma \psi(y_{ml}(\theta))) < \int_{\theta_{\min}}^{\theta} \beta_{ml}(s) \gamma ds + C(y_{ml}(\theta) - \gamma \theta). \]  

(55)

There are two possible cases to be considered. If \( \theta \geq \psi(y_{ml}(\theta)) \), then the inequality in Eq.(55) gives us

\[ 0 < \int_{\psi(y_{ml}(\theta))}^{\theta} \beta_{ml}(s) \gamma ds + C(y_{ml}(\theta) - \gamma \theta) - C(y_{ml}(\theta) - \gamma \psi(y_{ml}(\theta))) \]

\[ = \int_{\psi(y_{ml}(\theta))}^{\theta} CE_{ml}(s) ds + C(y_{ml}(\theta) - \gamma \theta) - C(y_{ml}(\theta) - \gamma \psi(y_{ml}(\theta))) \]

\[ = CE_{ml}(\theta) - CE_{ml}(\psi(y_{ml}(\theta))) + C(y_{ml}(\theta) - \gamma \theta) - C(y_{ml}(\theta) - \gamma \psi(y_{ml}(\theta))) \]

\[ = CE_{ml}(\theta) + C(y_{ml}(\theta) - \gamma \theta) - \{CE_{ml}(\psi(y_{ml}(\theta))) + C(y_{ml}(\theta) - \gamma \psi(y_{ml}(\theta)))\} \]

\[ = t(y_{ml}(\theta)) - t(y_{ml}(\theta)) = 0. \]
This is a contradiction. The proof of the remaining case is similar. This establishes the lemma.

The following lemma states that the procedure of constructing the single nonlinear compensation scheme involves the inverse of the production rule.\textsuperscript{30}

Lemma 4. For every feasible menu of linear contracts \((\alpha_{mt}(\cdot), \beta_{mt}(\cdot))\) and the corresponding certainty equivalent \(CE_{mt}(\cdot)\) with the boundary condition \(CE_{mt}(\theta_{min}) = \bar{a}\), the type-assignment function is written as the inverse function of the production rule: for each \(\bar{g} \in \mathcal{Y}\),

\[
\arg\min_{\theta_{\min}} \left[ \int_{\theta_{\min}}^{\theta_{\max}} \beta_{mt}(s) \gamma ds + C(\bar{g} - \gamma \hat{\theta}) \mid \hat{\theta} \in [\theta_{\min}, \theta_{\max}] \right] = \{ y_{mt}^{-1}(\bar{g}) \}.
\]

Proof of Lemma 4. Let \(\bar{g} \in \mathcal{Y}\). Firstly, I will show that for each type-assignment function \(\psi(\bar{g}) \in \Gamma(\bar{g})\), both composites \(1) y_{mt} \circ \psi : \mathcal{Y} \to \mathcal{Y}\) and \(2) \psi \circ y_{mt} : [\theta_{\min}, \theta_{\max}] \to [\theta_{\min}, \theta_{\max}]\) are identity functions. For the first identity, any interior optimum \(\psi(\bar{g})\) yields the following first-order condition:

\[
0 = \beta_{mt}(\psi(\bar{g})) + C(\bar{g} - \gamma \psi(\bar{g}))(\gamma) - \frac{\gamma}{k} (\beta_{mt}(\psi(\bar{g})) - \bar{g} + \gamma \psi(\bar{g}))
\]

\[
= \frac{\gamma}{k} (\epsilon_{mt}(\psi(\bar{g})) + \gamma \psi(\bar{g}) - \bar{g})
\]

and hence \(\bar{g} = y_{mt}(\psi(\bar{g})) = (y_{mt} \circ \psi)(\bar{g})\). This establishes that the composite \(y_{mt} \circ \psi : \mathcal{Y} \to \mathcal{Y}\) is the identity function.

It remains to show another identity. Without loss of generality, I may choose \(\psi(\bar{g}) = \min \Gamma(\bar{g})\) by the compactness of \(\Gamma(\bar{g})\) in Lemma 3. Since \(\theta \in \Gamma(\psi(\theta))\), it follows that \(\psi(y_{mt}(\theta)) \leq \theta\). Suppose, by way of contradiction, that \(\theta > \psi(y_{mt}(\theta))\). By the strict monotonicity of the production rule, \(y_{mt}(\theta) > y_{mt}(\psi(y_{mt}(\theta))) = (y_{mt} \circ \psi)(y_{mt}(\theta)) = y_{mt}(\theta)\), a contradiction.\textsuperscript{31} Therefore, it must be the case that \(\theta = \psi(y_{mt}(\theta)) = (\psi \circ y_{mt})(\theta)\), and hence the composite \(\psi \circ y_{mt} : [\theta_{\min}, \theta_{\max}] \to [\theta_{\min}, \theta_{\max}]\) is the identity function. This establishes the lemma.

The following proposition tells us that the menu of linear contracts itself is not a linear piece-rate compensation scheme.

\textsuperscript{30}The type-assignment approach examined by Goldman et al (1984) and subsequent analysis by Nöldeke and Samuelson (2007) solves screening problems in which direct revelation mechanism consists of a transfer rule and a decision rule by focusing on using the inverse of a decision rule referred to as the type assignment. They obtain the solution to the principal-agent problem from a point-wise maximization in terms of type function. Lemma 4 reveals that the method taken in the paper is closely related with the type-assignment approach.\textsuperscript{31}Since two parameters \(\gamma\) and \(k\) are strictly positive, and \(\beta_{mt}(\cdot)\) is non-decreasing, it follows that the production rule \(y_{mt}(\theta) = \gamma \theta + \epsilon_{mt}(\theta) = \gamma \theta + \beta_{mt}(\theta)\) is strictly increasing in \(\theta\).
Proposition 13. For every feasible menu of linear contracts \( \langle a_m t(\cdot), b_m t(\cdot) \rangle \) and the corresponding certainty equivalent \( CE_m t(\cdot) \) with the boundary condition \( CE(\theta_{\text{min}}) = \bar{w} \), the wage schedule \( t : \mathcal{Y} \rightarrow \mathbb{R} \) defined in Eq.(50) is a quota-based curvilinear compensation schedule, and its second derivative is given by

\[
t''(\bar{y}) = \frac{\dot{e}_{mt}(\Psi(\bar{y}))}{k[\gamma + \dot{e}_{mt}(\Psi(\bar{y}))]} \geq 0 \quad \text{with equality only if } \dot{e}_{mt}(\Psi(\bar{y})) = 0,
\]

where \( \Psi(\bar{y}) = y_{mt}^{-1}(\bar{y}) \) for each \( \bar{y} \in \mathcal{Y} \).

Proof of Proposition 13. By the previous lemmas, the type-assignment function is the inverse of the production rule. This implies that

\[
y'(\bar{y}) = \frac{1}{y_{mt}(\theta)} \quad \text{for } y_{mt}(\theta) = \bar{y}.
\]

Since the sales response function is of the form \( y_{mt}(\theta) = \gamma \theta + e_{mt}(\theta) \), it follows that

\[
\Psi(\bar{y}) = \frac{1}{\gamma + \dot{e}_{mt}(\theta)},
\]

and then

\[
1 - \gamma \Psi'(\bar{y}) = 1 - \frac{\gamma}{\gamma + \dot{e}_{mt}(\theta)} = \frac{\dot{e}_{mt}(\theta)}{\gamma + \dot{e}_{mt}(\theta)}.
\]

Since \( C''(e) = \frac{1}{k} \), the second derivative of the optimal wage schedule in Eq.(54) will be the following.

\[
t''(y) = C''(y - \gamma \Psi(y)) \times (1 - \gamma \Psi'(y))
\]

\[
= C''(y - \gamma \Psi(y)) \times \frac{\dot{e}_{mt}(\Psi(y))}{\gamma + \dot{e}_{mt}(\Psi(y))}
\]

\[
= \frac{\dot{e}_{mt}(\Psi(y))}{k[\gamma + \dot{e}_{mt}(\Psi(y))]} \geq 0 \quad \text{with equality only if } \dot{e}_{mt}(\Psi(y)) = 0.
\]

(56)

This establishes the proposition.

Recall now that the effort function is constant only over \([\theta_{\text{min}}, \bar{\theta}]\) if there is a bunching in the menu of linear contracts by Proposition 1. Therefore, the optimal wage schedule is constant over that interval. The main finding here is that if the reservation wage is set exogenously, then the optimal menu of linear contracts can be regarded as a single quota-based contract.

I shall use Propositions 12 and 13 to derive the single quota-based contract explicitly that implements the menu of linear contracts in Proposition 1. Recall that the personalized pay-performance sensitivity is given by \( \beta_{mt}(\theta) = 0 \) for all type \( \theta \in [\theta_{\text{min}}, \theta_{\text{min}}] \). Then, the wage schedule can be written as

\[
t(y) = \int_{\theta_{\text{min}}}^{\Psi(y)} \beta_{mt}(s) \gamma ds + C(y - \gamma \Psi(y)) + \bar{w}
\]

\[
= \int_{\theta_{\text{min}}}^{\Psi(y)} \frac{\gamma}{k + \rho c^2} (k - \gamma \eta(\theta_{\text{max}} - s)) ds + \frac{(y - \gamma \Psi(y))^2}{2k} + \bar{w}.
\]
I can interpret this wage schedule as a combination plan with a fixed salary $\tilde{w}$ and the remaining part represents commissions depending on output. It is worth to know whether the part of commissions is linearly, progressively, or degressively.\footnote{Regarding the piece-wise linearity of the optimal contract, Gibbons (1987) has already argued that the worker’s private information makes a linear contract suboptimal for risk-neutral and heterogeneous workers.} Recall that the type-assignment function $\psi(\cdot)$ is the inverse of the sales response function $y_{ml}(\cdot)$. Since the sales response function is written as

$$y_{ml}(\theta) = \gamma \theta + e_{ml}(\theta) = \begin{cases} 
\gamma \theta & \text{for } \theta \leq \theta_{\text{min}}^{\text{qb}}, \\
\gamma \theta + \frac{k^2}{k + \rho \sigma^2} \left(1 - \frac{\gamma - 1 - F(\theta)}{F(\theta)}\right) & \text{for } \theta > \theta_{\text{min}}^{\text{qb}}.
\end{cases}$$

Here, for each $\theta > \theta_{\text{min}}^{\text{qb}}$, substituting $(1 - F(\theta))/f(\theta) = \eta(\theta_{\text{max}} - \theta)$ yields that

$$y_{ml}(\theta) = \gamma \theta + \frac{k^2}{k + \rho \sigma^2} \frac{k - \gamma \eta (\theta_{\text{max}} - \theta)}{\eta(\theta_{\text{max}} - \theta)} = \gamma \theta \left(1 + \frac{k}{k + \rho \sigma^2}\right) + \frac{k(k - \gamma \eta \theta_{\text{max}})}{k + \rho \sigma^2},$$

and then I obtain the following type-assignment function:

$$\psi(y) = \begin{cases} 
\frac{y}{\gamma} & \text{for } y \leq \gamma \theta_{\text{min}}^{\text{qb}}, \\
\frac{(k + \rho \sigma^2)y - k(k - \gamma \eta \theta_{\text{max}})}{\gamma(\eta + 1)(k + \rho \sigma^2)} & \text{for } y > \gamma \theta_{\text{min}}^{\text{qb}}.
\end{cases}$$

Therefore, the optimal curvilinear contract with a quota $y_{ml}(\theta_{\text{min}}^{\text{qb}}) = \gamma \theta_{\text{min}}^{\text{qb}} = \gamma \left(\theta_{\text{max}} - \frac{k}{\eta}\right)$ is as follows:

$$t(y) = \begin{cases} 
\tilde{w} & \text{for } y \in [y_{ml}(\theta_{\text{min}}), y_{ml}(\theta_{\text{max}})], \\
\frac{(k + \rho \sigma^2)y - k(k - \gamma \eta \theta_{\text{max}})}{\gamma(\eta + 1)(k + \rho \sigma^2)} + \tilde{w} & \text{for } y \in (y_{ml}(\theta_{\text{min}}), y_{ml}(\theta_{\text{max}})].
\end{cases}$$

Moreover, the first- and the second-derivatives can be explicitly described as follows:\footnote{Notice that, in the expression of the first-derivative of the wage schedule, $k + \eta(y - \gamma \theta_{\text{max}}) > k + \eta(y_{ml}(\theta_{\text{min}}^{\text{ml}}) - \gamma \theta_{\text{max}}) = k + \eta\gamma(\theta_{\text{min}}^{\text{ml}} - \gamma \theta_{\text{max}}) = k + \gamma \eta(\theta_{\text{max}} - \theta_{\text{min}}^{\text{ml}}) = k - \gamma \eta \times \frac{k}{\eta} - k = 0$ for each $y > y_{ml}(\theta_{\text{min}}^{\text{ml}})$. Therefore, the wage schedule is strictly increasing beyond the quota $y_{ml}(\theta_{\text{min}}^{\text{ml}})$.}

$$t'(y) = \begin{cases} 
0 & \text{for } y \in [y_{ml}(\theta_{\text{min}}), y_{ml}(\theta_{\text{max}})], \\
\frac{k + \eta(y - \gamma \theta_{\text{max}})}{(\eta + 1)(k + \rho \sigma^2)} > 0 & \text{for } y \in (y_{ml}(\theta_{\text{min}}), y_{ml}(\theta_{\text{max}})],
\end{cases}$$

$$t''(y) = \begin{cases} 
0 & \text{for } y \in [y_{ml}(\theta_{\text{min}}), y_{ml}(\theta_{\text{max}})], \\
\frac{\eta}{(\eta + 1)(k + \rho \sigma^2)} > 0 & \text{for } y \in (y_{ml}(\theta_{\text{min}}), y_{ml}(\theta_{\text{max}})]
\end{cases}$$
Eq. (57)-Eq. (59) illustrate how the quota-based contract which is equivalent to the menu of linear contracts is affected by the information structure, the characteristics of the response function, etc. In the following examples, I shall illustrate how a properly designed quota-based contract looks like. The existence of a quota $y_{mt}(q^{mt}_{\min})$ depends on whether there is a bunching in the menu of linear contracts. The wage schedule is strictly increasing and strictly convex beyond the quota, and it is constant at the reservation wage up to the quota. In the following numerical experiments, I set the values of the parameters such as $k = 1, r = 2, \sigma^2 = 1, \theta_{\text{max}} = 5$ and $\theta_{\min} = 1$.

Example 1 (no bunching case). Consider the case that there is no bunching in the menu of linear contracts ($\lambda > 1$). The three alternative forms of contracts are well-defined only if $1 < \lambda < \frac{2\eta + 1}{\eta + 1}$ as shown in Figure 10. Since $\lambda = \frac{k}{\eta + 1} = \frac{1}{\eta + 1}$, it follows that $1 < \frac{1}{\eta + 1} < \frac{2\eta + 1}{\eta + 1}$. For each $\eta = 0.5, 1$ and 2, the value of $\gamma$ will be restricted. If $\eta = 0.5$ then $1 < \frac{1}{2} < \frac{3}{2}$ yields that $\frac{3}{8} < \gamma < \frac{1}{2}$. If $\eta = 1$ then $1 < \frac{1}{2} < \frac{3}{2}$ yields that $\frac{1}{8} < \gamma < \frac{1}{2}$. Finally, if $\eta = 2$ then $1 < \frac{1}{4} < \frac{3}{2}$ yields that $\frac{3}{16} < \gamma < \frac{1}{8}$. Therefore, I may consider the following pairs: $(\gamma, \eta) = (0.4, 0.5), (0.2, 1), (0.1, 2)$. For these combinations of $(\gamma, \eta)$, the value of $\lambda$ is equal to 1.25 (see Figure 10).

![Figure 10: $(\gamma, \eta) = (0.4, 0.5), (0.2, 1), (0.1, 2)$](image)

The optimal compensation scheme $t(\cdot)$ are depicted in Figures 11-13. There is no kink in the compensation schedule.

Example 2 (bunching case). Consider the case that there is a bunching in the menu of linear contracts ($\lambda \leq 1$). I need to identify the relevant pairs of $(\gamma, \eta)$ again. The three alternative forms of contracts are well-defined only if $1 < \lambda \leq 1$ as shown in Figure 14. Since $\lambda = \frac{k}{\eta + 1} = \frac{1}{\eta + 1}$, it follows that $\frac{1}{\eta + 1} < \frac{1}{\eta + 1} \leq 1$. For each $\eta = 0.6, 1$ and 2, this condition will restrict the range of the value of $\gamma$. If $\eta = 0.6$ then $\frac{1}{1.6} < \frac{1}{2.6} \leq 1$ yields that $\frac{5}{12} \leq \gamma < \frac{2}{3}$. If $\eta = 1$ then $\frac{1}{2} < \frac{1}{4} < \frac{1}{2}$ yields that $0.25 \leq \gamma < 0.5$. Finally, if $\eta = 2$ then $\frac{1}{5} < \frac{1}{1.5} \leq 1$ yields that $0.125 \leq \gamma < \frac{3}{16} = 0.1875$. For instance, I may attention to the following pairs: $(\gamma, \eta) = (0.5, 0.6), (0.3, 1), (0.15, 2)$. For these combinations of $(\gamma, \eta)$, the value of $\lambda$ is $\frac{5}{6}$ (see Figure 14).
The optimal compensation scheme \( t(\cdot) \) are depicted in Figures 15-17. These figures suggest that the menu of linear contracts can be substantially replicated by a single piece-wise linear contract.

Eventually, the compensation plan consists of a salary plus continuously sliding commission rates. As mentioned in Raju and Srinivasan (1996), intuitively, any linear contract (or piece-rate contract) is not able to incorporate the convex shape of the single quota-based plan as a representation of the menu of linear contracts, and thus it produces a significant amount of non-optimality. In contrast, a single piece-wise quota-based contract is able to capture the convexity of the curvilinear quota-based contract, and this results in the high performance of it.
7 Conclusion

In this paper, I have examined the performance of simple piece-wise linear contracts relative to the menu of linear contracts in a principal-agent problem under moral hazard and adverse selection. I have provided analytical upper and lower bounds on the performance measures. It turns out that non-uniform salesforce heterogeneity plays a crucial role in determining the performance of the piece-rate incentive scheme, while the quota-based contract is robust to the shifts in the distribution of private information. A primary finding is that a properly designed simple piece-wise linear contract always captures at least 73 percent of the incremental gain secured under the optimal menu of linear contracts on the entire region of the relevant parameter values. The paper have provided an explanation for the use of piece-wise-linear-threshold contracts in practice.

References


Appendix

Proof of Proposition 4

I need to calculate the expectation of $(k - \gamma \eta (\theta_{\text{max}} - \theta))^2$ over the interval $[\theta_{\text{min}}^m, \theta_{\text{max}}]$ in Eq.(13). Firstly, notice that this term will not vanish except for $\theta = \theta_{\text{min}}^m$.

\[
k - \gamma \eta (\theta_{\text{max}} - \theta) = \begin{cases} 
  k - \gamma \eta (\theta_{\text{max}} - \theta_{\text{min}}^m) = k - \gamma \eta \times \frac{k}{\gamma \eta} = k - k = 0 & \text{for } \theta = \theta_{\text{min}}^m, \\
  < k - \gamma \eta (\theta_{\text{max}} - \theta_{\text{min}}^m) = 0 & \text{for } \theta < \theta_{\text{min}}^m.
\end{cases}
\]

Then, I have

\[
\int_{\theta_{\text{min}}^m}^{\theta_{\text{max}}^m} (k - \gamma \eta (\theta_{\text{max}} - \theta))^2 f(\theta) d\theta = \int_{\theta_{\text{min}}^m}^{\theta_{\text{max}}^m} [(k - \gamma \eta \theta_{\text{max}})^2 + 2\gamma \eta (k - \gamma \eta \theta_{\text{max}}) \theta + \gamma^2 \eta^2 \theta^2] f(\theta) d\theta. \tag{60}
\]

The first term $(k - \gamma \eta \theta_{\text{max}})^2$ in Eq.(60) is independent of the private information:

\[
\int_{\theta_{\text{min}}^m}^{\theta_{\text{max}}^m} (k - \gamma \eta \theta_{\text{max}})^2 f(\theta) d\theta = (k - \gamma \eta \theta_{\text{max}})^2 F(\theta_{\text{min}}^m) = (k - \gamma \eta \theta_{\text{max}})^2 (1 - \text{Prob}(\theta > \theta_{\text{min}}^m)). \tag{61}
\]

Regarding the expectation of $2\gamma \eta (k - \gamma \eta \theta_{\text{max}})$ over the relevant interval will be

\[
\int_{\theta_{\text{min}}^m}^{\theta_{\text{max}}^m} 2\gamma \eta (k - \gamma \eta \theta_{\text{max}}) \theta f(\theta) d\theta = 2\gamma \eta (k - \gamma \eta \theta_{\text{max}}) \int_{\theta_{\text{min}}^m}^{\theta_{\text{max}}^m} \theta f(\theta) d\theta
= 2\gamma \eta (k - \gamma \eta \theta_{\text{max}}) \left( \mathbb{E}(\theta) - \text{Prob}(\theta > \theta_{\text{min}}^m) \left( \mathbb{E}(\theta) + \frac{\Delta - \delta}{\eta + 1} \right) \right). \tag{62}
\]

Finally, the expectation of $\gamma^2 \eta^2 \theta^2$ over that interval involves the expressions for the variance and the expectation of $\theta$.

\[
\int_{\theta_{\text{min}}^m}^{\theta_{\text{max}}^m} \gamma^2 \eta^2 \theta^2 f(\theta) d\theta = \gamma^2 \eta^2 \left( \text{Var}(\theta) + \mathbb{E}(\theta)^2 - \text{Prob}(\theta > \theta_{\text{min}}^m) \left( \theta_{\text{max}}^2 - \frac{2\theta_{\text{max}} \delta}{\eta + 1} + \frac{\delta^2}{2(\eta + 1)} \right) \right). \tag{63}
\]

Combining Eq.(61)-(63) to get
\[
\int_{q_{\text{min}}}^{q_{\text{max}}} (k - \gamma \eta (\theta_{\text{max}} - \theta))^2 f(\theta) d\theta
\]

\[
= (k - \gamma \eta \theta_{\text{max}})^2 \left(1 - \text{Prob}(\theta > \theta_{\text{min}}^{\text{mt}})\right) \\
+ 2\gamma \eta (k - \gamma \eta \theta_{\text{max}}) \left(\mathbb{E}(\theta) - \text{Prob}(\theta > \theta_{\text{min}}^{\text{mt}}) \left(\mathbb{E}(\theta) + \frac{\Delta - \delta}{\eta + 1}\right)\right) \\
+ \gamma^2 \eta^2 \left(\text{Var}(\theta) + \mathbb{E}(\theta)^2 - \text{Prob}(\theta > \theta_{\text{min}}^{\text{mt}}) \left(\theta_{\text{max}}^2 - \frac{2\theta_{\text{max}} \delta}{\eta + 1} + \frac{\delta^2}{2\eta + 1}\right)\right) \\
= (k - \gamma \eta \theta_{\text{max}})^2 - (k - \gamma \eta \theta_{\text{max}})^2 \text{Prob}(\theta > \theta_{\text{min}}^{\text{mt}}) \\
+ 2\gamma \eta (k - \gamma \eta \theta_{\text{max}}) \mathbb{E}(\theta) - 2\gamma \eta (k - \gamma \eta \theta_{\text{max}}) \text{Prob}(\theta > \theta_{\text{min}}^{\text{mt}}) \left(\mathbb{E}(\theta) + \frac{\Delta - \delta}{\eta + 1}\right) \\
+ \gamma^2 \eta^2 \text{Var}(\theta) + \gamma^2 \eta^2 \mathbb{E}(\theta)^2 \\
- \gamma^2 \eta^2 \text{Prob}(\theta > \theta_{\text{min}}^{\text{mt}}) \left(\theta_{\text{max}}^2 - \frac{2\theta_{\text{max}} \delta}{\eta + 1} + \frac{\delta^2}{2\eta + 1}\right). 
\]  

(64)

Notice that

\[
(k - \gamma \eta \theta_{\text{max}})^2 + 2\gamma \eta (k - \gamma \eta \theta_{\text{max}}) \mathbb{E}(\theta) + \gamma^2 \eta^2 \mathbb{E}(\theta)^2 = (k - \gamma \eta \theta_{\text{max}} + \gamma \eta \mathbb{E}(\theta))^2 \\
= (k - \gamma \eta (\theta_{\text{max}} - \mathbb{E}(\theta)))^2 \\
= (k - \gamma (\mathbb{E}(\theta) - \theta_{\text{min}}))^2 
\]

and

\[
\mathbb{E}(\theta) + \frac{\Delta - \delta}{\eta + 1} = \frac{\eta \theta_{\text{max}} + \theta_{\text{min}} + \theta_{\text{max}} - \theta_{\text{min}} - \delta}{\eta + 1} = \frac{(\eta + 1)\theta_{\text{max}} - \delta}{\eta + 1} = \theta_{\text{max}} - \frac{\delta}{\eta + 1}.
\]

The coefficients of Prob(\theta > \theta_{\text{min}}^{\text{mt}}) in Eq.(64) can be simplified as follows:
\[(k - \gamma \eta \theta_{\text{max}})^2 + 2 \gamma \eta (k - \gamma \eta \theta_{\text{max}}) \left( \mathbb{E}(\theta) + \frac{\Delta - \delta}{\eta + 1} \right) + \gamma^2 \eta^2 \left( \theta_{\text{max}}^2 - \frac{2 \theta_{\text{max}} \delta}{\eta + 1} + \frac{\delta^2}{2\eta + 1} \right) \]

\[= (k - \gamma \eta \theta_{\text{max}})^2 + 2 \gamma \eta (k - \gamma \eta \theta_{\text{max}}) \left( \theta_{\text{max}} - \frac{\delta}{\eta + 1} \right) + \gamma^2 \eta^2 \left( \theta_{\text{max}}^2 - \frac{2 \theta_{\text{max}} \delta}{\eta + 1} + \frac{\delta^2}{2\eta + 1} \right) \]

\[= \left( (k - \gamma \eta \theta_{\text{max}}) + \gamma \eta \left( \theta_{\text{max}} - \frac{\delta}{\eta + 1} \right) \right)^2 - \gamma^2 \eta^2 \delta^2 \left( \frac{1}{\eta + 1} - \frac{1}{\eta + 1} \right) \]

\[= \left( \frac{(\eta + 1)k - \gamma \eta \delta}{\eta + 1} \right)^2 + \gamma^2 \eta^2 \delta^2 \left( \frac{(\eta + 1)^2 - (2\eta + 1)}{(\eta + 1)^2(2\eta + 1)} \right) \]

\[= \left( \frac{(\eta + 1)k - \gamma \eta \delta}{1 + \eta} \right)^2 + \frac{\gamma^2 \eta^2 \delta^2}{(\eta + 1)^2(2\eta + 1)} \]

\[= \left( \frac{(\eta + 1)k - \gamma \eta \delta}{\eta + 1} \right)^2 + \frac{\gamma^2 \eta^2 \delta^2 \Delta^2}{(\eta + 1)^2(2\eta + 1) \Delta^2} \]

\[= \left( \frac{(\eta + 1)k - \gamma \eta \delta}{\eta + 1} \right)^2 + \gamma^2 \eta^2 \left( \frac{\delta}{\Delta} \right)^2 \text{Var}(\theta). \]

Finally, Eq.(60) can be written as

\[\int_{\theta_{\text{min}}}^{\theta_{\text{max}}} (k - \gamma \eta (\theta_{\text{max}} - \theta))^2 f(\theta) d\theta \]

\[= (k - \gamma (\mathbb{E}(\theta) - \theta_{\text{min}}))^2 + \gamma^2 \eta^2 \text{Var}(\theta) - \text{Prob}(\theta > \theta_{\text{min}}) \left[ \left( \frac{(1 + \eta)k - \gamma \eta \delta}{1 + \eta} \right)^2 + \gamma^2 \eta^2 \left( \frac{\delta}{\Delta} \right)^2 \text{Var}(\theta) \right]. \]

Therefore, the maximized expected profit under the menu of linear contracts will be
\[ \mathbb{E} \pi_{m\ell}(\gamma, \eta, \lambda) = \mathbb{E} \pi_{m\ell}(\gamma, \eta, \lambda) + \frac{\gamma^2 \theta^2 \text{Var}(\theta)}{2(k + \rho \sigma^2)} \]

\[ - \frac{1}{2(k + \rho \sigma^2)} \left[ (k - \gamma \eta (\theta_{\max} - \mathbb{E}(\theta)))^2 + \gamma^2 \theta^2 \text{Var}(\theta) \right] \]

\[ - \text{Prob}(\theta > \theta_{\text{min}}^{m\ell}) \left[ \left( \frac{(1 + \eta)k - \gamma \eta \delta}{1 + \eta} \right)^2 + \gamma^2 \theta^2 \left( \frac{\delta}{\Delta} \right)^2 \text{Var}(\theta) \right] \]

\[ = \mathbb{E} \pi_{m\ell}(\gamma, \eta, \lambda) - \frac{(k - \gamma \eta (\theta_{\max} - \mathbb{E}(\theta)))^2}{2(k + \rho \sigma^2)} \]

\[ + \frac{\text{Prob}(\theta > \theta_{\text{min}}^{m\ell})}{2(k + \rho \sigma^2)} \left[ \left( \frac{(1 + \eta)k - \gamma \eta \delta}{1 + \eta} \right)^2 + \gamma^2 \theta^2 \left( \frac{\delta}{\Delta} \right)^2 \text{Var}(\theta) \right] \]

Reorganizing the expression for \( \mathbb{E} \pi_{m\ell}(\gamma, \eta, \lambda) \) yields that

\[ \mathbb{E} \pi_{m\ell}(\gamma, \eta, \lambda) = \mathbb{E} \pi_{f\omega}(\gamma, \eta) + \frac{\text{Prob}(\theta > \theta_{\text{min}}^{m\ell})}{2(k + \rho \sigma^2)} \left[ \left( \frac{(1 + \eta)k - \gamma \eta \delta}{1 + \eta} \right)^2 + \gamma^2 \theta^2 \left( \frac{\delta}{\Delta} \right)^2 \text{Var}(\theta) \right]. \quad (65) \]

The second term on the right-hand side in Eq.(65) is written as

\[ \frac{\text{Prob}(\theta > \theta_{\text{min}}^{m\ell})}{2(k + \rho \sigma^2)} \left[ \left( \frac{(1 + \eta)k - \gamma \eta \delta}{1 + \eta} \right)^2 + \gamma^2 \theta^2 \left( \frac{\delta}{\Delta} \right)^2 \text{Var}(\theta) \right] \]

\[ = \frac{1}{2(k + \rho \sigma^2)} \left( \frac{\delta}{\Delta} \right)^{\frac{3}{2}} \left[ (k - \gamma \eta \delta)^2 + \gamma^2 \theta^2 \left( \frac{\delta}{\Delta} \right)^2 \text{Var}(\theta) \right]. \]

Notice that

\[ k - \frac{\gamma \eta \delta}{1 + \eta} = \gamma \eta \left( \frac{k}{\gamma \eta} - \frac{\delta}{1 + \eta} \right) = \gamma \eta \left( \delta - \frac{\delta}{1 + \eta} \right) = \gamma \eta \delta \left( \frac{1 + \eta - 1}{1 + \eta} \right) = \frac{\gamma \eta^2 \delta}{1 + \eta}. \]

Factoring the inside the square brackets to get
\[
\left(k - \frac{\gamma \eta \delta}{1 + \eta}\right)^2 + \gamma^2 \eta^2 \left(\frac{\delta}{\Delta}\right)^2 \text{Var}(\theta) = \left(\frac{\gamma \eta \delta}{1 + \eta}\right)^2 + \gamma^2 \eta^2 \left(\frac{\delta}{\Delta}\right)^2 \frac{\Delta^2 \eta^2}{(1 + \eta)^2 (1 + 2\eta)}
\]

\[
= \frac{\gamma^2 \eta^4 \delta^2}{(1 + \eta)^2} \left(1 + \frac{1}{1 + 2\eta}\right)
\]

\[
= \frac{2\gamma^2 \eta^4 \delta^2}{(1 + \eta)(1 + 2\eta)}.
\]

Therefore, Eq.(65) becomes:

\[
\mathbb{E}\pi_{\text{ml}}(\gamma, \eta, \lambda, \lambda)_{\lambda \leq 1} = \mathbb{E}\pi_{\text{fw}}(\gamma, \eta) + \frac{1}{2(k + \rho \sigma^2)} \left(\frac{\delta}{\Delta}\right)^{\frac{1}{4}} \frac{2\gamma^2 \eta^4 \delta^2}{(1 + \eta)(1 + 2\eta)}
\]

\[
= \mathbb{E}\pi_{\text{fw}}(\gamma, \eta) + \frac{1}{k + \rho \sigma^2} \left(\frac{\delta}{\Delta}\right)^{\frac{1}{4}} \left(\frac{(1 + \eta) \gamma^2 \eta^2 \delta^2}{\Delta^2}\right) \left(\frac{\Delta^2 \eta^2}{(1 + \eta)^2 (1 + 2\eta)}\right)
\]

\[
= \mathbb{E}\pi_{\text{fw}}(\gamma, \eta) + \frac{(1 + \eta) \gamma^2 \eta^2}{k + \rho \sigma^2} \left(\frac{\delta}{\Delta}\right)^{\frac{1 + 2\eta}{\frac{1}{4}}} \text{Var}(\theta).
\]

This establishes the proof of Proposition 4. \(\blacksquare\)