Journal of Economic Theory and Econometrics, Vol. 32, No. 2, Jun. 2021, 74–91

Compensation Disparity between Risk Averse Agents under Adverse Selection and Moral Hazard*

Sung Hyun Kim†

Abstract We extend the LEN moral hazard model to allow for adverse selection and derive the optimal incentive contract menu. We show that both moral hazard and adverse selection separately cause compensation disparity between agents with different degrees of risk aversion. We also show that adverse selection aggravates the compensation disparity when more risk averse agents form a minority of the agent population.

Keywords Incentive Contract, Pay Gap, LEN Model

JEL Classification D82, J31, M52

*I thank two anonymous referees for comments and criticisms. I also thank Meg Meyer at Nuffield College, Oxford for stimulating lectures on the LEN model and June Kim for listening to my ideas in this paper.

†Department of Economics, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Republic of Korea, E-mail: sungkim@ewha.ac.kr

Received April 26, 2021, Revised May 31, 2021, Accepted June 2, 2021

1. INTRODUCTION

This paper examines how the total compensation may be different across different types of agents, in a contracting model where both adverse selection and moral hazard are present. A typical analysis of a contracting model focuses on the optimal contract, the terms of which include a level of base salary and a performance incentive schedule. We explore a less discussed question of total compensation and its disparity among agents.

As a first step towards answering the question, we place our model in a standard and familiar setting and seek to develop intuitive insights from the simple model. In particular, we investigate how adverse selection and moral hazard separately may affect the extent of compensation disparity.

In a pure moral hazard model, the optimal contract gives a more risk averse agent a lower-powered incentive contract, so he exerts lower effort. We show that he receives lower total compensation as a result. But the incentive contract takes the degree of risk aversion as a known parameter. If the degree of risk aversion is private information, the principal needs to offer an incentive-compatible menu of contracts to prevent adverse selection.

What does the optimal contract menu look like? What will happen to compensation disparity? The answers are not obvious and are not readily available in the literature. Our finding is that compensation disparity is larger under adverse selection and is worse if the proportion of more risk averse agents is lower (a "minority") in the agent population. As a potential application of the result, we offer a sketchy discussion of how the "pay gap" between social groups with different observable traits and slightly different composition of unobservable types might arise from information asymmetry.

Our finding may have social relevance in the following sense. If a certain group of agents, say those from low income background or with less education or even of female gender (Sarin and Wieland, 2016), are perceived (correctly or incorrectly) by the employer as more risk averse on average, then this group may be *guided* in the hiring and contracting process to choose those jobs that offer lower compensation. Moreover, if the share of the group in the workforce is small in the beginning and gradually increases, the "pay gap" would decrease over time but still remain sizable when the group is no longer a minority.

Someone looking at the compensation data (especially on an aggregated level) might interpret this as prevalent discrimination against the group. Another observer looking at the same data might point out that the group in question shows lower performance (lower effort, lower output) and the pay gap is the result of voluntary choices by agents. Our analysis suggests that both positions are problematic. In our model, the employer does not engage in explicitly discriminatory practices but the pay gap is real, arising from information asymmetry.

The modeling setup we employ is the LEN (linear, exponential, normal) model[1] of the moral hazard literature, combined with the binary type contracting model of the adverse selection literature. Both models are standard in each literature and combination of the two yields a familiar-looking but interesting form of the optimal contract menu. Our analysis focuses on the agent's degree of risk aversion as private type, but we believe that the general insight is not limited to risk aversion and can be extended to other dimensions of agent characteristics, such as the marginal cost of effort or the reservation wage.

2. THE MODEL

Most ingredients of the model are standard, so will be described succinctly. We have a two-stage game between a principal ("the firm") and an agent ("the worker"). The second stage is a standard moral hazard model, which has a well-known form of optimal incentive contract if it were a stand-alone model. We extend the model by introducing adverse selection in the first stage, which affects the second stage incentive contract.

Before the game begins, the worker's type is determined and observed by the worker only. The description of a type could potentially include the risk attitude, marginal cost of effort, reservation wage level, productivity, etc but we focus on the Arrow-Pratt coefficient of absolute risk aversion, denoted by θ hereafter.

In stage 1, the firm offers a set of linear contracts $\{(\alpha_i, \beta_i)\}$, which the worker accepts or rejects. The worker who rejects all offers gets the reservation utility or equivalently the reservation wage $\overline{w}$. In stage 2, the worker who accepted an offer chooses the effort level e. The outcome is realized as $y = e + \varepsilon$, where $\varepsilon \sim N(0, v)$ is a normally distributed random noise with variance v. The firm observes y, but cannot separate e from ε. The firm pays $w = \alpha_i + \beta_i y$ to the worker who accepted the contract (α_i, β_i). The agent's payoff function takes the exponential or CARA form $u(x;\theta) = -\exp(-\theta x)$, where $x = w - \frac{1}{2}ce^2$ is the net compensation consisting of random pay $w = \alpha_i + \beta_i y$ from the contract and disutility of effort $-\frac{1}{2}ce^2$. This constitutes the LEN model, with linear compensation, exponential utility and normally distributed noise (see Bolton and Dewatripont, 2005, Section 4.2; Kim, 2016; Kirkegaard, 2017). The firm is risk neutral and maximizes the expected value of outcome, net of compensation to the worker.

[1]This model assumes linear compensation scheme, exponential CARA utility for the agent and normally distributed noise in observed performance. See Section 2 for a fuller description.

It is well known for moral hazard models that (i) the first-best (without moral hazard) contract involves the principal bearing all risk and the risk averse agent receiving a fixed pay and (ii) the second-best (recognizing moral hazard) contract puts some risk on the agent in the form of an incentive scheme based on observable performance. Since the agent is risk averse, the risky incentive scheme is less desirable, particularly more so to a more risk averse agent. In fact, it can be shown for the LEN model that (iii) the optimal intensity of incentive compensation is inversely related to the degree of risk aversion of the agent ("the incentive intensity principle", Milgrom and Roberts, 1992, p.221), as well as the marginal disutility of effort and the variance of performance noise (see Section 3 for the explicit formula).

In our model, the firm is faced with the pool of potential workers whose degree of risk aversion is not observable. The firm has to take into account the fact that different types of workers can choose different contracts. Does the incentive intensity principle still apply? Is the "gap" between different intensities for different types larger or smaller? Do more risk averse agents in fact receive lower compensation? We will examine these questions next.

Since there are two layers of information asymmetry, we will ignore the first-best case (where there is no hidden information) and treat the pure moral hazard case as the benchmark. To use the terminology in Myerson (1982), our benchmark scenario requires *obedience* (in choosing the action intended by the principal) only, while the actual model requires both *obedience* and *honesty* (revealing private type by choice of contracts).

3. THE OPTIMAL CONTRACT

3.1. THE BENCHMARK: OBSERVABLE TYPE

Suppose, for the moment, that θ is observable. Then for each realized type θ, all the relevant parameters are given and the stage 2 is the standard LEN model, with the optimal incentive intensity (the incentive intensity principle formula):

$$\beta_i^o = \frac{1}{1+\theta_i cv}$$

and the worker's effort choice is

$$e_i^o = \frac{\beta^o}{c} = \frac{1}{c(1+\theta_i cv)}$$

(see Bolton and Dewatripont, 2005, pp.137–139 for derivation of the formulas).

In stage 1, the firm offers the contract (α_i^o, β_i^o) for each type i where the worker receives the reservation wage (binding participation constraint): then we have the base salary level

$$\alpha_i^o = \overline{w} + \frac{\theta_i cv - 1}{2c(1+\theta_i cv)^2}$$

so that the certainty equivalent of the compensation scheme (α_i^o, β_i^o) for type θ_i equals the reservation wage $\overline{w}$.

Notice that if $\theta_i cv > 1$ then $\alpha_i^o > \overline{w}$. In other words,

$$\widehat{\theta} \equiv \frac{1}{cv} \tag{1}$$

is a threshold level such that if $\theta > \widehat{\theta}$, the worker needs to be assured of a base salary exceeding the reservation wage $\overline{w}$ in a pure moral hazard model.

In this observable type case, the firm gets the expected outcome $E[y] = e_i^o$ and incurs the expected cost

$$\alpha_i^o + \beta_i^o E[y] = \overline{w} + \frac{1}{2c(1+\theta_i cv)} = \overline{w} + \frac{1}{2} e_i^o$$

hence the firm's expected profit from a worker of type θ_i is

$$\Pi_i^o = (1-\beta_i^o)e_i^o - \alpha_i^o = \frac{1}{2c(1+\theta_i cv)} - \overline{w} = \frac{1}{2} e_i^o - \overline{w}$$

The firm would want to hire the worker only when $\Pi_i^o \geq 0$. Let's assume this.

Assumption 1. For each type θ_i, we have $e_i^o \geq 2\overline{w} \iff \overline{w}c(1+\theta_i cv) \leq 1/2$

This puts a higher bound on the degree of risk aversion θ. (It also puts higher bounds on $\overline{w}$, c and v as well. For example, if $\theta_i = \widehat{\theta} = \frac{1}{cv}$ from (1), then $\overline{w}c \leq 1/4$.) The assumption is made to guarantee that every worker would be employed under complete information regarding θ so that the model is meaningful, but is not critically used in our analysis.

3.2. THE OPTIMAL CONTRACT UNDER BINARY PRIVATE TYPE $\theta \in \{\theta_L, \theta_H\}$

We now check if it is possible for the firm to use a menu of separating incentive contracts when the worker's type is unobservable. Suppose that the only type dimension is the agent's risk aversion coefficient θ and it can take one of two values $\theta_L < \theta_H$, with probability of θ_L being p. Let us make an assumption using $\widehat{\theta}$ from (1):

Assumption 2. $\theta_L < \widehat{\theta} = \frac{1}{cv} < \theta_H$

Under the assumption, type θ_L is only moderately risk averse and need not be assured of α_L^o exceeding $\overline{w}$ under the contract (α_L^o, β_L^o), while type θ_H is more risk averse and requires $\alpha_H^o > \overline{w}$ in the pure moral hazard setting. This greatly simplifies our analysis below.

If the firm offers a menu of linear incentive contracts that truthfully reveal the worker's type, then the stage 2 subgame is *seemingly* identical to the benchmark LEN model, so we might expect the optimal menu to be $\{(\alpha_L, \beta_L), (\alpha_H, \beta_H)\}$ such that, for $i = L, H$,

$$\beta_i = \beta_i^o = \frac{1}{1+\theta_i cv}, \quad \alpha_i = \alpha_i^o = \overline{w} + \frac{\theta_i cv - 1}{2c(1+\theta_i cv)^2}, \quad e_i = e_i^o = \frac{1}{c(1+\theta_i cv)}$$

Since $\theta_L < \theta_H$, we have $\beta_L > \beta_H$ and $e_L > e_H$. By Assumption 2, we have $\alpha_H > \alpha_L$.

We go back to stage 1 and check whether this menu is indeed incentive compatible. Let $A(\theta_j|\theta_i)$ denote the certainty equivalent ("effective compensation") of the linear compensation menu $\{(\alpha_i, \beta_i)\}$ for a worker of type θ_i when he claims he is type θ_j. Then[2]

$$A(\theta_j|\theta_i) = \alpha_j + \beta_j e_j - \frac{1}{2} c e_j^2 - \frac{1}{2}\theta_i \beta_j^2 v = \alpha_j + \frac{1}{2c}\beta_j^2 - \frac{1}{2}\theta_i \beta_j^2 v$$

since $e_j = \beta_j / c$. If less risk averse θ_L-type worker faced with the menu $\{(\alpha_i^o, \beta_i^o)\}$ were to choose (α_H^o, β_H^o), his effective compensation would be

$$\begin{aligned}
A(\theta_H|\theta_L) &= \alpha_H^o + \beta_H^o e_H - \frac{1}{2} c e_H^2 - \frac{1}{2}\theta_L (\beta_H^o)^2 v \\
&= \overline{w} + \frac{\theta_H cv - 1}{2c(1+\theta_H cv)^2} + \frac{1}{c(1+\theta_H cv)^2} - \frac{1}{2c(1+\theta_H cv)^2} - \frac{1}{2}\theta_L \frac{v}{(1+\theta_H cv)^2} \\
&= \overline{w} + \frac{(\theta_H - \theta_L)cv}{2c(1+\theta_H cv)^2} > \overline{w} = A(\theta_L|\theta_L)
\end{aligned}$$

He would prefer to pretend to be of type θ_H! Hence, the benchmark "optimal" menu $\{(\alpha^o, \beta^o)\}$ is *not* incentive compatible in the combined adverse selection-moral hazard model.

[2]For CARA utility function, the certainty equivalent of a normally distributed random variable with mean μ and variance v is $CE = \mu - \frac{1}{2}v^2$. See Bolton and Dewatripont (2005) pp.137–139. Also see Kim (2016) for discussion of the certainty equivalent for more general distributions.

Note that

$$A(\theta_j|\theta_i) - A(\theta_j|\theta_j) = \frac{1}{2}(\theta_j - \theta_i)\beta_j^2 v > 0 \text{ if } \theta_j > \theta_i$$

so we always have

$$A(\theta_H|\theta_L) > A(\theta_H|\theta_H) \tag{2}$$

In words, less risk averse θ_L always enjoys a higher payoff than θ_H from the same contract.

Recognizing the incentive compatibility constraints, the firm's problem is

$$\begin{aligned}\max_{\alpha_j,\beta_j}\ & p[(1-\beta_L)e_L - \alpha_L] + (1-p)[(1-\beta_H)e_H - \alpha_H] \\ &= p[\frac{1}{c}(1-\beta_L)\beta_L - \alpha_L] + (1-p)[\frac{1}{c}(1-\beta_H)\beta_H - \alpha_H]\end{aligned}$$

subject to

$$\begin{aligned} A(\theta_L|\theta_L) &\geq \overline{w} && (PC_L) \\ A(\theta_H|\theta_H) &\geq \overline{w} && (PC_H) \\ A(\theta_L|\theta_L) &\geq A(\theta_H|\theta_L) && (IC_L) \\ A(\theta_H|\theta_H) &\geq A(\theta_L|\theta_H) && (IC_H) \end{aligned}$$

By familiar arguments for adverse selection models (see Salanie, 2005, chapter 2), we expect (PC_L) and (IC_H) to be slack and (PC_H) and (IC_L) to be binding. For example, we can see that

$$A(\theta_L|\theta_L) \underbrace{\geq}_{(IC_L)} A(\theta_H|\theta_L) \underbrace{>}_{(2)} A(\theta_H|\theta_H) \underbrace{\geq}_{(PC_H)} \overline{w} \implies A(\theta_L|\theta_L) > \overline{w} \ \ (\text{slack } PC_L)$$

That (PC_H) and (IC_L) should be binding follows easy arguments: if they are slack, the firm can reduce the slack and increase payoff without affecting other constraints. Assume, for the moment, that (IC_H) holds. (We will check that (IC_H) is slack once we derive the optimal contract formulas.) From binding (PC_H) and (IC_L) we obtain

$$\begin{aligned} \alpha_H &= \overline{w} - \frac{1}{2c}\beta_H^2 + \frac{1}{2}\theta_H\beta_H^2 v = \overline{w} + \frac{1}{2c}(\theta_H cv - 1)\beta_H^2 \\ \alpha_L &= \alpha_H + \frac{1}{2c}(\beta_H^2 - \beta_L^2) - \frac{1}{2}\theta_L(\beta_H^2 - \beta_L^2)v \\ &= \alpha_H + \frac{1}{2c}(\theta_L cv - 1)(\beta_L^2 - \beta_H^2) \\ &= \overline{w} + \frac{1}{2c}(\theta_L cv - 1)\beta_L^2 + \frac{1}{2c}(\theta_H - \theta_L)cv\beta_H^2 \end{aligned} \tag{3}$$

In (3), it is obvious that $\alpha_H \gtrless \alpha_L$ if and only if $\theta_L cv \lessgtr 1$, assuming that $\beta_L > \beta_H$ (which we will shortly confirm). By Assumption 2, we have $\theta_L cv < 1$, hence $\alpha_H > \alpha_L$. In other words, in the optimal contract menu, more risk averse θ_H should receive a higher base salary.

Incorporating binding constraints and ignoring non-binding constraints, we may now rewrite the principal's problem as an unconstrained problem:

$$\max_{\beta_L,\beta_H} \; p[\frac{1}{c}(1-\beta_L)\beta_L - \alpha_L] + (1-p)[\frac{1}{c}(1-\beta_H)\beta_H - \alpha_H]$$
$$= p[\frac{1}{c}\beta_L - \frac{1+\theta_L cv}{2c}\beta_L^2 - \frac{1}{2}(\theta_H - \theta_L)\beta_H^2 v] + (1-p)[\frac{1}{c}\beta_H - \frac{1+\theta_H cv}{2c}\beta_H^2] - \overline{w}$$

Solving this problem yields the optimal incentive intensities as our first result:

Proposition 1. *In the optimal contract under both adverse selection and moral hazard, less risk averse θ_L-type worker faces the same incentive intensity as under pure moral hazard, while more risk averse θ_H-type worker faces a lower incentive intensity, as given by the following expressions.*

$$\begin{aligned} \beta_L^* &= \frac{1}{1+\theta_L cv} && = \beta_L^o \\ \beta_H^* &= \frac{1}{\frac{p}{1-p}(\theta_H - \theta_L)cv + (1+\theta_H cv)} < \frac{1}{1+\theta_H cv} && = \beta_H^o < \beta_L^o = \beta_L^* \end{aligned} \tag{4}$$

The proof involves a straightforward application of the first-order conditions of the unconstrained maximization problem above and is omitted. (4) says that in order to satisfy the incentive compatibility constraints, the firm must make the lower-powered contract for the more risk averse worker even more low-powered in performance pay.

For the contract specified by (3) and (4), that (IC_H) is slack is checked easily:

$$\begin{aligned} A(\theta_H|\theta_H) - A(\theta_L|\theta_H) &= (\alpha_H + \frac{1}{2c}\beta_H^2 - \frac{1}{2}\theta_H\beta_H^2 v) - (\alpha_L + \frac{1}{2c}\beta_L^2 - \frac{1}{2}\theta_H\beta_L^2 v) \\ &= (\alpha_H - \alpha_L) + (\beta_L^2 - \beta_H^2)\frac{\theta_H cv - 1}{2c} \\ &= (\beta_L^2 - \beta_H^2)\frac{(\theta_H - \theta_L)cv}{2c} > 0 \end{aligned}$$

which is positive under the optimal contract, because $\beta_L^* > \beta_H^*$ and $\theta_H > \theta_L$.

We can incorporate (4) into (3) to record more explicit expressions of the base salary:

$$
\begin{aligned}
\alpha_H^* &= \overline{w} + \frac{1}{2c}(\theta_H cv - 1)\left(\frac{1-p}{1+\theta_H cv - p(1+\theta_L cv)}\right)^2 \\
\alpha_L^* &= \alpha_H^* + \frac{1}{2c}(\theta_L cv - 1)(\beta_L^2 - \beta_H^2) \\
&= \alpha_H^* + \frac{1}{2c}(\theta_L cv - 1)\frac{(1+\theta_H cv)^2 - 2p(1+\theta_L cv)(1+\theta_H cv) + (2p-1)(1+\theta_L cv)^2}{(1+\theta_L cv)^2(1+\theta_H cv - p(1+\theta_L cv))^2}
\end{aligned}
\tag{5}
$$

Finally, the effort levels under the optimal contracts are

$$
e_L^* = \frac{\beta_L^*}{c} = \frac{1}{c(1+\theta_L cv)}, \quad e_H^* = \frac{\beta_H^*}{c} = \frac{1}{\frac{p}{1-p}(\theta_H - \theta_L)c^2 v + c(1+\theta_H cv)} \tag{6}
$$

4. ANALYSIS OF THE OPTIMAL CONTRACT AND COMPENSATION

4.1. THE INCENTIVE INTENSITY β_H FOR θ_H

The key difference of the optimal contract $\{(\alpha_i^*, \beta_i^*)\}$ from the benchmark $\{(\alpha_i^o, \beta_i^o)\}$ lies in the incentive intensity β_H for the more risk averse type. The effect of adverse selection on the incentive intensity may be measured by how β_H^* differs from β_H^o. Let's see how β_H^* is affected by the probability p of type θ_L. When p is high, θ_H is a "minority" in the population of potential workers.

Proposition 2. *The optimal incentive intensity β_H^* for more risk averse θ_H-type worker is smaller than the benchmark β_H^o for all $p \in (0,1)$ and decreases as p increases.*

Proof. $\beta_H^* < \beta_H^o$ is seen by simple inspection and was already noted in (4) in Proposition 1.

$$
\begin{aligned}
\frac{\partial \beta_H^*}{\partial p} &= \frac{-(1+\theta_H cv) + p(1+\theta_L cv) + (1-p)(1+\theta_L cv)}{(1+\theta_H cv - p(1+\theta_L cv))^2} \\
&= \frac{(\theta_L - \theta_H)cv}{(1+\theta_H cv - p(1+\theta_L cv))^2} < 0 \quad \square
\end{aligned}
$$

If we plug in $p = 0$ into β_H^* in (4), we obtain the identical expression as β_H^o. (If all workers are of θ_H-type, then the pure moral hazard case obtains.) On the other hand, if we plug in $p = 1$, we get $\beta_H^* = 0$. (If all workers are of θ_L-type, then the 'non-existent' θ_H-type would have gotten a fixed salary.) Of course, either $p = 0$ or $p = 1$ collapses the model into a single-type case, so the expression (4) technically applies to $0 < p < 1$ only. But since the expression is continuous in p, the previous statements apply to very extreme binary distributions: if $p \approx 0$, then $\beta_H^* \approx \beta_H^o < \beta_L^* = \beta_L^o$ and if $p \approx 1$, then $\beta_H^* \approx 0$.

If only a very small proportion ($p \approx 0$) of the potential worker population is of type θ_L, the firm offers a two-tier incentive scheme almost identical to the benchmark case, where types are observable, to attract the small number of less risk averse workers who are willing to take some risk. The incentive power for (a small number of) highly risk averse agents are close to that for the observable type case. On the other hand, if most ($p \approx 1$) workers are of less risk averse type θ_L, then the firm offers an essentially fixed (and very low, since $\alpha_H^* \approx \overline{w}$) compensation scheme for a small number of highly risk averse workers while most workers are induced to choose an incentive compensation scheme. When p is higher (more workers are less risk averse), then it is important to keep them away from low-powered incentive contract (intended for highly risk averse workers) by making it even lower-powered and unattractive.

Table 1 is a summary ($p = 1/2$ is included for illustration).

optimal incentive intensity (unobservable type case)		benchmark (observable type case)
$\beta_L^* = \dfrac{1}{1+\theta_L cv}$		$= \beta_L^o$
$\vee$		
$\beta_H^* \begin{cases} \approx \dfrac{1}{1+\theta_H cv} \\ = \dfrac{1}{1+\theta_H cv + (\theta_H - \theta_L)cv} \\ \approx 0 \end{cases}$	$(p \approx 0)$ $(p = \frac{1}{2})$ $(p \approx 1)$	$= \beta_H^o$

Table 1: Comparison of incentive intensity β_i

Corollary 1. *The equilibrium effort e_H^* chosen by more risk averse θ_H-type worker is lower than the effort e_L^* chosen by less risk averse θ_L-type worker. The gap $(e_L^* - e_H^*)$ increases as p increases.*

Proof. Follows from $e_i^* = \beta_i^*/c$. □

4.2. THE BASE SALARY α_I

Let us now examine the optimal base salary α_H^* and α_L^* and how they respond as p changes. While we derived fuller expressions for α_i in (5), now that we know how β_H^* responds to p, it is easier to work with (3), reproduced in shorter form here:

$$\alpha_H^* = \overline{w} + \frac{1}{2c}(\theta_H cv - 1)(\beta_H^*)^2 \tag{3a}$$

$$\alpha_L^* = \alpha_H^* + \frac{1}{2c}(\theta_L cv - 1)((\beta_L^*)^2 - (\beta_H^*)^2) \tag{3b}$$

$$= \overline{w} + \frac{1}{2c}(\theta_L cv - 1)(\beta_L^*)^2 + \frac{1}{2c}(\theta_H - \theta_L)cv(\beta_H^*)^2 \tag{3c}$$

By the assumption $\theta_H cv > 1$, we ensured that, under pure moral hazard, θ_H-type worker would receive a base salary exceeding reservation wage ($\alpha_H^o > \overline{w}$). This remains true with adverse selection added: in (3a), $\alpha_H^* > \overline{w}$ again by $\theta_H cv > 1$. Since β_H^* decreases as p increases (Proposition 2), α_H^* also decreases as p increases.

As for α_L^*, since $\theta_L cv < 1$ by assumption and $\beta_L^* > \beta_H^*$, we know $\alpha_L^* < \alpha_H^*$. Although we had $\alpha_L^o < \overline{w}$, whether $\alpha_L^* \gtrless \overline{w}$ is ambiguous. In (3c), we see that if θ_L is very close to $1/cv$, then $\alpha_L^* > \overline{w}$ but if θ_L is sufficiently low, then $\alpha_L^* < \overline{w}$. We do know that α_L^* decreases as p increases. Moreover, α_L^* decreases faster than α_H^* as can be seen from (3b).

These findings are collected in the following proposition.

Proposition 3. *The base salary α_H^* for more risk averse θ_H-type worker exceeds the reservation wage $\overline{w}$ and decreases as p increases. The base salary α_L^* for less risk averse θ_L-type worker is lower than α_H^* and decreases faster than α_H^* as p increases. In a more succinct representation, we have*

$$\begin{aligned} \alpha_H^* &= \overline{w} + \alpha_H(p) \\ \alpha_L^* &= \alpha_H^* - \alpha_L(p) = \overline{w} + \alpha_H(p) - \alpha_L(p) \end{aligned} \tag{7}$$

where $\alpha_H(p) > 0$ for all $p \in (0,1)$ and $\alpha_H'(p) < 0$ and $\alpha_L(p) > 0$ for all $p \in (0,1)$ and $\alpha_L'(p) > 0$.

Proof. Let $\alpha_H(p) \equiv \frac{1}{2c}(\theta_H cv - 1)(\beta_H^*)^2$ and $\alpha_L(p) \equiv -\frac{1}{2c}(\theta_L cv - 1)((\beta_L^*)^2 - (\beta_H^*)^2)$. Then the results follows. □

Corollary 2. *If $p \approx 1$, then $a_L^* < \overline{w}$.*

Proof. $\lim_{p\to 1}\alpha_H(p) = 0$ because $\beta_H^* \to 0$, while $\lim_{p\to 1}\alpha_L(p) > 0$. □

To recapitulate, more risk averse θ_H-type worker is ensured of a base salary exceeding the reservation wage but the surplus is smaller when p is higher (less θ_H and more θ_L in the population). Less risk averse θ_L-type always gets a lower base salary than θ_H and the gap is bigger when p is higher. In fact, a θ_L-type worker's base salary falls short of the reservation wage if p is sufficiently high.

4.3. THE TOTAL COMPENSATION

The expected total compensation is determined by the contract terms α, β and the chosen effort level e. Before looking at the total compensation explicitly, let us tabulate what we know about these individual factors. Since we noted in Corollary 1 that the equilibrium effort choice e_i^* moves together with β_i, we need only look at α and β.

benchmark	unobservable type case		
(observable type)	$p \approx 0$	$0 < p < 1$	$p \approx 1$
$\beta_L^o > \beta_H^o$	$\beta_L^o = \beta_L^* > \beta_H^o \approx \beta_H^*$	β_H^* decreases	$\beta_H^* \approx 0$
$\alpha_L^o < \overline{w} < \alpha_H^o$	$\overline{w} < \alpha_H^o \approx \alpha_H^*$	α_H^* decreases	$\alpha_H^* \approx \overline{w}$
	$(\alpha_L^o \neq)\alpha_L^* < \alpha_H^*$	α_L^* decreases	$\alpha_L^* < \overline{w}$

Table 2: Comparison of contract terms

A θ_H-type worker chooses a lower powered incentive contract and a lower level of effort in equilibrium. Such a voluntary low performance choice gets more pronounced when he is a "minority". But we cannot yet conclude that he will be paid less in total, because he gets a higher base salary than the more ambitious θ_L-type worker.

Our final result establishes that the higher base salary advantage cannot overcome the gap driven by incentives and efforts. The proof is not as straightforward as our other results.

Proposition 4. *Let $W_L(p)$ and $W_H(p)$ be the expected total compensation as functions of the probability p of type θ_L, for type θ_L and θ_H, respectively, from the optimal contract under adverse selection and moral hazard. Let $\Delta(p) \equiv W_L(p) - W_H(p)$ be the compensation gap between the two types. Then*

(a) $\Delta(p) > 0$ *i.e.* $W_L(p) > W_H(p)$ *for all* $0 < p < 1$*: More risk averse θ_H-type worker always receives lower total compensation than θ_L-type worker.*

(b) $W_H'(p) < 0$*: As the probability of less risk averse type* θ_L *increases (i.e.* θ_H *becomes a minority), more risk averse* θ_H*-type worker receives lower and lower compensation.*

(c) $\Delta'(p) > 0$*: As* p *increases, the compensation gap between types increases.*

Proof. (a) Since

$$\beta_L^* e_L^* - \beta_H^* e_H^* = \frac{1}{c} \cdot \frac{(1+\theta_H cv)^2 - 2p(1+\theta_L cv)(1+\theta_H cv) + (2p-1)(1+\theta_L cv)^2}{(1+\theta_L cv)^2(1+\theta_H cv - p(1+\theta_L cv))^2}$$

$$\alpha_L^* - \alpha_H^* = \frac{1}{2c}(\theta_L cv - 1)\frac{(1+\theta_H cv)^2 - 2p(1+\theta_L cv)(1+\theta_H cv) + (2p-1)(1+\theta_L cv)^2}{(1+\theta_L cv)^2(1+\theta_H cv - p(1+\theta_L cv))^2}$$

we have

$$\begin{aligned}
W_L - W_H &= (\alpha_L^* - \alpha_H^*) + (\beta_L^* e_L^* - \beta_H^* e_H^*) \\
&= \frac{1}{2c}(1+\theta_L cv)\frac{(1+\theta_H cv)^2 - 2p(1+\theta_L cv)(1+\theta_H cv) + (2p-1)(1+\theta_L cv)^2}{(1+\theta_L cv)^2(1+\theta_H cv - p(1+\theta_L cv))^2} \\
&= \frac{1}{2c(1+\theta_L cv)} \cdot \frac{[1+\theta_H cv - p(1+\theta_L cv)]^2 - (1-p)^2(1+\theta_L cv)^2}{(1+\theta_H cv - p(1+\theta_L cv))^2} \\
&= \frac{1}{2c(1+\theta_L cv)}\left[1 - \underbrace{\left(\frac{1+\theta_L cv - p(1+\theta_L cv)}{1+\theta_H cv - p(1+\theta_L cv)}\right)}_{R\,<\,1\ (\because\ \theta_L<\theta_H)}{}^2\right] > 0 \qquad (8)
\end{aligned}$$

(b) The expression for W_H can be written as

$$\begin{aligned}
W_H &= \alpha_H^* + \beta_H^* e_H^* \\
&= \overline{w} + \frac{1}{2c}(\theta_H cv - 1)\left(\frac{1-p}{1+\theta_H cv - p(1+\theta_L cv)}\right)^2 + \frac{1}{c}\left(\frac{1-p}{1+\theta_H cv - p(1+\theta_L cv)}\right)^2 \\
&= \overline{w} + \frac{1+\theta_H cv}{2c}(\underbrace{\frac{1-p}{1+\theta_H cv - p(1+\theta_L cv)}}_{=\,P})^2
\end{aligned}$$

where only the part denoted by P depends on p, and

$$\frac{\partial P}{\partial p} = \frac{-1-\theta_H cv + p(1+\theta_L cv) + (1-p)(1+\theta_L cv)}{(1+\theta_H cv - p(1+\theta_c v))^2} = \frac{(\theta_L - \theta_H)cv}{(1+\theta_H cv - p(1+\theta_c v))^2} < 0$$

(c) This is *not* a direct consequence of (b), because $W_L'(p) \neq 0$. From Table 2, β_L^* remains constant while α_L^* falls as p increases, which combined with (b) implies our conclusion. More specifically, from (8) in the proof of item (a), only the part denoted by $R(<1)$ depends on p,

$$R = \frac{1+\theta_L cv - p(1+\theta_L cv)}{1+\theta_H cv - p(1+\theta_L cv)} = \frac{1-p}{K-p}, \text{ where } K = \frac{1+\theta_H cv}{1+\theta_L cv} > 1$$

Observe

$$\frac{\partial R}{\partial p} = \frac{-(K-p)+(1-p)}{(K-p)^2} = \frac{1-K}{(K-p)^2} < 0$$

Therefore the gap $W_L - W_H$ increases as p increases. □

4.4. GROUPS WITH OBSERVABLE TRAITS: AN EXPLORATION OF "GENDER PAY GAP"

Suppose that there are groups with different *observable traits*. Following Spence (1974)'s terminology, we shall call the observable trait an *index*. The examples of *indices* include gender, race, physical appearances, geographical origins and the past achievements (information that can be verifiably revealed in résumés). Spence (1974) has shown that apparently discriminatory practices may arise from interactions between indices and unobservable characteristics.

For an illustration, let the index be gender and call the groups "male" (M) and "female" (F). The gender pay gap is a complex and controversial issue with many aspects. We use this only as an illustration of our model. We refer the reader to Blau and Kahn (2017) for a recent comprehensive survey on theories and empirics of gender pay gap. For theoretical and empirical research into gender pay gap related with informational issues, see, e.g. Dohmen and Falk (2011), Santos-Pinto (2012)[3], and Sarin and Wieland (2016).

If gender is not correlated with the unobservable risk attitude (the dummy index case), in other words, if male and female workers have an identical composition of risk attitude types, the gender pay gap cannot arise in our model, unless the firm actively engages in a discriminatory practice. This is because, unlike in Spence (1974)'s signaling model with multiple equilibria, our contracting model has a unique equilibrium for each set of parameters.

If gender is correlated with the risk attitude, then the index contains some information about the private information. Consider the following example.

[3] I thank an anonymous referee for alerting me to this reference.

Example:

Suppose that M's have slightly more of θ_L types and F's have slightly more of θ_H types as reported in table (a) below (see Sarin and Wieland, 2016). Assume that table (a) is common knowledge. Then from our analysis, we have the optimal contracts and the total expected compensation outlined in tables (b) and (c). (Vertical positions reflect relative sizes and are not to the scale.)

	M	F	
θ_L	0.6	0.4	0.5
θ_H	0.4	0.6	0.5
	0.5	0.5	

(a) composition

ND	M	F
β_L^*	β_L^*	β_L^*
		$\beta_H^*(0.4)$
$\beta_H^*(0.5)$		
	$\beta_H^*(0.6)$	

(b) contracts

ND	M	F
	$W_L(0.6)$	
$W_L(0.5)$		
W^M		$W_L(0.4)$
W^F		$W_H(0.4)$
$W_H(0.5)$		
	$W_H(0.6)$	

(c) compensation

In (b) and (c), 'ND' refers to no distinction, or gender-blind contracting. In gender-blind contracting, the gender pay gap is observed in the aggregate because there are more θ_L workers among M so that $W^M = 0.6W_L(0.5) + 0.4W_H(0.5) > W^F = 0.4W_L(0.5) + 0.6W_H(0.5)$. The employer does not discriminate on gender at all, and only indirectly discriminates against the more risk averse agents (both M and F) by lower-powered incentive contract.

If the gender index is used explicitly in contracting (gender-specific contracting), then while the higher-powered incentive contract is the same for both M and F, the (smaller number of) highly risk averse M workers will be led to lower-powered alternative contract.

One prediction is that we will observe a higher compensation disparity, or variance, among M workers than among F workers, when the employer explicitly recognizes gender in hiring and contracting. The employer does discriminate on gender but whether we will observe an aggregate gender pay gap is not obvious. It depends on parametric details. What is obvious is that θ_L-type M workers are better off and θ_H-type M workers are worse off, and θ_L-type F workers are worse off and θ_H-type F workers are better off than under ND case. □

This example is only meant as an illustration of potential application of our results. While highly simplified, it does offer some interesting empirical predictions and policy implications. For instance, in the example, requiring gender-blind contracting does not completely remove the gender pay gap observed in the aggregate data. But the pay gap is between risk aversion types and the gender pay gap is indirectly induced by (slightly different) compositions of risk aversion types in different genders. Furthermore, it is not obvious whether gender-blind contracting would mitigate or aggravate the gender pay gap.

One takeaway of the example is a warning against the use of aggregated data in discussing the pay gap (as is sometimes done in popular press). Apparent pay gaps may or may not be actual pay gaps. When looking at the aggregate data only, one might find a gap where there is none and might miss it where there is. To delve into the pay gap possibly arising from information asymmetry would require more solid empirical investigation as well as extension of our theoretical analysis. We discuss some directions of extensions in the final section.

5. DISCUSSION AND CONCLUDING REMARKS

We have outlined an argument that a minority group of agents with higher risk aversion may choose a job with lower incentive power and lower total compensation. If the minority group's share increases in the population, the compensation disparity is mitigated but remains significant. From Proposition 4,

$$\Delta(0) = W_L(0) - W_H(0) = \frac{1}{2c(1+\theta_L cv)}\left[1 - \frac{(1+\theta_L cv)^2}{(1+\theta_H cv)^2}\right] > 0 \tag{9}$$

measures the compensation disparity arising from moral hazard because as $p \to 0$, the optimal contract terms approach those of pure moral hazard model, while

$$\Delta(p) - \Delta(0) = \frac{1}{2c(1+\theta_L cv)}\left[\frac{(1+\theta_L cv)^2}{(1+\theta_H cv)^2} - \frac{(1+\theta_L cv - p(1+\theta_L cv))^2}{(1+\theta_H cv - p(1+\theta_H cv))^2}\right] > 0 \tag{10}$$

measures the compensation disparity arising from adverse selection as function of p. The source of the composite compensation disparity $\Delta(p) = [\Delta(p) - \Delta(0)] + \Delta(0)$ is information asymmetry. While highly simplified, (9) and (10) point to a way to measure these disparities, or at least to analyze how the disparities are related to model parameters (θ_i, c, v, and p).

Our general point is that information asymmetry (about any agent characteristic) may be a source of the pay gap observed. To empirically measure the contribution of information in total, and adverse selection and moral hazard in separate, would require a more comprehensive model. The disutility of effort c may be considered a private characteristic, which would probably yield a similar conclusion. More challenging is to let the reservation wage $\overline{w}$ to be different and private information (Choi and Park, 2021). We should ultimately consider multi-dimensional type space and selection. (Dohmen and Falk, 2011; Kim, 2019)

On the normative side, many questions arise. Is it equitable to pay someone less because he is more risk averse? What are social or policy implications of our analysis? If we consider a person's risk attitude exogenous, determined by, say, biological or psychological factors, then there is not much we can (or should) do about the pay gap. Technically, it is not different from a pay gap arising from innate talent. If we still want to eliminate the gap because of social concerns, then our analysis suggests that the effort be directed at mitigating information asymmetry, e.g. gathering and making available information regarding private characteristics. If we can discern social factors that affect risk attitude, efforts might be directed at them. Why does someone become more risk averse? Perhaps because of his family background or wealth level or social circumstances. More social and behavioral research will be helpful in answering such questions.

REFERENCES

Blau, F.D. and L.M. Kahn (2017). "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature* 55(3), 789-865.

Bolton, P. and M. Dewatripont (2005). *Contract Theory*, MIT Press.

Choi, G.S. and J. Park (2021). "Optimal Wage Contracts under Reference-Dependent Preferences," *Journal of Economic Theory and Econometrics* 32(1), 54-79.

Dohmen, T. and A. Falk (2011). "Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender," *American Economic Review* 101, 556-590.

Kim, S.H. (2019). "Two-Dimensional Self-Selection of Borrowers," *Korean Economic Review* 35(1), 125-161.

Kim, S.H. (2016). "Risks with Semi-Infinite Support: Characterizations and Applications," *Journal of Economic Theory and Econometrics* 27(4), 1-24.

Kirkegaard, R. (2017). "A Unifying Approach to Incentive Compatibility in Moral Hazard Problems," *Theoretical Economics* 12, 25-51.

Milgrom, P. and J. Roberts (1992). *Economics, Organization & Management*, Prentice Hall.

Myerson, R.B. (1982). "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics* 10, 67-81.

Salanie, B. (2005). *The Economics of Contracts* (2nd ed.), MIT Press.

Santos-Pinto, L. (2012). "Labor Market Signaling and Self-Confidence: Wage Compression and the Gender Pay Gap," *Journal of Labor Economics* 30(4), 873-914.

Sarin, R. and A. Wieland (2016). "Risk Aversion for Decisions under Uncertainty: Are There Gender Differences?," *Journal of Behavioral and Experimental Economics* 60, 1-8.

Spence, A.M. (1974). *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*, Harvard University Press.