GCOE Discussion Paper Series

Global COE Program
Human Behavior and Socioeconomic Dynamics

Discussion Paper No.225

Fair Contracts

Shingo Ishiguro

November 2011

GCOE Secretariat
Graduate School of Economics
OSAKA UNIVERSITY
1-7 Machikaneyama, Toyonaka, Osaka, 560-0043, Japan

Fair Contracts

Shingo Ishiguro*
Graduate School of Economics
Osaka University

November 17, 2011

Abstract

In this paper we present an axiomatic approach to characterize the optimal contracts, which we call "fair contracts," in the general moral hazard model. The two main axioms we employ are incentive efficiency and no–envyness. The incentive efficiency requires that agents of organization select the Pareto efficient contracts among all possible incentive compatible contracts. No–envyness is equity requirement to ensure that each agent does not envy contracts of others in the same organization. We then show that, due to the tension between incentive efficiency and no-envyness, fair contracts have the very simple feature that risk averse agents are offered the fixed wage to choose only the least costly action.

Keywords: Moral Hazard, Incentive Contracts, Fairness.
JEL Classification Number: D82, D86

*Correspondence: Shingo Ishiguro, 1-7 Machikaneyama, Toyonaka, Osaka 560 0043, Japan. e-mail: ishiguro@econ.osaka-u.ac.jp

1 Introduction

In this paper we investigate the moral hazard model in which agents choose unobservable actions and provide an axiomatic approach to characterize the optimal contracts, which we call "fair contracts," in the view point of both efficiency and equity. The standard approach to the moral hazard problem (for example see Grossman and Hart (1983), Rogerson (1985), Jewitt (1988)) has focused only on the efficiency in the sense that the principal chooses a contract to maximize her expected payoff subject to the set of constraints as follows. (i) incentive compatibility (IC) constraint: contract must give the agent the proper incentive to choose a right action because his action is not observable. (ii) individual rationality (IR) constraint: contract should give the agent at least the reservation utility. (iii) Budget balanced (BB): total incomes of both parties sum up to the realized total revenue. We then say that a contract satisfies *incentive efficiency* (IE) if it satisfies (IC), (BB) and there exist no other Pareto improving contracts which satisfy (IC), (BB). In the standard principal-agent model a contract is selected in favor of the principal from all (IE) contracts by satisfying (IR) of the agent.

We provide a slightly different approach to address the optimal contract design problem in the moral hazard models. We impose several axioms which must be taken into account when designing contracts to agents in organization and then characterize the contracts satisfying these axioms. The axioms we will employ include not only the standard efficiency criterion such as (IE) mentioned above but also a sort of equity concern among agents, which we will call *no-envyness* (NE). Here (NE) requires that each agent does not envy any contracts offered to others within the same organization. This is analogous to the concept of "no-envyness" used in the literature of social choice theory (see Thosmson (2010) for a comprehensive survey on this topic): a resource allocation is said be no-envy if every individual does not envy the consumption bundles allocated to others.[1] We will modify this idea in the context of the principal-agent model and utilize it as an equity condition in order to characterize the optimal contracts.

In addition to (IE), (IR) and (NE), we also consider the possibility that the wage scheme of each agent is non-decreasing in his own verifiable performance (output). We call this axiom *monotonicity* (MON). This is also a reasonable requirement because most wage schemes used in practice satisfy this condition. We can also give theoretical justification for this condition by considering ex post moral hazard of agents as follows. Each agent can secretly dispose his own output ex post after it is realized. Then, if his wage is decreasing in some parts of his outputs, the agent has the incentive to discard some amount of the realized output and obtain higher wages. Thus (MON) must be satisfied for ruling out such ex post moral hazard (see Innes (1990) and Matthews (2001) for a similar argument).

We call a contract profile of agents in organization *fair contract* when it satisfies

[1] See also Foley (1967), Pazner and Schmeidler (1974, 1978) and Varian (1974) for the classic works on equity in economics.

all these axioms (IE), (NE), (MON) and (IR).[2] Then we show that fair contracts have the very simple feature as follows: every risk averse agent is offered a fixed wage so that he chooses the least costly action. On the other hand, risk neutral agents are motivated to choose the actions which maximize the surplus of organization, given all risk averse agents choose the least costly action. Hence the optimal contracts ("fair contracts" in our notion) become drastically simple once we take into account not only the incentive efficiency but also non-envyness for contracting in organizations.

The basic intuition behind our characterization result is the following: since risk averse agents are assumed to be heterogeneous with respect to their abilities for performing tasks, they must be offered different wage schemes which differ in the degrees of how much their wages reflect their verifiable performances. If a wage scheme responds more highly to the changes of the performances than another, it is said to be higher-powered incentive scheme than the other. Put differently, a higher-powered incentive scheme has "steeper slope" of reward with respect to the performances than a lower-powered scheme. Of course, higher-powered incentive scheme involves more risk on risk averse agent than lower-powered one. Then consider a fair contract which induces risk averse agents with different abilities to work hard. High ability agents can improve the probability distributions of their performances at lower marginal action costs than the low abilities. Thus, less risk is needed for the high ability agents to work as hard as the low abilities. Then, in the view point of incentive efficiency the high ability agents should be offered lower powered incentive scheme than the low abilities because the former can work as hard as the latter at smaller risk. However, then the high ability agents envy the higher-powered incentive scheme offered to the low ability agents because the former can exploit their (high) ability under the wage scheme which more highly responds to the verifiable performances. When facing such conflict between incentive efficiency and no-envyness, the organization finds it impossible to have risk averse agents with different abilities work hard together. As a consequence, any fair contract must involve a fixed wage for different risk averse agents who are induced to choose only the least costly action.

Our model has the relative advantage to the standard moral hazard models in that we can provide a simple characterization result even in general production environment. In the standard approach to the moral hazard problem, it is quite difficult to give sharp predictions about exact features of optimal contracts: how do optimal wages vary with verifiable performances? What shapes does the optimal wage schedule display? Is it linear or non-linear? In contrast our result shows that the optimal contracts become drastically simple once we take into account not only efficiency but also some sort of equity among agents in organization.

Beside the theoretical interest, it is also important to extend the standard

[2]This is also analogous to the concept of "fair allocation" used in the literature of social choice theory. An allocation is said to be fair if it satisfies Pareto efficiency and no-envyness (Thomson (2010)). In our context the Pareto efficiency is replaced by incentive efficiency (IE) due to the presence of the moral hazard problem.

principal-agent models to encompass some notions of equity or fairness. This is because in real world there are organizations which may seek different objectives from maximizing the principal's profit, e.g., non-profit organizations, governments, and state–owned enterprises. These organizations may take into account the welfare of employees, citizens and so on, and may pursue not only efficiency but also some kinds of fairness, e.g., governments would try to design "fair" tax systems. Even in profit organizations equity is often one of the major concerns as well as efficiency and profitability. For example, top management of corporations might try to avoid conflict among employees who are concerned about their relative wages in the same workplaces.

Related Literature

There are two strands of the literature which are related to our paper. First, several recent papers incorporate the aspects of fairness into the standard principal-agent models.[3] However, in these papers equity concern and other-regarding preferences are directly introduced into the payoff functions of agents in some specific ways: for example, the payoff function of an agent is assumed to depend on relative incomes between him and his colleagues. Our approach is different from these papers because we do not impose direct specifications of fairness concern on the payoff functions of agents but rather we use an axiomatic approach by using (NE) as a natural axiom. Thus our model can avoid the problem regarding what specification of the payoff function is the most plausible to describe the preference over equity and fairness. Second, some papers show the optimality of simple contracts in the moral hazard environments. Holmström and Milgrom (1987) show that the optimal reward to agent becomes linear with respect to the eventual outcomes in the dynamic model in which agent controls the drift rate of the stochastic process of outputs (see also Hellwig and Schmidt (2002) for further elaboration of the model). Our paper is different from theirs because we focus on the standard static moral hazard environment only except for incorporating the concept of no-envyness as an equity concern. Holmström and Milgrom (1994) also show that the optimal contract to agent becomes fixed wage in the multi-task context in which agent performs multiple tasks at a time. Although our model also shows the optimality of the fixed wage as in theirs, our approach is different from theirs again because we investigate the tension between incentive efficiency and no-envyness, which results in the lower-powered incentive schemes.

The remaining sections are organized as follows: in Section 2 we set up the basic model of moral hazard in an organization with multiple agents. In Section 3 we provide a characterization result that the optimal contracts which satisfy incentive efficiency and no–envyness become the fixed wage scheme for all risk averse agents

[3]See for example Bartling (2010), Englmaier and Wamback (2010), Fehr and Schmidt (2000, 2004), Itoh (2004), Neilson and Stone (2010) and Rey-Biel (2008) for recent development of agency models incorporating inequity aversion.

when we impose some technical conditions on the probability distributions of the verifiable performances. In Section 4 we will show more general characterization result without imposing the specific conditions on the information structure, given the heterogeneity among risk averse agents is sufficiently small. In Section 5 we will proceed to show the existence proof of fair contract. Our proof is constructive. We find a fair contract which satisfies all the axioms we have mentioned, (IE), (NE), (MON) and (IR), under certain conditions. In Section 6 we will discuss the robustness and extensions of our results: first, we show that our characterization results are not substantially changed by dropping the axiom of monotonicity (MON). Second, we will discuss the implications about the priority of efficiency and equity for designing incentive contracts. Our concept of fair contracts is interpreted as the one we are capturing the situation in which agents of organization first seek the incentive efficiency and then, as the second step, they select the contracts which satisfy no-envyness as well as other axioms. We will call such contracts *efficiency–first–optimal contracts*. One might however think that in some organizations equity may be taken as the first priority before considering the incentive efficiency. This might be suitable in non–profit organizations and public sectors. In such situations agents of organization first identify the set of no–envy contracts and then, as the second step, they select the Pareto efficient contracts among all possible no-envy contracts together with other axioms. We will call such contracts *equity-first-optimal contracts*. We then show that the set of equity–first-optimal contracts is not smaller than the set of efficiency–first–optimal contracts. This result implies that there may exist the possibilities that high–powered incentive contracts are offered to risk averse agents even when heterogeneity among agents is so small that all efficiency–first–optimal contracts must have the fixed wage schemes of risk averse agents. Thus incentive schemes offered in organizations which take equity as the first priority are more likely to be higher powered than those in organizations which take efficiency as the first priority.

2 Model

We consider an organization (or a team) which consists of N agents. Let I be the set of all agents where $\#I = N$. Let I_r be the set of risk averse agents where $N_r = \#I_r$ and I_n the set of risk neutral agents where $N_n = \#I_n$. Thus $I = I_r \cup I_n$. We assume that there exist at least two risk averse agents and one risk neutral agent in organization:

Assumption 1. $N_r \geq 2$ and $N_n \geq 1$.

Agent i chooses an unobservable action $a^i \in A$ where A is finite and $\#A = M + 1 \geq 2$. We denote by $a_n \in A$ a generic element of A where $a_0 \equiv 0 \leq a_1 \leq a_2 \leq \cdots \leq a_M$. Let $\mathbf{a} \equiv (a^1, a^2, ..., a^N)$ be a vector of actions taken by all agents.

Agent i generates a verifiable performance $y^i \in Y$. We can interpret y^i as output of agent i which is used to yield the revenue of organization as we will explain more below. We assume that Y is a finite set and $\#Y = L \geq 2$. Let y_n^i be a generic element of Y. We denote by $P_n^i(a) \in (0, 1)$ the probability of the performance y_n^i of agent i being realized conditional on his action $a \in A$. Let $\mathbf{y} \equiv (y^1, y^2, ..., y^N)$ be a vector of the realized performances of all agents. We assume that the performances $\mathbf{y}$ are independently distributed among agents.

Agents are heterogeneous in terms of their abilities for performing the tasks assigned to them. Specifically, agents differ from each other in terms of the probability distributions of their performances $(P_n^i(a^i))_{n=1}^L$ and their action costs $G_i(a)$. Agent i has the following standard utility function which is separable between utility on income w and disutility on action $a \in A$:

$$u_i(w) - G_i(a) \tag{1}$$

where G_i is increasing and we normalize $G_i(0) = 0$. Here we assume that the utility functions on incomes are identical across all risk averse agents. Thus we set $u_i = u$ for all $i \in I_r$. Let $\phi \equiv u^{-1}$ be the inverse function of u. We also suppose that the utility function on income of risk neutral agent is given by $u_i(w) = w$ for each $i \in I_n$.

Remark. We can allow all A, Y and u to depend on identities of agents. However, our main results are not substantially changed as long as the differences in these characteristics are small.

The organization can generate the total revenue $R(\mathbf{y})$ which depends on a profile of performances (outputs) of agents $\mathbf{y}$. R is increasing in each argument. In what follows we will use notations as follows: let $E_y[\cdot|\mathbf{a}]$ be expectation over the performances of all agents $\mathbf{y} \in Y^N$ conditional on an action profile of them $\mathbf{a}$. Let $E_{y^i}[\cdot|a^i]$ be expectation over the verifiable performance y^i of agent i conditional on his action a^i. Let also $E_{y^{-i}}[\cdot|a^{-i}]$ denote expectation over the verifiable performances of all other agents than i conditional on their actions $a^{-i} = (a^j)_{j \neq i}$.

Since the performances (or outputs) of agent $\mathbf{y}$ are verifiable but their actions $\mathbf{a}$ are not, a wage scheme of agent i should be contingent only on the realizations of $\mathbf{y}$. A contract of agent i is defined by $C^i = \{w^i(\mathbf{y}), \hat{a}^i\}$ which specifies $w^i(\mathbf{y})$ the wage scheme depending on realization of verifiable performances $\mathbf{y}$ and an action $\hat{a}^i \in A$ to be taken. Let $C = (C^i)_{i \in I}$ be a profile of contracts. We call a wage scheme $w^i(\mathbf{y})$ *fixed wage* when

$$w^i(\mathbf{y}) = \overline{w}$$

for some constant $\overline{w}$ for all $\mathbf{y} \in Y^N$.

We suppose that the wage schemes $\{w^i(\mathbf{y})\}_{i \in I}$ must be budget balanced ex post, i.e. the total wages of all agents should be equal to the organization revenue, as follows:

$$\sum_{i \in I} w^i(\mathbf{y}) = R(\mathbf{y}) \quad \forall\, \mathbf{y} \in Y^N. \tag{BB}$$

We denote by $\mathcal{C}$ the set of all contracts satisfying (BB). We will restrict our attention to the class of the wage schemes in $\mathcal{C}$ in what follows.

Next we will consider several axioms to be satisfied in organization.

First, we define the efficiency concept in the presence of the moral hazard problem. One reasonable notion of efficiency is *incentive efficiency* (IE) which is defined as the Pareto efficient contracts satisfying *incentive compatibility* (IC) of agents. (IC) requires the following: since actions of agents are not observable, a contract profile $C \in \mathcal{C}$ should give them the proper incentives to choose specified actions $\hat{\mathbf{a}} = (\hat{a}^i)_{i \in I}$. In other words a contract profile $C \in \mathcal{C}$ should require that every agent follows the instruction to choose specified action $\hat{a}^i$ given all others doing so:

$$E_y[u_i(w^i(\mathbf{y}))|\hat{a}^i, \hat{a}^{-i}] - G_i(\hat{a}^i) \geq E_y[u_i(w^i(\mathbf{y}))|a', \hat{a}^{-i}] - G_i(a') \ \ \forall\, a' \neq a^i. \tag{IC}$$

Here agents simultaneously choose their actions and hence each of them takes the actions of others as given when he chooses his action because, by assumption, agents cannot observe their actions each other. Thus an action profile $\hat{\mathbf{a}}$ should be a Nash equilibrium in the game in which agents simultaneously choose their actions.

We call a contract profile $C \in \mathcal{C}$ *incentive feasible* if it satisfies (IC). We denote by $\mathcal{C}^{\mathcal{F}} \subset \mathcal{C}$ the set of all incentive feasible contracts.

Although there are several ways to define the efficiency in the presence of private information, we follow the argument by Holmström and Myerson (1983) and define the efficiency on the ground of contract satisfying (IC). More formally, take an incentive feasible contract $C \in \mathcal{C}^{\mathcal{F}}$ and define the resulting expected utility of agent i as

$$U^i(C^i) \equiv \sum_n E_y[u_i(w^i(\mathbf{y}))|\hat{\mathbf{a}}] - G_i(a^i). \tag{2}$$

Then, our first axiom is *incentive efficiency* (IE) as follows:

Incentive Efficiency (IE): *An incentive feasible contract $C \in \mathcal{C}^{\mathcal{F}}$ is **incentive efficient** if there exist no other incentive feasible contract $C' \in \mathcal{C}^{\mathcal{F}}$ with $C' \neq C$ such that $U^i(C^{i'}) \geq U^i(C^i)$ for all $i \in I$ with strict inequality being held for at least one agent.*

In the standard principal-agent model it has been assumed that one party, say the principal, chooses her optimal contract among all (IE) contracts given the agent receives at least the reservation utility. In other words the optimal contract is selected in favor of the principal on the Pareto frontier constrained by (IC). In this paper we do not make any restrictions on bargaining powers of agents in organization but simply we require that a contract must be incentive efficient.

In addition to (IE), we impose an axiom representing a sort of equity requirement, which we call *no–envyness* (NE), as follows:

$$E_y[u_i(w^i(\mathbf{y}))|\hat{a}^i, \hat{a}^{-i}] - G_i(\hat{a}^i) \geq \max_{a' \in A} E_y[u_i(w^j(\mathbf{y}))|a', \hat{a}^{-i}] - G_i(a') \ \ \forall\, j \neq i, \ \ a' \in A. \tag{NE}$$

(NE) states that each agent i does not envy the wage scheme of others $w^j(\mathbf{y})$ compared to his own scheme $w^i(\mathbf{y})$. If agent i were offered the wage scheme of agent j, he imagines that he could receive the expected utility corresponding to the right hand side of (NE) given he will choose his action $a' \in A$ appropriately and all others choose the specified actions $\hat{a}^{-i}$. (NE) then requires that agent i does not envy the contract offered to others.

Note that (NE) is analogous to the concept of no-envyness used in the literature of social choice theory (Thomson (2010)). In our model each agent imagines what would happen if he were offered a contract of other agent before he chooses his action. In this sense our notion of no-envyness is imposed at the ex ante stage before agents choose actions. However, an alternative condition for no-envyness is also possible. For example, some agent may envy others after they chose actions and the final performances were realized. We will discuss this ex post case in the concluding remarks later.

We also impose individual rationality (IR) as an axiom to be satisfied. We denote by $\overline{U}_i$ the reservation utility of agent i who would receive outside the organization if he rejected the contract C^i and left the organization. Then a contract C^i should give agent i at least the reservation utility $\overline{U}^i$, i.e. it should be individually rational (IR):

$$E_y[u_i(w^i(\mathbf{y}))|\hat{\mathbf{a}}] - G_i(\hat{a}^i) \geq \overline{U}^i \ \ \forall\, i \in I. \tag{IR}$$

Finally, we impose *monotonicity* (MON) that the wage scheme of each agent is non-decreasing in his own verifiable performance (output). This is stated formally as follows. For all $i \in I$,

$$w^i(y^{i''}, y^{-i}) \geq w^i(y^{i'}, y^{-i}) \ \ \forall\, y^{i''} > y^{i'},\ \forall\, y^{-i} \tag{MON}$$

Many contracts used in practice satisfy such monotonicity and hence this is a reasonable condition. As in Innes (1990) and Matthews (2001), we can also give a theoretical justification for this as well: each agent can secretly dispose his output y^i before he provides it to the organization.[4] Then, if $w^i(y'', y^{-i}) < w^i(y', y^{-i})$ for some $y'' > y'$, agent i will dispose $y'' - y' > 0$ and obtain higher wage $w^i(y', y^{-i})$ even when y'' is realized. Thus w^i must be non-decreasing in y^i.

Remark. We will discuss below that we can dispense with (MON) (see Section 6).

We call a contract profile $C = (C^i)_{i \in I} \in \mathcal{C}$ *fair contract* if it satisfies all the above axioms (IE), (NE), (IR) and (MON). Let C^f denote the set of all fair contracts. It may be the case that $C^f = \emptyset$.

[4] We assume here that the supplied output *after* the decision to dispose was made is verifiable but it is not verifiable what outputs each agent actually produced *before* the decision to dispose is made.

3 Characterization of Fair Contracts

In this section we will characterize the fair contracts. To obtain sharper characterization result, we modify the notion of incentive efficiency slightly as follows:

Modified Incentive Efficiency (IE*): *Let $\tilde{\mathcal{C}}^{\mathcal{F}} \subset \mathcal{C}$ be the set of all contracts satisfying (IC) and (MON). Then a contract profile $C \in \tilde{\mathcal{C}}^{\mathcal{F}}$ is said to be modified incentive efficient, denoted by (IE*), if there exist no other contarcts $C' \in \tilde{\mathcal{C}}^{\mathcal{F}}$ with $C' \neq C$ such that $U^i(C^{i'}) \geq U^i(C^i)$ for all $i \in I$ with strict inequality being held for at least one agent.*

Let C^f_* denote the set of all contracts satisfying (IE*), (NE) and (IR).

Then we can verify that $C^f \subseteq C^f_*$.

Lemma 1. $C^f \subseteq C^f_*$.

Proof. Take $C \in C^f$ but suppose that $C \notin C^f_*$. Then, since C satisfies (NE) and (IR) but $C \notin C^f_*$, C must not satisfy (IE*). Since C satisfies (IC) and (MON), we have $C \in \tilde{\mathcal{C}}^{\mathcal{F}}$. Then, since C is not modified incentive efficient, there must exist a different contract $C' \neq C$ where $C' \in \tilde{\mathcal{C}}^{\mathcal{F}}$ such that C' Pareto dominates C. However, since $C' \in \tilde{\mathcal{C}}^{\mathcal{F}}$ implies that C' satisfies (IC), we must have $C' \in \mathcal{C}^{\mathcal{F}}$. Then, C' is incentive feasible and Pareto dominates C but this contradicts the fact that C is incentive efficient. Q.E.D.

In this section we will focus on the set C^f_* which is larger than C^f. Then we will give a characterization result on C^f_* which is sharper than the characterization result on the set C^f. With a slight abuse of definition, we call $C \in \tilde{\mathcal{C}}^{\mathcal{F}}$ fair contract in the same way as defined in the previous section. Then we will show that all fair contracts $C \in \tilde{\mathcal{C}}^{\mathcal{F}}$ must involve fixed wages offered to risk averse agents under certain conditions. Then Lemma 1 implies that all contracts in C^f must also have fixed wages offered to risk averse agents.

We take any fair contract $\hat{C} \in C^f_*$ where $\hat{C}^i = \{w^i(\mathbf{y}), \hat{a}^i\}$ for each $i \in I$ and fix it in this section.

We first show the following result:

Lemma 2. *Any fair contract $\hat{C}^i$ offered to risk averse agent $i \in I_r$ who is induced to choose a higher action $\hat{a}^i > 0$ than the least costly one (zero) must be individualistic in the sense that $\hat{w}^i(\mathbf{y}) = \hat{w}^i(y^i)$ for all $i \in I_r$.*

Proof. Appendix.

Since the performance signals are statistically independent among agents, it is efficient for wage scheme of every risk averse agent to depend solely on his own performances. If this is not the case for some risk averse agent, it is possible to design the new wage scheme of that agent such that it induces the same expected utility and the same action as under the original scheme while reducing the risk imposed on him by eliminating the dependence of all other agents' performances from his wage scheme. Then some risk neutral agents are better off by reducing the risk on the risk averse agents while unchanging their incentives. Thus (IE*) requires that the wage scheme of every risk averse agent must be individualistic.

Due to Lemma 2 we will denote by $w^i(y^i)$ the wage scheme offered to risk averse agent i, which solely depends on his own performance $y^i = y_n \in Y$.

Next we consider the following implementation problem where the expected wage of risk averse agent $j \in I_r$ is minimized when implementing an action $a \in A$ from him:

Problem (CM$^j - a$)

$$\min_{w(y_n)} \sum_n P^j_n(a)w(y_n)$$

subject to

$$w(y'') \geq w(y') \;\; \forall \, y'' > y' \qquad \text{(MON)}$$

$$\sum_n P^j_n(a)u(w(y_n)) - G_j(a) \geq \sum_n P^j_n(a')u(w(y_n)) - G_j(a'), \;\; \forall \, a' \neq a \qquad \text{(IC)}$$

$$\sum_n P^j_n(a)u(w(y_n) - G_j(a) \geq \hat{U}^j \qquad \text{(IR}^*\text{)}$$

where

$$\hat{U}^j \equiv \sum_n P^j_n(a)u(\hat{w}^j(y_n)) - G_j(\hat{a}^j)$$

denotes the expected payoff of agent i under the supposed fair contract $\hat{C}^j$.

Let $\hat{w}^j(y|a)$ be the wage scheme which solves the above problem (CM$^j - a$) to implement $a \in A$ from risk averse agent j. Let $a = \hat{a}^j$ in Problem (CM$^j - a$). Then the optimal wage scheme $\hat{w}^j(y|a)$ which solves Problem (CM$^j - \hat{a}^j$) exists because the fair contract $\hat{C}^j$ satisfies all (MON), (IC) and (IR*) by definition and hence the constraint set is non-empty.[5] Also the optimal scheme $\hat{w}^j(y|a)$ is unique.[6]

Since a contract $\hat{C}^i$ must be incentive efficient, it must solve Problem (CM^i − $\hat{a}^i$) for each risk averse agent $i \in I_r$. Otherwise there exists a different contract

[5] By a standard argument (Grossman and Hart (1983)), we can verify that the constraint set is compact as well. Then the existence of the optimal wage scheme directly follows from the continuity of the objective function.

[6] To see the uniqueness, define $w_n \equiv \phi(u_n)$ for each $y_n \in Y$ and note that (IC) and (IR*) are both linear constraints with respect to (u_n). Thus the constraint set, denoted by Γ, is convex: take any $(u'_n)_n$ and $(u''_n)_n$ in Γ, and $\lambda u'_n + (1-\lambda)u''_n$ for $\lambda \in (0,1)$. Then we can verify that $\phi(\lambda u'_n + (1-\lambda)u''_n) - \phi(\lambda u'_{n-1} + (1-\lambda)u''_{n-1}) \geq \phi'(\lambda u'_{n-1} + (1-\lambda)u''_{n-1})(\lambda(u'_n - u'_{n-1}) + (1-\lambda)(u''_n - u''_{n-1})) \geq 0$ due to convexity of ϕ and $(u'_n), (u''_n) \in \Gamma$ so that they satisfy (MON). Thus $\lambda u'_n + (1-\lambda)u''_n \in \Gamma$ and hence Γ is convex.

$C^i \neq \hat{C}^i$ which satisfies (MON), (IC) and (IR*) for agent i but yields a smaller expected wage than the minimum attained in Problem (CM$^i - \hat{a}^i$). By replacing $\hat{C}^i$ by such alternative contract C^i, some risk neutral agent can be better off because the organization can save the wage payment to the risk averse agent and gives some risk neutral agent such saved amount while still implementing the same action profile.

This is formally shown as in the following lemma:

Lemma 3. *Suppose that* $\hat{C}^i = \{\hat{w}^i(y^i), \hat{a}^i\}$ *is a fair contract of risk averse agent* $i \in I_r$. *Then* $\hat{C}^i$ *should solve Problem (CM*$^i - \hat{a}^i$*) to implement* $\hat{a}^i$ *from agent i, i.e.,* $\hat{w}^i(y) = \hat{w}^i(y|\hat{a}^i)$ *for each* $y \in Y$ *and for each risk averse agent* $i \in I_r$.

Proof. Appendix.

From now on, we assume that risk averse agents are heterogeneous in decreasing order in terms of their abilities. To this end, we introduce some notation as follows: let $\Delta G_i(\hat{a}, a) \equiv G_i(\hat{a}) - G_i(a) > 0$ be the marginal action cost of agent i for $\hat{a} > a$. Let $F_n^i(a)$ be the distribution function of agent i's performance conditional on his action $a \in A$, defined by $F_n^i(a) \equiv \sum_{l \leq n} P_l^i(a)$.

Then we impose the following weak condition:

Assumption 2. *(i) For any* $i \in I_r$ *and any* $\hat{a} > a$, *we have* $F_n^i(a) > F_n^i(\hat{a})$ *for any* $n \leq L - 1$. *(ii) For any* $i \neq j$, $i, j \in I_r$, *and any* $a \in A$, $F_n^i(a) > F_n^j(a)$ *holds for all* $y_n \in Y$ *when* $i > j$.

Assumption 2 (i) simply states that higher action can improve the probability distribution of performances for any risk averse agent in the sense of the first order stochastic dominance. Assumption 2 (ii) says that the risk averse agent indexed by lower number j can improve the probability distribution of the performances in the sense of the first order stochastic dominance than the risk averse agent indexed by higher number $i > j$.

In addition to Assumption 2 we impose further restrictions on the ability ranking among different risk averse agents. We define the following: for any two risk averse agents $k = i, j \in I_r$, $i \neq j$, and any two actions $\hat{a} \neq a$,

$$\lambda_n^k(\hat{a}, a|i, j) \equiv \frac{F_n^i(a) - F_n^j(\hat{a})}{F_n^k(a) - F_n^k(\hat{a})} \tag{3}$$

and

$$\overline{\lambda}^k(\hat{a}, a|i, j) \equiv \max_n \lambda_n^k(\hat{a}, a|i, j),$$

$$\underline{\lambda}^k(\hat{a}, a|i, j) \equiv \min_n \lambda_n^k(\hat{a}, a|i, j).$$

Here note that $\lambda_n^k(\hat{a}, a|i, j)$ is well-defined due to Assumption 2.

To see what $\lambda_n^k(\hat{a}, a|i, j)$ says about, suppose that risk averse agent i chooses an action $a \in A$ and, as a thought experiment, imagine that his type is changed from i to j as well as his action is also changed from a to $\hat{a}$. Then the distribution function of the performance y^i is changed from $F_n^i(a)$ to $F_n^j(\hat{a})$. Such total change $F_n^j(\hat{a}) - F_n^i(a)$ can be decomposed into either

$$F_n^i(a) - F_n^j(\hat{a}) = (F_n^i(a) - F_n^i(\hat{a})) + (F_n^i(\hat{a}) - F_n^j(\hat{a})) \tag{4}$$

or

$$F_n^i(a) - F_n^j(\hat{a}) = (F_n^i(a) - F_n^j(a)) + (F_n^j(a) - F_n^j(\hat{a})). \tag{5}$$

Such decomposition is made along with two different roots: first, type is changed from i to j by fixing an action. Second, action is changed from a to $\hat{a}$ by fixing a type. In this respect $\lambda_n^k(\hat{a}, a|i, j)$ can be interpreted as the inverse measure about how much the change in agent k's action ($a \to \hat{a}$) contributes to the total change of $F_n^i(a) - F_n^j(\hat{a})$ relative to the change of types ($i \to j$). Then $1/\underline{\lambda}^k(\hat{a}, a|i, j)$ (resp. $1/\overline{\lambda}^k(\hat{a}, a|i, j)$) represents the maximum (resp. minimum) influence of such relative contribution of agent k's action to the total change $F_n^i(a) - F_n^j(\hat{a})$.

Then we make the following assumption:

Assumption 3. *For any two risk averse agents $i, j \in I_r$ and any action $\hat{a} > a$, if $i > j$, then*

$$\frac{\Delta G_i(\hat{a}, a)}{\Delta G_j(\hat{a}, a)} > \frac{\overline{\lambda}^j(\hat{a}, a|i, j)}{\underline{\lambda}^i(\hat{a}, a|i, j)}.$$

The left hand side of the inequality in Assumption 3 is the marginal action cost of agent i relative to agent j. Its right hand side is the ratio between the maximal influence of the relative contribution by agent i's action and the minimal influence of the relative contribution by agent j's action as we have defined above. Thus the right hand side represents the upper bound for the relative contribution of agent i's action as compared to agent j. Then Assumption 3 says that agent i incurs larger marginal action cost than agent j even when he performs his task at the best compared to agent j. In this sense we can say that agent i is less efficient than agent j.

Given all these assumptions, by recalling that $\#I_r = N_r$ and $\#A = M + 1$, we can show the following result.

Theorem 1. *Suppose that Assumption 1-3 hold. Suppose also that $M + 1 \leq N_r$. Then any fair contract must have a fixed wage for at least $N_r - M$ risk averse agents.*

Proof. Appendix.

Theorem 1 states that offering a fixed wage to risk averse agents becomes prevalent feature in large organizations in which the number of risk averse agents is so large relative to the number of actions each agent can take.

The intuition behind Theorem 1 is as follows. Since the number of actions $M+1$ is larger than the number of risk averse agents N_r ($M+1 \geq N_r$), if only less than $N_r - M$ risk averse agents choose the least costly action, there must be at least two risk averse agents who choose the same action $\hat{a} > 0$. Let these agents be i and j. Then, by Assumption 3 we can label them as $i > j$ without loss of generality. Here agent i is less efficient than j in the sense of Assumption 3. Then we can show that, when both i and j choose the same and higher action than the least costly one, (IC) of more efficient agent (agent j) must be slack: the wage scheme of less efficient agent i must be higher-powered than that of more efficient agent j when implementing the same action $\hat{a} > 0$ from both agents. However, agent j then envies agent i since the former can exploit more gains of higher ability from higher-powered scheme than the latter. Thus no-envyness and incentive efficiency cannot be compatible with each other. Therefore, fair contract must involve a fixed wage for at least $N_r - M$ risk averse agents.

When the number of verifiable performances $y^i \in Y$ is two ($\#Y = 2$), Assumption 3 is greatly simplified as follows. When $\#Y = 2$, we have $y = y_1$ (low outcome) or $y = y_2$ (high outcome). Then Assumption 3 is satisfied as long as $\Delta G_i(\hat{a}, a)/\Delta G_j(\hat{a}, a) > (P_2^i(\hat{a}) - P_2^i(a))/(P_2^j(\hat{a}) - P_2^j(a))$ for all $\hat{a} > a$ because then we have $\underline{\lambda}^k(\hat{a}, a) = \overline{\lambda}^k(\hat{a}, a) = (F_1^i(a) - F_1^j(\hat{a}))/(F_1^k(a) - F_1^k(\hat{a}))$ and $F_1^k(a) = 1 - P_2^k(a)$. Here $P_2^k(\hat{a}) - P_2^k(a)$ denotes the marginal increase in success probability of obtaining high outcome y_2 from exerting higher action $\hat{a} > a$. Thus this condition simply states that the marginal action cost of agent i relative to that of agent j is larger than the marginal increase in success probability of agent i relative to that of agent j.

4 More General Characterization: Small Heterogeneity

We will next consider an alternative condition to dispense with both $N_r \geq M+1$ and Assumption 3. In particular, we will show that the same characterization result as Theorem 1 holds even when we drop $N_r \geq M+1$ and Assumption 2 if, instead, we assume that heterogeneity among risk averse agents is sufficiently small.

We first define the *marginal* expected revenue of organization with respect to action of agent i, given action of agent j, a^j, and the actions of all others, denoted by $a^{-i-j} \equiv (a^k)_{k \neq i,j}$, as follows:

$$\Delta_i(a^j | a^{i''}, a^{i'}) \equiv E_y[R(\mathbf{y}) | a^{i''}, a^j, a^{-i-j}] - E_y[R(\mathbf{y}) | a^{i'}, a^j, a^{-i-j}].$$

Since A and I_r are finite, there exists some scalar $\alpha > 0$ such that

$$|\Delta_i(z | z'', z') - \Delta_j(z | z'', z')| \leq \alpha$$

for all $z, z'', z \in A$ and all $i, j \in I_r$.

Then we make the following assumption:

Assumption 4. *For any two risk averse agents $i, j \in I_r$, $i \neq j$, and α defined above, (i) there exists some $\tilde{\beta} > 0$ such that*

$$\Delta_i(z''|z'', z') - \Delta_i(z'|z'', z') \geq \tilde{\beta}$$

for all $z'' > z'$ and (ii) $\beta \equiv \tilde{\beta} - \alpha > 0$

The first part of Assumption 4 says that the marginal expected revenue of organization with respect to action of each risk averse agent is increasing in actions of others. Put differently, actions of agents are complement (supermodular) with each other. The second part of Assumption 4 then means that such complementarity gain $\tilde{\beta}$ outweighs the asymmetry α among risk averse agents.

We also weaken Assumption 3 as follows:

Assumption 5. *For any two risk averse agents $i, j \in I_r$, if $i > j$, then*

$$\Delta G_i(\hat{a}, a) > \Delta G_j(\hat{a}, a)$$

for all $\hat{a} > a$.

Assumption 5 simply states that risk averse agents are ranked in decreasing order in terms of only their marginal action costs.

Since A and I are finite, there exists some $(\varepsilon, \delta) > 0$ such that

$$\delta \geq \left| \Delta G_j(a'', a') - \Delta G_i(a'', a') \right| \tag{6}$$

for all $i, j \in I_r$ and all $a'', a' \in A$, and

$$\varepsilon \geq \left| P_n^i(a) - P_n^j(a) \right| \tag{7}$$

for all $i, j \in I_r$, all $y_n \in Y$ and all $a \in A$.

Then we can show the following result:

Theorem 2. *Suppose that Assumption 1, 4 and 5 hold. Then there exists some non–empty set $D \subset \Re^2_+$ such that for all $(\varepsilon, \delta) \in D$ any fair contract $\hat{C}$ must have fixed wages for all risk averse agents who are induced to choose the least costly actions $a^i = 0$ for all $i \in I_r$.*

Proof. Appendix.

Note that Theorem 2 states that *all* risk averse agents are offered fixed wages in any fair contract when heterogeneity among them are sufficiently small in the sense that ε and δ defined in (6) and (7) are both small enough. Thus we can dispense with Assumption 2 and 3 and $M+1 \geq N_r$ which were used for proving Theorem 1.

The reason for this is as follows.

First, when heterogeneity among risk averse agents is small, any fair contract must have the feature that the wage schemes of risk averse agents which implement actions $(\hat{a}^i)_{i\in I_r}$ become close to each other. More formally, recall Lemma 3 that the fair contract $\hat{w}^i(y)$ for risk averse agent i must solve Problem $(\mathrm{CM}^i - \hat{a}^i)$: let $\hat{w}^i(y^i|\hat{a}^i)$ be the optimal wage scheme which implements action $\hat{a}^i$ from risk averse agent i in Problem $(\mathrm{CM}^i - \hat{a}^i)$. Here we must have $\hat{w}^i(y) = \hat{w}^i(y|\hat{a}^i)$. Then we show that $\hat{w}^i(y|\hat{a})$ becomes close to $\hat{w}^j(y|\hat{a})$ for any $y \in Y$ when their heterogeneity is small enough (both ε and δ are small), provided agent i and j choose the same action $\hat{a} \in A$.

Second, we show that all risk averse agents are actually induced to choose the same action at any fair contract if the heterogeneity is small enough. This follows from (IE*) and Assumption 4. By making use of the complementarity condition in Assumption 4, we can obtain the result that the actions of risk averse agents which satisfy (IE*) must be symmetric. To see this, let the organizational residual surplus define by the expected total revenue $E_y[R(\mathbf{y})|\mathbf{a}]$ minus the total wages of all risk averse agents $\sum_{i\in I_r} E_{y^i}[\hat{w}^i(y^i)|a^i]$:

$$H(\mathbf{a}) \equiv E_y[R(\mathbf{y})|\mathbf{a}] - \sum_{i\in I_r} E_{y^i}[\hat{w}^i(y^i)|\hat{a}^i].$$

This should be shared among all risk neutral agents due to (BB). Then (IE*) requires that the actions of risk averse agents $(a^i)_{i\in I_r}$ must maximize $H(\mathbf{a})$ because otherwise other actions can be implemented from risk averse agents and hence risk neutral agents can be better off, which contradicts (IE*). We then show that, if Assumption 4 holds, such optimal action becomes symmetric among all risk averse agents in the sense that $a^i = a^j$ for all $i, j \in I_r$.[7] Thus we can ensure that all risk averse agents choose the same action $\hat{a}^i = \hat{a}^j$ in any fair contract when the heterogeneity among them is sufficiently small, given Assumption 4.

Third, we can show that Assumption 3 is satisfied when risk averse agents are heterogeneous but such heterogeneity is larger for the differences in their marginal action costs than the differences in the impacts of their actions on the probability distributions of performances. In other words both ε and δ are small enough but δ is larger than ε.

By combining all these facts, we can ensure Assumption 3 as well as the condition that all risk averse agents choose the same action $\hat{a}^i = \hat{a}^j$ for all $i, j \in I_r$ when the heterogeneity among risk averse agents lies in some small range. Then, by using

[7]Without Assumption 4, the optimal actions which maximize H over $(a^i)_{i\in I_r}$ may be asymmetric even when the expected revenue function $E_y[R(\mathbf{y})|\mathbf{a}]$ is symmetric with respect to $(a^i)_{i\in I_r}$ and $\hat{w}^i(\cdot|\cdot)$ is also symmetric, i.e. $\hat{w}^i(\cdot|\cdot) = \hat{w}(\cdot|\cdot)$ for all $i \in I_r$. This is essentially because H may have multiple maximizes over $(a^i)_{i\in I_r}$ some of which are asymmetric.

Theorem 1, we can show that the fair contract must offer a fixed wage to each risk averse agent.

Next we characterize the optimal contracts offered to risk neutral agents. Note first that the wage scheme of each risk averse agent in any fair contract is individualistic by Lemma 2. Recall that $\hat{w}^i(y^i)$ is the wage scheme of risk averse agent i in a fair contract which implements action $\hat{a}^i$ from him. Due to (BB), the risk neutral agents then obtain the residual revenue $R(\mathbf{y})$ after subtracting total wages paid to all risk averse agents $\sum_{i \in I_r} \hat{w}^i(y^i)$.

Then we define the actions of risk neutral agents, denoted by $(\tilde{a}^i)_{i \in I_n}$, which maximize the expected residual surplus of organization after fixing the action profile $\mathbf{a}^r \equiv (\hat{a}^i)_{i \in I_r}$ of risk averse agents:

$$\hat{a}^i \in \arg\max_{a^i \in A} E_y[R(\mathbf{y})|\mathbf{a}^r, a^i, \hat{a}^{-i}] - \sum_{i \in I_n} G_i(a^i) - \sum_{i \in I_r} E_{y^i}[\hat{w}^i(y^i)|\hat{a}^i].$$

We can show the following result:

Theorem 3. *In any fair contract the risk neutral agents must choose the actions $(\hat{a}^i)_{i \in I_n}$ so as to maximize the residual surplus of organization, given action profile of risk averse agents $\mathbf{a}^r = (\hat{a}^i)_{i \in I_r}$.*

Proof. Appendix.

5 Existence of Fair Contract

In this section we show the conditions under which there exists a fair contract. Instead of showing $C_f^* \neq \emptyset$, we will return back to the original set of fair contracts C^f and proceed to show the stronger result that $C^f \neq \emptyset$. Due to Lemma 1, we know that $C^f \subseteq C_*^f$. Thus, if we can show that $C_f \neq \emptyset$, this implies $C_*^f \neq \emptyset$. Note here that any fair contract in C_f is defined in terms of incentive efficiency (IE) but not its modified version (IE*).

Our strategy for existence proof is as follows: first, we set the wage scheme of every risk averse agent as a fixed wage, i.e., $w^i(y^i) = \overline{w}$ for all $y^i \in Y$. Thus every risk averse agent chooses the least costly action, $a^i = 0$ for all $i \in I_r$.

We then show that offering a fixed wage to every risk averse agent becomes incentive efficient when we take the fixed wage $\overline{w}$ to be large enough. To see this, note that, if the fixed wage $\overline{w}$ is not incentive efficient, there must exist a profile of other contracts $(C^i)_{i \in I_r}$ for risk averse agents such that it solves the following problem:

Problem (OP):

$$\max_{(C^i)_{i \in I_r}} E[R(\mathbf{y})|\mathrm{a}^r, \mathrm{a}^n] - \sum_{i \in I_r} P_n^i(a^i) w_n^i$$

subject to

$$\sum_n P_n^i(a^i)u(w^i(y_n)) - G_i(a^i) \geq \sum_n P_n^i(a)u(w^i(y_n)) - G_i(a) \;\; \forall\, a \neq a^i, \;\; \forall\, i \in I_r \quad \text{(IC)}$$

$$\sum_n P_n^i(a^i)u(w^i(y_n)) - G_i(a^i) \geq u(\overline{w}) \;\; \forall\, i \in I_r \qquad \text{(IR}^*\text{)}$$

where $\mathbf{a}^r \equiv (a^i)_{i\in I_r}$ and $\mathbf{a}^n \equiv (a^i)_{i\in I_n}$ denote vectors of actions taken by risk averse and neutral agents respectively. Here (IC) denotes incentive compatibility constraint for risk averse agent i and (IR*) denotes the constraint such that risk averse agent i is not worse off compared to the case that he receives the fixed wage $\overline{w}$. Note that the expected wage paid to risk averse agent i, i.e., $\sum_n P_n^i(a^i)w_n^i$, is bounded below from $\phi(G_i(a^i) + u(\overline{w}))$ due to (IR*) and concavity of u. Then such lower bound for the expected wage of risk averse agent becomes large when $\overline{w}$ is taken to be large as well. This is more likely to be so when higher action a^i is implemented because then higher payment must be compensated to risk averse agent for agreeing on an alternative wage scheme. Thus risk neutral agents cannot be better off by offering other wage schemes than the fixed wage $\overline{w}$ to risk averse agent when the status quo utility $u(\overline{w})$ is large enough. Then the optimal solution to Problem (OP) involves the implementation of only the least costly action from risk averse agents.

Second, given this result, we find the wage schemes of risk neutral agents which are induced to choose the actions maximizing the residual surplus of organization after subtracting the fixed wages of risk averse agents. More formally, we define the optimal actions of risk neutral agents, denoted by $\tilde{\mathbf{a}}^n \in A^{N_n}$, which maximize the total residual of surplus given all risk averse agents choosing the least costly action $a^i = 0$, as follows:

$$\tilde{\mathbf{a}}^n \in \arg\max_{\mathbf{a}^n \in A^{N_n}} RS(\mathbf{a}^n) \equiv E_y[R(\mathbf{y})|\mathbf{0}, \mathbf{a}^n] - \sum_{i\in I_n} G_i(a^i) \tag{8}$$

where $\mathbf{0}$ denotes N_r–dimensional vector of zeros. Then we can show that there exists a single optimal action $\tilde{a} \equiv \tilde{a}^i$ for any risk neutral agent $i \in I_n$ when the heterogeneity among risk neutral agents is small enough. Given this, we will design the wage scheme which implements the optimal single action $\tilde{a}$ from every risk neutral agent.

The most difficult part of the proof is that we need to show that the wage schemes constructed in the above way must satisfy no–envyness condition (NE). To deal with this subtle issue, we will design the wage schemes of agents by exploiting the risk aversion of risk averse agents such that they would face risky lottery if they were offered the wage scheme of risk neutral agent, which prevents them from envying risk neutral agents. On the other hand, the wage scheme of risk neutral agent is designed such that his expected payoff is equal to the fixed wage offered to risk averse agent. Thus risk neutral agent does not envy risk averse agents.

Now we will proceed the formal argument for the proof of existence of fair contract.

First, we make the following assumptions:

Assumption 6. *For any action $a \in A$ and any risk neutral agents $i, j \in I_n$,*

$$\left| G_i(a) - G_j(a) \right| < \frac{1}{2}\beta$$

where $\beta > 0$ is given in Assumption 4.

Assumption 7. $u'(+\infty) = 0$.

Assumption 6 says that the difference between action costs of risk neutral agents are not so large relative to the complementarity gains of actions β defined in Assumption 4. Assumption 7 ensures that $\phi'(+\infty) = +\infty$.

Then we can show the following:

Lemma 4. *Suppose that Assumption 4 and 6 are satisfied. Then there exists a single optimal action $\tilde{a} \in A$ such that $\tilde{a}^i = \tilde{a}$ maximizes the residual surplus of organization $RS(\mathbf{a}^n)$ for all risk neutral agents $i \in I_n$.*

Proof. See Appendix.

For such optimal symmetric action $\tilde{a}$, we make the following condition:

Assumption 8. *For **some** risk neutral agent $l \in I_n$ there exists no $M+1$–dimensional vector $(\lambda_m) \geq \mathbf{0}$, not all zero, such that $\sum_m \lambda_m = 1$,*

$$\sum_m \lambda_m F_n^l(a_m) \leq F_n^l(\tilde{a})$$

and

$$\sum_m \lambda_m G_l(a_m) < G_l(\tilde{a}) + \delta$$

where $\delta > 0$ is given in (6).

Assumption 8 says that there exists some risk neutral agent, say l, who cannot improve the probability distribution of his performances $F_n^l(a)$ in the sense of first order stochastic dominance by using a mixes strategy λ_m over $(a_m)_{m=1}^{M+1} \neq \tilde{a}$ at lower expected action cost than the action $\tilde{a}$. This condition is more likely to be satisfied when heterogeneity among risk averse agents δ is small enough, as long as $F_n^l(a)$ satisfies Convexity of Distribution Function Condition (CDFC) and G_l is convex.[8]

[8]CDFC means that $F_n^i(a)$ is convex function of a. Suppose that $F_n^l(a)$ satisfies CDFC and G_l

Assumption 8 then ensures that there exists a non-decreasing wage scheme which implements the optimal action $\tilde{a}$ from every risk neutral agent. Thus (MON) is satisfied.

Finally, we make the following assumption:

Assumption 9. $E_y[R(\mathbf{y})|\mathbf{a}] = \Phi(\mathbf{a}) + F$

Assumption 9 says that the expected revenue of organization can be decomposed into two parts: one is the part which varies with unobservable actions of agents and the other is the part which is independent of these actions. Here F can be interpreted as the base revenue which is independent of non–contractible action choices.[9]

Then we can show the following result:

Theorem 4. *Suppose that Assumption 1 and 6-9 are satisfied. Then, if the base revenue of organization F is sufficiently large, there exists a fair contract $\hat{C}$ which has the following features:*

- *the risk averse agents are offered the fixed wage scheme which induces them to choose the least costly action $a^i = 0$, and*
- *the risk neutral agents are offered the wage scheme which induce them to choose the optimal action $\tilde{a}$ each so as to maximize the residual surplus of organization*

$$E[R(\mathbf{y})|\mathbf{0}, \mathbf{a}^n] - \sum_{j \in I_n} G_j(a^j).$$

given all risk averse agents choose the least costly action $\mathbf{a}^r = \mathbf{0}$.

Proof. See Appendix.

The requirement that the base revenue F is large can ensure that the fixed wage offered to each risk averse agent can be large enough without affecting the incentives of risk neutral agents. Then, large enough fixed wage makes the payoff of risk averse agents large as well so that the designed contract profile becomes incentive efficient because it solves Problem (OP).

is convex. Then $F_n^l(\sum_n \lambda a(m)) < \sum_m \lambda_m F_n^l(a(m)) \leq F_n^l(\tilde{a})$ implies that $a' \equiv \sum_m \lambda_m a(m) > \tilde{a}$. However, then we have $\sum_m \lambda_m G_l(a(m)) > G_l(a') > G_l(\tilde{a})$ which implies that $\sum_m \lambda_m G_l(a(m)) > G_l(\tilde{a})$. The last inequality cannot hold when δ is small enough.

[9]This does not necessarily mean that F is independent of any kinds of agents' actions. It could depend on contractible actions taken by agents.

6 Discussion and Robustness

6.1 Efficiency First or Equity First?

When we have so far characterized fair contracts in Section 3 and 4, we have defined the set of fair contracts as C_*^f which are the contracts satisfying (IE*), (NE) and (IR). In this definition (IE*) was referred to as the set of incentive feasible (i.e., budget balanced and incentive compatible) and monotonic contracts which are not Pareto dominated by all other incentive feasible and monotonic contracts. Thus we have imposed Pareto efficiency only to the set of incentive feasible and monotonic contracts which are not necessarily no–envy. This definition of fair contracts can be thus interpreted as follows: as the first step, agents of organization seek the efficiency constrained by incentive compatibility and monotonicity by ruling out the consideration of no-envyness. Then, as the second step, they select among all these contracts the one satisfying no–envyness (and (IR)). In other words, efficiency is taken as the first priority to be satisfied in organization whereas equity concern is the objective to be accomplished after efficiency is attained. In this section, by taking into account the above interpretation, we will call fair contracts which we have so far defined *efficiency–first optimal contracts*. We will denote by C_E^* the set of efficiency–first optimal contracts. Of course, C_E^* coincides with the set of fair contracts $\hat{C}$ which we have characterized in Section 4.

However, in contrast to this case, one might think that equity can be considered as the first priority to be satisfied in organization whereas efficiency is the second objective. For example, suppose that agents of organization first identify the set of all no–envy contracts. Next, the agents find the Pareto efficient contracts among all possible no–envy, incentive feasible and monotonic contracts. We define by C_E^{**} the set of all the efficient contracts defined in this way and satisfying (IR). We will call this *equity–first optimal contracts*.

Then, we can show the following result:

Theorem 5. $C_E^* \subseteq C_E^{**}$.

Proof. Take any efficiency–first optimal contract $C \in C_E^*$. Then we will show that $C \in C_E^{**}$. To see this, suppose that $C \notin C_E^{**}$. Then, $C \in C_E^*$ implies that C satisfies all (IC), (NE), (IR), (BB) and (MON). Thus, the fact that $C \notin C_E^{**}$ implies that there must exist some contract $C' \notin C$ such that C' satisfies (IC) and (NE), and it Pareto dominates C. In particular C' satisfies (IC) but Pareto dominates C. However, since $C \in C_E^*$, C satisfies (IC), and there exist no other contracts which satisfy (IC), and Pareto dominate C. C' contradicts to this fact. Q.E.D.

Thus, by Theorem 5, we may have $C_E^{**} \setminus C_E^* \neq \emptyset$ so that there may exist some equity–first optimal contract which does not have the property shown in Theorem 1-2 even when heterogeneity among agents is so small. We will show this by the following example:

Example. Suppose that $I_r = \{1,2\}$ and $I_n = \{p\}$. Here agent p is the principal. Suppose also that $A = \{0,1\}$ and $Y = \{y_2, y_1\}$ where $y_2 > y_1$. We also simplify the model by assuming that $R(y^1, y^2, y^p) = R(y^1, y^2)$. Thus the principal's action does not affect the expected revenue of organization. We thus set $a^p = 0$. Without loss of generality we take the principal as the residual claimant who receives all the residual surplus of organization after subtracting the wages to risk averse agents. Thus (BB) is satisfied and we ignore this.

Now consider the following optimization problem which minimizes the expected wage of the principal when implementing actions $\mathbf{a} = (a^1, a^2)$ from risk averse agents subject to (IC), (MON) and (NE), given each risk averse agent is guaranteed a certain utility level $\hat{U}^i$:

$$\min \sum_{i=1,2} E_y[w^i(\mathbf{y})|\mathbf{a}]$$

subject to

$$E_y[u(w^i(\mathbf{y}))|\mathbf{a}] - G_i(a^i) \geq E_y[u_i(w^i(\mathbf{y}))|a, a^j] - G_i(a), \;\; \forall\, a \neq a^i \qquad \text{(IC)}$$

$$E_y[u(w^i(\mathbf{y}))|\mathbf{a}] - G_i(a^i) \geq \hat{U}^i \qquad \text{(PE)}$$

$$E_y[u(w^i(\mathbf{y}))|\mathbf{a}] - G_i(a^i) \geq E_y[u(w^j(\mathbf{y}))|a, a^j] - G_i(a), \;\; \forall\, a \qquad \text{(NE)}$$

$$w^i(y_2, y^j) \geq w^i(y_1, y^j) \;\; \forall\, i \neq j \qquad \text{(MON)}$$

Note that the contract which solves the above problem becomes an equity–first optimal contract because it is Pareto efficient among all feasible contracts satisfying (IC), (NE) and (MON).

Let $EW^{**}(a^1, a^2)$ be the expected total wage in the solution to the above problem. Then we define the principal's expected payoff as follows:

$$\Pi^p(a^1, a^2) \equiv E_y[R(y^1, y^2)|a^1, a^2] - EW^{**}(a^1, a^2).$$

Then, we can verify that implementing high actions $(a^1, a^2) = (1,1)$ from both risk averse agents maximize the principal's expected payoff $\Pi^p(a^1, a^2)$ when $\Delta(1,1) \equiv E_y[R(\mathbf{y})|1,1] - \max_{\mathbf{a} \neq 1} E_y[R(\mathbf{y})|\mathbf{a}]$ is sufficiently large. This shows that the contract which solves the above problem gives risk averse agents high–powered incentive to choose high action instead of low action. Thus there may exist some equity–first optimal contract having the feature that high–powered incentive schemes are offered to risk averse agents when $\Delta(1,1)$ is sufficiently large.

The above example is in contrast to Theorem 1–2 that all efficiency–first optimal contracts must have fixed wage schemes offered to risk averse agents. This gives an interesting testable implication as follows: rewards of agents are more sensitive to their objective performances in organizations in which equity among them is taken as the first priority to be satisfied than in organizations in which efficiency is the first priority to be satisfied.

6.2 Robustness of the Results

We have so far paid our attention to the class of monotonic wage schemes (MON). Although (MON) is a reasonable and realistic restriction, one might think that, once we remove (MON) from the consideration of possible axioms, the set of contracts is expanded so that all axioms may be satisfied at non-fixed wages for risk averse agents.

In this subsection we will discuss the robustness of our results when we remove the axiom of monotonicity (MON).

First, we can show that our characterization results, Theorem 1 and Theorem 2, still remain true even if we drop (MON) from the axioms to be considered when we impose the monotone likelihood ratio property (MLRP) and convexity of distribution function condition (CDFC) on the probability distribution of each risk averse agent's performances $P_n^i(a)$. MLRP means that the likelihood ratio $P_n^i(a')/P_n^i(a)$ of agent i's performance is increasing in y_n conditional on $a' > a$. As we have already mentioned, CDFC means that the distribution function $F_n^i(a) \equiv \sum_{l \leq n} P_n^i(a)$ is convex with respect to action levels $a \in A$. Then, we modify Problem $(\mathrm{CM}^j - \hat{a}^j)$ by dropping (MON) from the constraint set. Then, by a similar argument to Lemma 3, we can show that any fair contract $\hat{C}^j$ of risk averse agent j must solve this modified problem. Otherwise, it is not incentive efficient. Then we can easily verify that the optimal wage scheme which solves this modified problem becomes monotone increasing by the standard argument when MLRP and CDFC are satisfied (see Grossman and Hart (1983)). Thus, as long as MLRP and CDFC are met, we can reach the same characterization results as Theorem 1-3.

Second, we will investigate the implications when we drop (MON) from the axioms and still do not impose the restrictions on the probability distributions such as MLRP and CDFC. We will then show that, even when we drop (MON), any fair contract still has the fixed wage offered to every risk averse agent, provided the heterogeneity among them is sufficiently small and the wage of each agent is bounded below from some amount. We impose the limited liability (LL) instead of (MON): for some $\underline{w} > -\infty$,

$$w^i(\mathbf{y}) \geq \underline{w}, \;\; \forall\, i \in I, \; \forall\, \mathbf{y} \in Y^N \tag{LL}$$

Then we call a contract profile C incentive feasible if it satisfies (BB), (IC) and (LL). Let $\mathcal{C}^{\mathcal{L}}$ be the set of all incentive feasible contracts. We focus on this set $\mathcal{C}^{\mathcal{L}}$. Similarly, we define the incentive efficient contract as follows: with a slight abuse of definition, a contract $C \in \mathcal{C}^{\mathcal{L}}$ is said to be incentive efficient, denoted by (IE**), if no other contract $C \in \mathcal{C}^{\mathcal{L}}$ exists such that it Pareto improve the expected utilities of agents. A fair contract C is defined analogously: C is said to be fair if it satisfies (IE**), (NE) and (IR).

Then we can show the following result:

Theorem 5. *Suppose that Assumption 1, 4 and 5 hold. Suppose also that (MON)* *is replaced by (LL). Then there exists some non-empty set $D \subset \Re^2_+$ such that for all $(\varepsilon, \delta) \in D$ any fair contract $\hat{C}$ involves a fixed wage offered to all risk averse agents.*

Proof. See Appendix.

By (LL), we can ensure that the utility payment for every risk averse agent $u(w(y_n))$ is bounded below from $\underline{u} \equiv u(\underline{w})$. Also (BB) requires that the utility payment $u(w(y_n))$ is bounded above by $u(\overline{R} - (N-1)\underline{w})$ where $\overline{R} \equiv \max_{\mathbf{y}} R(\mathbf{y})$. Thus the utility payment $u_n \equiv u(w(y_n))$ is bounded for each $y_n \in Y$. Then we can show that, when heterogeneity among risk averse agents is small enough so that ε is sufficiently small relative to δ, Assumption 3 is satisfied and hence we can reach the same characterization result as Theorem 1 and 2.

7 Concluding Remarks

In this paper we have characterized fair contracts which are incentive efficient and no–envy in general moral hazard environments. Then we have shown that fair contracts must have the feature that risk averse agents are offered the fixed wage scheme. On the other hand, risk neutral agents are motivated to work in order to maximize the residual surplus of organization after subtracting the fixed wages of risk averse agents. Then the tension between incentive efficiency and equity as no–envy makes characterization of contracts drastically simple. We have also shown the conditions under which a fair contract exists.

Our concept of no–envyness is based on ex ante view before agents choose their actions and hence final outcomes are realized. One might think that no–envy constraint is binding even ex post after the final performances are realized. In such case ex post no–envy constraint becomes more stringent than the ex ante one: suppose that the performances of agents $\mathbf{y} \in Y^N$ are realized. Then agent i does not envy agent j ex post if and only if $u(w^i(y^i)) \geq u(w^j(y^i))$, *conditional on his own performance* y^i. Then, we must have $w^i(y) = w^j(y)$ for any $y \in Y$ and all $i \neq j$. Hence we can reach the same conclusions as Theorem 1 and 2.

8 Appendix

8.1 Proof of Lemma 2

Suppose that some fair contract $\hat{C}^i$ is not individualistic for some risk averse agent $k \in I_r$, i.e., $\hat{w}^k(\mathbf{y})$ depends on y^{-k}. Let $\hat{u}^k(\mathbf{y}) \equiv u^k(\hat{w}^k(\mathbf{y}))$. Then we define the new contract $\tilde{u}^k(y^k)$ for that agent k as

$$\tilde{u}^k(y^k) \equiv E_{y^{-k}}[\hat{u}^k(y^k, y^{-k})|\hat{a}^{-k}]$$

and consider an alternative contract $C \neq \hat{C}$ as follows:

- For the risk averse agent k, the new contract $C = \{\tilde{u}^k, \hat{a}^k\}$ is offered.
- Pick one risk neutral agent, say agent n. Then, for all other agents $i \neq n, k$, we define a new wage scheme as

$$\tilde{w}^i(y^i) \equiv E_{y^{-i}}[\hat{w}^i(y^i, y^{-i})|\hat{a}^{-i}] \ \ \forall \ \ y^i \in Y.$$

- For the risk neutral agent n, the following wage scheme $\tilde{w}^n(\mathbf{y})$ is offered

$$\tilde{w}^n(\mathbf{y}) \equiv R(\mathbf{y}) - \sum_{i \in I, i \neq k} \tilde{w}^i(\mathbf{y}) - \tilde{w}^k(y^k)$$

where $\tilde{w}^k(y^k) \equiv \phi(\tilde{u}^k(y^k))$.

First, note that such alternative wage scheme $(\tilde{u}_n^k)_{y^n \in Y}$ offered to risk averse agent k satisfies (MON) because the original scheme $(u_n^k)_{y^n \in Y}$ satisfies (MON) and y^i and y^{-i} are statistically independent. By the same reason, the new scheme $\tilde{w}^i(y^i)$ offered to agent $i \neq k, n$ is also non-decreasing and hence (MON) is satisfied. Since $\tilde{w}^i(y^i)$, $i \neq k, n$, is independent of the performance of the selected risk neutral agent n, $\tilde{w}^n(\mathbf{y})$ is also non-decreasing in y^n because R is non-decreasing in y^n. Thus (MON) is satisfied for the risk neutral agent n as well.

Second, since (IC) is satisfied at $\hat{C}^k$, we have

$$E_y[\hat{u}^k(\mathbf{y})|\hat{a}^k, \hat{a}^{-k}] - G_k(\hat{a}^k) \geq E_y[\hat{u}^k(\mathbf{y})|a^k, \hat{a}^{-k}] - G_k(a^k), \ \ \forall \, a_k \neq \hat{a}_k,$$

which, by independence of y^i and y^{-i}, implies that

$$E_{y^k}[\tilde{u}^k(y^k)|\hat{a}^k] - G_k(\hat{a}^k) \geq E_{y^k}[\tilde{u}^k(y^k)|a^k] - G_k(a^k), \quad \forall \, a^k \neq \hat{a}^k.$$

Thus the new contract $\tilde{u}^k$ satisfies (IC) for the risk averse agent k who chooses the same action $\hat{a}^k$, given $\hat{a}^{-k}$. All other agents than k and n also choose the same actions $\hat{a}^{-k-n}$ as well. It is also true that all agents except n will obtain the same expected payoffs under the new contract as under the original one.

Now consider the incentive of the risk neutral agent n. Agent n would receive the following expected payoff under the new contract $\tilde{w}^n(\mathbf{y})$:

$$\begin{aligned}
\max_{a_n \in A} E_y[\tilde{w}^n(\mathbf{y})|a^n, \hat{a}^{-n}] - G_n(a^n) \ &\geq \ E_y[\tilde{w}^n(\mathbf{y})|\hat{a}^n, \hat{a}^{-n}] - G_n(\hat{a}^n) \\
&= \ E_y[R(\mathbf{y})|\hat{\mathbf{a}}] - \sum_{i \in I, i \neq k, n} E_y[\hat{w}^i(\mathbf{y})|\hat{a}^i] \\
&\quad -E_{y_k}[\tilde{w}^k(y_k)|\hat{a}_k] - G_n(\hat{a}_n) \\
&> \ E_y[R(\mathbf{y})|\hat{\mathbf{a}}] - \sum_{i \in I, i \neq k, n} E_y[\hat{w}^i(\mathbf{y})|\hat{\mathbf{a}}] - E_y[\hat{w}^k(y)|\hat{\mathbf{a}}] - G_n(\hat{a}^n). \\
&= \ E_y[\hat{w}^n(y)|\hat{\mathbf{a}}] - G_n(\hat{a}^n)
\end{aligned}$$

Here the last equality follows from (BB) and the strict inequality from the risk aversion of agent k (concavity of u), the Jensen's inequality and the dependence of $\hat{w}^k(y^k, y^{-k})$ on y^{-k}:

$$\begin{aligned}
&E_y[\hat{w}^k(y^k, y^{-k})|\hat{\mathbf{a}}] \\
&= E_y[\phi(\hat{u}^k(y^k, y^{-k}))|\hat{a}^k, \hat{a}^{-k}] \\
&= E_{y^k}[E_{y^{-k}}[\phi(\hat{u}^k(y^k, y^{-k}))|\hat{a}^{-k}]\hat{a}^k] \\
&> E_{y^k}[\phi(E_{y^{-k}}[\hat{u}^k(y^k, y^{-k})|\hat{a}^{-k}])|\hat{a}^k] \\
&= E_{y^k}[\phi(\tilde{u}^k(y^k))|\hat{a}^k] \\
&= E_{y^k}[\tilde{w}^k(y^k)|\hat{a}^k]
\end{aligned}$$

respectively. Thus the risk neutral agent n can be strictly better off under the new contract $\tilde{w}^n(\mathbf{y})$, which contradicts the fact that $\hat{C}$ is incentive efficient. Q.E.D.

8.2 Proof of Lemma 3

Consider Problem (CM$^j - \hat{a}^j$) for implementing the action $a = \hat{a}^j$ specified in the fair contract $\hat{C}^j$ from risk averse agent j. As we have argued in the main text, the constraint set of such problem is non-empty, compact and convex. Thus the optimal wage scheme which solves Problem (CM$^j - \hat{a}^j$) exists and is unique. We denote by $\hat{w}^j(y|\hat{a}^j)$ the optimal wage scheme which solves Problem (CM$^j - \hat{a}^j$).

Suppose that for some risk averse agent $j \in I_r$ the the wage scheme $\hat{w}^j(y)$ specified in the fair contract $\hat{C}^j = \{\hat{w}^j(y), \hat{a}^j\}$ differs from the optimal wage scheme $\hat{w}^j(y|\hat{a}^j)$: $\hat{w}^j(y) \neq \hat{w}^j(y|\hat{a}^j)$ for some $y \in Y$. Then, note that $\hat{w}^j(y)$ satisfies (IC) and (IR*) in Problem (CM$^j - \hat{a}^j$) so that it is feasible in Problem (CM$^j - \hat{a}^j$) and that $\hat{w}^j(y|\hat{a}^j)$ is uniquely determined to solve the problem (CM$^j - \hat{a}^j$) given $\hat{U}^j$. Thus we must have

$$E_{y^j}[\hat{w}^j(y^j)|\hat{a}^j] \equiv \sum_n P_n^j(\hat{a}^j)\hat{w}^j(y_n) > E_{y^j}[\hat{w}^j(y^j|\hat{a}^j)|\hat{a}^j] \equiv \sum_n P_n^j(\hat{a}^j)\hat{w}^j(y_n|\hat{a}_j).$$

Then consider an alternative contract offered to the above risk averse agent j as $C^j = \{\hat{w}^j(y^j|\hat{a}^j), \hat{a}^j\}$. Without loss of generality we can also assume that the wage schemes of other risk neutral agents than n depend solely on their own performances because (y^i, y^{-i}) are statistically independent and hence their action incentives are not changed by removing the dependence of the others' performances on their wage schemes. Thus all others than n choose the same actions $\hat{a}^i$ as before.

But now consider an alternative contract offered to the selected risk neutral agent $n \in I_n$ as $C^n = \{\tilde{w}^n(\mathbf{y}), \tilde{a}^n\}$ where

$$\tilde{w}^n(\mathbf{y}) = R(\mathbf{y}) - \sum_{l \neq j} \hat{w}^l(y^l) - \hat{w}^j(y^j|\hat{a}^j)$$

and $\tilde{a}^n$ is defined as the action to maximize the expected utility of agent n:

$$\tilde{a}^n \in \arg\max_{a \in A} E_y[\tilde{w}^n(\mathbf{y})|a, \hat{a}^{-n}] - G_n(a).$$

Then such risk neutral agent n can be strictly better off because

$$\begin{aligned}
& E_y[\tilde{w}^n(\mathbf{y})|\tilde{a}^n, \hat{a}^{-n}] - G_n(\tilde{a}^n) \\
\geq \ & E_y[\tilde{w}^n(\mathbf{y})|\hat{a}^n, \hat{a}^{-n}] - G_n(\hat{a}^n) \\
= \ & E_y[R(\mathbf{y})|\hat{\mathbf{a}}] - \sum_{l \neq j} E_{y^l}[\hat{w}^l(y^l)|\hat{a}^l] - E_{y^j}[\hat{w}^j(y^j|\hat{a}^j)|\hat{a}^j] - G_n(\hat{a}^n) \\
> \ & E_y[R(\mathbf{y})|\hat{\mathbf{a}}] - \sum_{i \in I} E_y[\hat{w}^i(y^i)|\hat{a}^i] - G_n(\hat{a}^n) \\
= \ & E_y[\hat{w}^n(\mathbf{y})|\hat{a}^n] - G_n(\hat{a}^n)
\end{aligned}$$

where the first inequality follows from the definition of $\tilde{a}^n$ which maximizes the agent n's expected utility, the second inequality from $E_{y^j}[\hat{w}^j(y^j)|\hat{a}^j] > E_{y^j}[\hat{w}^j(y^j|\hat{a}^j)|\hat{a}^j]$ and the final equality from (BB) respectively. Q.E.D.

8.3 Proof of Theorem 1

Take any fair contract $\hat{C}$. Suppose contrary to the claim that there are less than $N_r - M$ risk averse agents who choose the least costly action (zero). Then, by our supposition that $\#A = M + 1 \leq \#I_r = N_r$, there must exist at least two risk averse agents, say i and j, who choose the same action, say $\hat{a} > 0$. Without loss of generality we suppose that $i > j$ in the sense of Assumption 3.

Let $\hat{u}_n^k \equiv u(\hat{w}^k(y_n))$, $k = i, j$, for each $y_n \in Y$. Then we obtain the following series of expressions:

$$
\begin{aligned}
&\sum_n P^j_n(\hat{a})\hat{u}^j_n - G_j(\hat{a}) \\
&\geq \sum_n P^j_n(\hat{a})\hat{u}^i_n - G_j(\hat{a}) \\
&= \sum_n P^i_n(\hat{a})\hat{u}^i_n - G_i(\hat{a}) + \left\{ G_i(\hat{a}) - G_j(\hat{a}) + \sum_n (P^j_n(\hat{a}) - P^i_n(\hat{a}))\hat{u}^i_n \right\} \\
&\geq \sum_n P^i_n(a)\hat{u}^j_n - G_i(a) + \left\{ G_i(\hat{a}) - G_j(\hat{a}) + \sum_n (P^j_n(\hat{a}) - P^i_n(\hat{a}))\hat{u}^i_n \right\} \\
&= \sum_n P^j_n(a)\hat{u}^j_n - G_j(a) \\
&+ \left\{ G_j(a) - G_i(a) + \sum_n (P^i_n(a) - P^j_n(a))\hat{u}^i_n \right\} + \left\{ G_i(\hat{a}) - G_j(\hat{a}) + \sum_n (P^j_n(\hat{a}) - P^i_n(\hat{a}))\hat{u}^j_n \right\} \\
&= \sum_n P^j_n(a)\hat{u}^j_n - G_j(a) \\
&+ \left\{ \Delta G_i(\hat{a}, a) - \Delta G_j(\hat{a}, a) + \sum_n (P^j_n(\hat{a}) - P^i_n(\hat{a}))\hat{u}^i_n - \sum_n (P^j_n(a) - P^i_n(a))\hat{u}^j_n \right\}
\end{aligned}
$$

for any $a < \hat{a}$. Here the first inequality follows from (NE) for agent j and the second inequality from (NE) of agent i respectively. Thus we can reach the following inequality:

$$\sum_n P^j_n(\hat{a})\hat{u}^j_n - G_j(\hat{a}) \geq \sum_n P^j_n(a)\hat{u}^j_n - G_j(a) + W(\hat{a}, a|\hat{\mathbf{u}}^i, \hat{\mathbf{u}}^j) \tag{A1}$$

where $\hat{\mathbf{u}}^k \equiv (\hat{u}^k_n)^L_{n=1}$ denotes the vector of the utility payments for $k = i, j$ and

$$W(\hat{a}, a|\hat{\mathbf{u}}^i, \hat{\mathbf{u}}^j) \equiv \left\{ \Delta G_i(\hat{a}, a) - \Delta G_j(\hat{a}, a) + \sum_n (P^j_n(\hat{a}) - P^i_n(\hat{a}))\hat{u}^i_n - \sum_n (P^j_n(a) - P^i_n(a))\hat{u}^j_n \right\}.$$

Since the above inequality (A1) holds for any action $a < \hat{a}$, we take an action $a < \hat{a}$ such that (IC) of agent j is binding:

$$\sum_n P^j_n(\hat{a})\hat{u}^j_n - G_j(\hat{a}) = \sum_n P^j_n(a)\hat{u}^j_n - G_j(a).$$

Such action a must exist: otherwise the action $\hat{a}$ must be the smallest among all relevant actions when solving Problem ($\text{CM}^j - \hat{a}$) because (IC) of agent j in that problem is not binding at any $a < \hat{a}$. Thus the optimal wage which solves Problem ($\text{CM}^j - \hat{a}$), denoted $\hat{w}^j(y|\hat{a})$, must be the fixed wage but then this does not satisfy (IC). Thus (IC) of agent j must be binding at some $a < \hat{a}$.

Now we show that $W(\hat{a}, a|\hat{\mathbf{u}}^i, \hat{\mathbf{u}}^j) > 0$ for such action $a < \hat{a}$ that (IC) of agent j is binding. To see this, note that $\hat{\mathbf{u}}^i = (\hat{u}_n^i)_{n=1}^L$ must satisfy (IC) of agent i:

$$\sum_n P_n^i(\hat{a})\hat{u}_n^i - G_i(\hat{a}) \geq \sum_n P_n^i(a)\hat{u}_n^i - G_i(a)$$

which can be written by

$$\sum_n (F_n^i(a) - F_n^i(\hat{a}))\Delta u_n^i \geq \Delta G_i(\hat{a}, a)$$

where $\Delta u_n^i \equiv \hat{u}_n^i - u_{n-1}^i$ for each $i = 2, ..., L$. By (MON) $\Delta u_n^i \geq 0$ must hold for each $i = 2, ..., L$.

Then, since $F_n^i(a) > F_n^i(\hat{a})$ for each n when $\hat{a} > a$, we can show that

$$\begin{aligned}
\sum_n (P_n^j(\hat{a}) - P_n^i(\hat{a}))\hat{u}_n^i &= \sum_n (F_n^i(\hat{a}) - F_n^j(\hat{a}))\Delta u_n^i \\
&= \sum_n \left(\frac{F_n^i(\hat{a}) - F_n^j(\hat{a})}{F_n^i(a) - F_n^i(\hat{a})}\right)(F_n^i(a) - F_n^i(\hat{a}))\Delta u_n^i \\
&= \sum_n \left(\frac{F_n^i(a) - F_n^j(\hat{a})}{F_n^i(a) - F_n^i(\hat{a})} - 1\right)(F_n^i(a) - F_n^i(\hat{a}))\Delta u_n^i \\
&= \sum_n \lambda_n^i(\hat{a}, a|i, j)(F_n^i(a) - F_n^i(\hat{a}))\Delta u_n^i \\
&\geq (\underline{\lambda}^i(\hat{a}, a|i, j) - 1)\sum_n (F_n^i(a) - F_n^i(\hat{a}))\Delta u_n^i \\
&\geq (\underline{\lambda}^i(\hat{a}, a|i, j) - 1)\Delta G_i(\hat{a}, a)
\end{aligned}$$

where the last inequality follows from (IC) of agent i and $\underline{\lambda}^i > 1$ (due to $F_n^j(\hat{a}) < F_n^i(a)$ for $i > j$ by Assumption 2 (ii)).

Next, the binding (IC) of agent j can be written by

$$\sum_n (F_n^j(a) - F_n^j(\hat{a}))\Delta u_n^j = \Delta G_j(\hat{a}, a).$$

By using this and $F_n^j(a) > F_n^j(\hat{a})$ for $\hat{a} > a$, we can show that

$$\begin{aligned}
\sum_n (P_n^j(a) - P_n^i(a))\hat{u}_n^j &= \sum_n (F_n^i(a) - F_n^j(a))\Delta u_n^j \\
&= \sum_n \left(\frac{F_n^i(a) - F_n^j(a)}{F_n^j(a) - F_n^j(\hat{a})}\right)(F_n^j(a) - F_n^j(\hat{a}))\Delta u_n^j \\
&= \sum_n \left(\frac{F_n^i(a) - F_n^j(\hat{a})}{F_n^j(a) - F_n^j(\hat{a})} - 1\right)(F_n^j(a) - F_n^j(\hat{a}))\Delta u_n^j \\
&= \sum_n \lambda_n^j(\hat{a}, a|i,j)(F_n^j(a) - F_n^j(\hat{a}))\Delta u_n^j \\
&\leq (\overline{\lambda}^j(\hat{a}, a|i,j) - 1)\sum_n (F_n^j(a) - F_n^j(\hat{a}))\Delta u_n^j \\
&= (\overline{\lambda}^j(\hat{a}, a|i,j) - 1)\Delta G_j(\hat{a}, a)
\end{aligned}$$

By combining the above inequalities, we obtain

$$\begin{aligned}
&\sum_n (P_n^j(\hat{a}) - P_n^i(\hat{a}))\hat{u}_n^i - \sum_n (P_nj(a) - P_n^i(a))\hat{u}_n^j \\
&= \sum_n (F_n^i(\hat{a}) - F_n^j(\hat{a}))\Delta u_n^i - \sum_n (F_n^i(a) - F_n^j(a))\Delta u_n^j \\
&\geq (\underline{\lambda}^i(\hat{a}, a|i,j) - 1)\Delta G_i(\hat{a}, a) - (\overline{\lambda}^j(\hat{a}, a|i,j) - 1)\Delta G_j(\hat{a}, a) \\
&> -\Delta G_i(\hat{a}, a) + \Delta G_j(\hat{a}, a)
\end{aligned}$$

due to Assumption 3. Thus we have established that $W(\hat{a}, a|\hat{\mathbf{u}}^i, \hat{\mathbf{u}}^j) > 0$ for the action $a < \hat{a}$ such that (IC) of agent j is binding.

Then inequality (A1) implies that

$$\sum_n P_n^j(\hat{a})\hat{u}_n^j - G_j(\hat{a}) > \sum_n P_n^j(a)\hat{u}_n^j - G_j(a)$$

which contradicts the supposition that (IC) of agent j is binding at $a < \hat{a}$.

Thus our first supposition that there are less than $N_r - M$ risk averse agents who choose the least costly action is false. This completes our proof. Q.E.D.

8.4 Proof of Theorem 2

First, note that Assumption 3 is satisfied when ε is small enough because then $\overline{\lambda}^k(\hat{a}, a) \to 1$ and $\underline{\lambda}^k(\hat{a}, a) \to 1$ for all $i, j \in I_r$ and all $\hat{a}, a \in A$. Then Assumption 5 implies Assumption 3 when we take ε to be small enough given δ.

By Lemma 3, the wage scheme $\hat{w}^j(y^j)$ specified in fair contract $\hat{C}^j$ must be equal to $\hat{w}^j(y^j|\hat{a}^j)$ for any risk averse agent $j \in I_r$ which solves problem $(\mathrm{CM}^j - \hat{a}^j)$. Now replace agent j by agent i in Problem ($\mathrm{CM}^j - \hat{a}^j$), denoted by Problem ($\mathrm{CM}^i - \hat{a}^j$) which minimizes the expected wage paid to risk averse agent i, i.e. $\sum_n P_n^i(\hat{a}^j)w(y_n)$, subject to (IC) and (IR*) for agent i (but not agent j) who is induced to choose the action $\hat{a}^j$ specified in the fair contract offered to agent j (but not agent i). In general, such problem may have no optimal solutions because the constraint set may be empty as opposed to Problem ($\mathrm{CM}^j - \hat{a}^j$). However, we can show that when heterogeneity among risk averse agents is sufficiently small ((ε, δ) is small enough) the optimal wage scheme which solves the problem ($\mathrm{CM}^i - \hat{a}^j$) exists and is very close to $\hat{w}^j(y|\hat{a}^j)$ the optimal wage scheme which solves Problem ($\mathrm{CM}^j - \hat{a}^j$).

Lemma A1. *There exist some $\delta > 0$ and $\varepsilon > 0$ such that*

$$\left| \sum_n P_n^j(\hat{a}^j)\hat{w}^j(y_n|\hat{a}^j) - \sum_n P_n^i(\hat{a}^j)\hat{w}^i(y_n|\hat{a}^j) \right| < \frac{1}{2}\beta$$

for any risk averse agents $i, j \in I_r$, $i \neq j$, where β is given in Assumption 4.

Proof. Take any fair contract $\hat{C}$ and the corresponding expected utility of agent i as $\hat{U}^i$.

Then we first show the following:

Claim A1-1. *$|\hat{U}^i - \hat{U}^j|$ goes to zero as the difference between agents i and j (captured by δ and ε) becomes very small.*

Proof. Let $\hat{w}_n^k \equiv \hat{w}^k(y_n)$ and $\Delta u_n^i \equiv \hat{u}_n^k - \hat{u}_{n-1}^k$ for each $y_n \in Y$.

Then we derive

$$\begin{aligned}
\hat{U}^j &\equiv \sum_n P_n^j(\hat{a}^j)u(\hat{w}_n^j) - G_j(\hat{a}^j) \\
&\geq \sum_n P_n^j(\hat{a}^i)u(\hat{w}_n^i) - G_j(\hat{a}^i) \text{ (by (NE))} \\
&= \sum_n P_n^j(\hat{a}^i)u(\hat{w}_n^i) - G_i(\hat{a}^i) + \{G_i(\hat{a}^i) - G_j(\hat{a}^i)\} \\
&= \sum_n (1 - F_n^j(\hat{a}^i))\Delta u_n^i - G_i(\hat{a}^i) + \{G_i(\hat{a}^i) - G_j(\hat{a}^i)\} \\
&= \sum_n (1 - F_n^i(\hat{a}^i))\Delta u_n^i - G_i(\hat{a}^i) + \{G_i(\hat{a}^i) - G_j(\hat{a}^i)\} + \sum_n (F_n^i(\hat{a}^i) - F_n^j(\hat{a}^i))\Delta u_n^i \\
&= \hat{U}^i + \{G_i(\hat{a}^i) - G_j(\hat{a}^i)\} + \sum_n (F_n^i(\hat{a}^i) - F_n^j(\hat{a}^i))\Delta u_n^i \\
&\geq \hat{U}^i + \{G_i(\hat{a}^i) - G_j(\hat{a}^i)\} + (\underline{\lambda}^i(\hat{a}^i, a) - 1)\Delta G_i(\hat{a}^i, a) \\
&\geq \hat{U}^i - \tilde{\varepsilon} + \{G_i(\hat{a}^i) - G_j(\hat{a}^i)\}
\end{aligned}$$

where $\tilde{\varepsilon} > 0$ is defined as

$$\tilde{\varepsilon} \geq |(\underline{\lambda}^i(\hat{a}, a) - 1)\Delta G_i(\hat{a}, a)| \tag{A2}$$

for all $\hat{a}, a \in A$ and all $i \in I_r$. Note that $\tilde{\varepsilon}$ can be close to zero as $\varepsilon \to 0$ because then $\underline{\lambda}^i(\hat{a}, a) \to 1$.

Thus we obtain

$$\hat{U}^j \geq \hat{U}^i - \tilde{\varepsilon} + \{G_i(\hat{a}^i) - G_j(\hat{a}^i)\} \tag{A3}$$

Similarly, we have

$$\hat{U}^i \geq \hat{U}^j - \tilde{\varepsilon} + \{G_j(\hat{a}_j) - G_i(\hat{a}_j)\}. \tag{A4}$$

By adding (A3) and (A4), we derive

$$\begin{aligned}
&\{G_i(\hat{a}^j) - G_j(\hat{a}^j)\} + \tilde{\varepsilon} + \hat{U}^i \\
&\geq \hat{U}^j \\
&\geq \hat{U}^i - \tilde{\varepsilon} + \{G_i(\hat{a}^i) - G_j(\hat{a}^i)\} \\
&\geq \hat{U}^i - \tilde{\varepsilon} + \{G_i(\hat{a}^j) - G_j(\hat{a}^j)\} - \delta
\end{aligned}$$

from which $\hat{U}^i - \hat{U}^j$ goes to zero as δ and ε go to zero (hence $\tilde{\varepsilon} \to 0$). Also, when $\varepsilon = 0$ and $\delta = 0$, we have $\hat{U}^i = \hat{U}^j$. Thus there exists some small $\tilde{\rho} > 0$ such that $|\hat{U}^i - \hat{U}^j| \leq \tilde{\rho}$ for any $i, j \in I_r$, when $(\varepsilon, \delta) \to 0$. Q.E.D.

Now, we pick any risk averse agents $i, j \in I_r$. Then we consider the modified problem of Problem ($\text{CM}^j - a$) which implements $a \in A$ from risk averse agent j by keeping (MON) while replacing (IR) and (IC*) by

$$\sum_n P_n^j(a) u(w(y_n)) - G_j(a) \geq \hat{U}^j - \rho \tag{IR$^*_\rho$}$$

and

$$\sum_n P_n^j(a) u(w(y_n)) - G_j(a) \geq \sum_n P_n^j(a') u(w(y_n)) - G_j(a') - \rho, \;\; \forall\, a' \neq a \tag{IC$_\rho$}$$

respectively where $\rho > 0$ (we will give a precise definition of ρ later.)

We call such modified problem of ($\text{CM}^j - a$) Problem ($\text{CM}^j_\rho - a$). Then we can show the following result:

Claim A1-2. *The solution to Problem ($\text{CM}^j_\rho - a$), denoted $(\hat{w}^j_\rho(y_n|a))_n$, exists and is uniquely determined if the constraint set is non-empty.*

Proof. To save notation, we drop superscript j to index agent j in the following proof. Let $w(y_n) \equiv w_n$ for each $y_n \in Y$. We change the variable as $u_n \equiv u(w_n)$ and hence $w_n = \phi(u_n)$ where $\phi = u^{-1}$ is inverse of u. Then (MON) is replaced by $\phi(u_n) \geq \phi(u_{n-1})$ because of $y_n > y_{n-1}$.

Let $\Gamma(a)$ be the constraint the set of $\mathbf{u} \equiv (u_n)_n$ satisfying all (IC_ρ), (IR^*_ρ) and (MON) in Problem ($\text{CM}^j_\rho - a$). Then we show that $\Gamma(a)$ is a convex set. Take $\mathbf{u}, \mathbf{u}' \in \Gamma(a)$ and let $\mathbf{u}'' \equiv \lambda \mathbf{u} + (1-\lambda)\mathbf{u}'$ for a scalar $\lambda \in (0,1)$. Then we can see that $\mathbf{u}''$ satisfies (IC_ρ) and (IR^*_ρ) because these constraints are linear with $\mathbf{u}$. Thus it suffices to show that $\mathbf{u}''$ satisfies (MON). To see this, note that

$$\begin{aligned}\phi(u''_n) - \phi(u''_{n-1}) &\geq \phi'(u''_{n-1})(u''_n - u''_{n-1}) \\ &= \phi'(u''_{n-1})[\lambda(u_n - u_{n-1}) + (1-\lambda)(u'_n - u'_{n-1})] \\ &\geq 0\end{aligned}$$

due to $\phi' > 0$ and $u'_n \geq u'_{n-1}$ and $u_n \geq u_{n-1}$ by (MON). Thus $\mathbf{u}'' \in \Gamma(a)$ and hence $\Gamma(a)$ is a convex set.

Note that $\Gamma(a)$ is closed because ϕ is continuous. Also $\Gamma(a)$ can be bounded (see Grossman and Hart (1983)). Thus $\Gamma(a)$ is compact set. Since the objective function $\sum_n P^j_n(a)\phi(u_n)$ is continuous and convex with respect to u_n, a solution to Problem ($\mathrm{CM}^j_\rho - a$) exists and is unique if the constraint set $\Gamma(a)$ is non-empty. Q.E.D.

Thus, if the constraint set of Problem ($\mathrm{CM}^j_\rho - a$) is non-empty, by the Berge's Maximum Theorem, $\hat{w}^j_\rho(y|a)$ is continuous in ρ.

Recall that $\hat{w}^j(y|\hat{a}^j)$ is the optimal wage scheme which solves Problem ($\mathrm{CM}^j - \hat{a}^j$). We denote $\hat{u}^i_n \equiv u(\hat{w}^i(y_n))$ and $\hat{u}^i(y_n|a) \equiv u^i(\hat{w}^i(y_n|a))$. Then Lemma 3 shows that the fair contract $\hat{w}^j(y)$ must be equal to $\hat{w}^j(y|\hat{a}^j)$. Thus we have $\hat{u}^i_n \equiv \hat{u}^i(y_n|\hat{a}^i)$ for each $i \in I_r$.

Then we show the following claim.

Claim A1-3. *There exists some $K < +\infty$ such that $|\hat{U}^i| \leq K$ for all $i \in I_r$.*

Proof. By (IR), we have $\hat{U}^i \geq \overline{U}^i$. Thus $\hat{U}^i$ is bounded below.

Second, by using Problem ($\mathrm{CM}^i - \hat{a}^i$) and convexity of ϕ, we also obtain $\sum_n P^i_n(\hat{a}^i)\phi(\hat{u}^i_n) = \sum_n P^i_n(\hat{a}^i)(\hat{u}^i(y_n|\hat{a}^i)) \geq \phi\left(\sum_n P^i_n(\hat{a}^i)\hat{u}^i(y_n|\hat{a}^i)\right) \geq \phi(G_i(\hat{a}^i) + \hat{U}^i)$. Thus, if $\hat{U}^i \to +\infty$ for some $i \in I_r$, the expected wage of such agent i goes to infinity, i.e., $\sum_n P^i_n(\hat{a}^i)\phi(\hat{u}^i_n) \to +\infty$, so that the payoff of some risk neutral agent must be negatively infinite but this violates (IR) of that agent. Thus $\hat{U}^i$ must be bounded above as well. Q.E.D.

As we have already mentioned, $\hat{w}^i(y|\hat{a}^i)$ which solves Problem ($\mathrm{CM}^i - \hat{a}^i$) for implementing $\hat{a}^i$ exists and is unique. Thus $\hat{w}^i(y|\hat{a}^i)$ is well-defined. Then $\hat{w}^i(y|\hat{a}^i)$ is bounded for all $\hat{U}^i \in [\overline{U}^i, K]$, i.e., there exists some $M < +\infty$ such that $|u(\hat{w}^i(y|\hat{a}^i))| \leq M$ for all $y \in Y$ and all $\hat{U}^i \in [\overline{U}^i, K]$.

Then we can show the following:

Claim A1-4. *For all $a \in A$ and all $i \neq j$, $i, j \in I_r$, we have*

$$\left|\sum_n P^i_n(a)\hat{u}^i(y_n|\hat{a}^i) - G_i(a) - \left\{\sum_n P^j_n(a)\hat{u}^i(y_n|\hat{a}^i) - G_j(a)\right\}\right| \to 0$$

as $(\varepsilon, \delta) \to 0$

Proof. We obtain

$$\begin{aligned}
&\left|\sum_n (P_n^i(a) - P_n^j(a))\hat{u}^i(y_n|\hat{a}^i) + (G_j(a) - G_i(a))\right| \\
&\le \sum_n \left|P_n^i(a) - P_n^j(a)\right|\hat{u}^i(y_n|\hat{a}^i) + \left|G_j(a) - G_i(a)\right| \\
&\le \sum_n \left|P_n^i(a) - P_n^j(a)\right| M + \left|G_j(a) - G_i(a)\right|
\end{aligned}$$

which goes to zero as $(\varepsilon, \delta) \to 0$ for all $a \in A$ (note that A is a finite set). Q.E.D.

By Claim A1-4, we can find some small enough $\hat{\rho} > 0$ such that

$$\left|\sum_n P_n^i(a)\hat{u}^i(y_n|\hat{a}^i) - G_i(a) - \left\{\sum_n P_n^j(a)\hat{u}^i(y_n|\hat{a}^i) - G_j(a)\right\}\right| \le \hat{\rho} \tag{A5}$$

when (ε, δ) is small enough.

In what follows we define $\rho \equiv \max\{\tilde{\rho} + \hat{\rho}, 2\hat{\rho}\}$.

Now we consider Problem $(\text{CM}_\rho^j - \hat{a}^i)$ for the defined value $\rho > 0$ and then show that the wage scheme $\hat{w}^i(y|\hat{a}^i)$ is feasible in Problem $(\text{CM}_\rho^j - \hat{a}^i)$ for implementing $\hat{a}^i$, i.e., it satisfies (MON), (IR_ρ^*) and (IC_ρ). It is clear that (MON) is satisfied because $\hat{w}^i(y|\hat{a}^i)$ is monotonic by definition. To see that it also satisfies (IR_ρ^*) and (IC_ρ), note first that

$$\begin{aligned}
\sum_n P_n^j(\hat{a}^i)u(\hat{w}^i(y_n|\hat{a}^i)) - G_j(\hat{a}^i) &\ge \sum_n P_n^i(\hat{a}^i)u(\hat{w}^i(y_n|\hat{a}^i)) - G_i(\hat{a}^i) - \hat{\rho} \\
&\ge \hat{U}^i - \hat{\rho} \\
&\ge \hat{U}^j - (\tilde{\rho} + \hat{\rho}) \\
&\ge \hat{U}^j - \rho
\end{aligned}$$

where the first inequality follows from Claim 4-4, the second inequality from the fact that $\hat{w}^i(y|\hat{a}^i)$ satisfies (IR^*) of Problem $(\text{CM}^i - \hat{a}^i)$ to implement $\hat{a}^i$ from agent i, the third inequality from $|\hat{U}^j - \hat{U}^i| \le \tilde{\rho}$ and the last inequality from the definition of ρ respectively. This shows that $\hat{w}^i(y|\hat{a}^i)$ satisfies (IR_ρ^*) in Problem $(\text{CM}^j - \hat{a}^i)$ for implementation of $\hat{a}^i$.

Also we have

$$\begin{aligned}
\sum_n P_n^j(\hat{a}^i)u(\hat{w}^i(y_n|\hat{a}^i)) - G_j(\hat{a}^i) &\geq \sum_n P_n^i(a)u(\hat{w}^i(y_n|a)) - G_i(a) - \hat{\rho} \\
&\geq \sum_n P_n^i(a')u(\hat{w}^i(y_n|a)) - G_i(a') - \hat{\rho} \\
&\geq \sum_n P_n^j(a')u(\hat{w}^i(y_n|a)) - G_j(a') - 2\hat{\rho} \\
&\geq \sum_n P_n^j(a')u(\hat{w}^i(y_n|a)) - G_j(a') - \rho
\end{aligned}$$

where the first and third inequalities follow from the definition of $\hat{\rho}$ (see (A5)), the second inequality from the fact that $\hat{w}^i(y|\hat{a}^i)$ satisfies (IC) of Problem $(\text{CM}^i - \hat{a}^i)$ to implement $\hat{a}^i \in A$ from agent i and the last inequality from the definition of ρ respectively.

The above argument shows that $\hat{w}^i(y|\hat{a}^i)$ satisfies (IR^*_ρ), (IC^*_ρ) and (MON) of Problem $(\text{CM}^j_\rho - \hat{a}^i)$. Thus $\hat{w}^i(y|\hat{a}^i)$ is feasible in Problem $(\text{CM}^j_\rho - \hat{a}^i)$. This implies that the constraint set of Problem $(\text{CM}^j_\rho - \hat{a}^i)$ is non-empty for implementation of $\hat{a}^i$. Thus the optimal wage scheme $\hat{w}^j_\rho(y|\hat{a}^i)$ which solves such problem is well-defined. Also, by optimality of $\hat{w}^j_\rho(y|\hat{a}^i)$, we have

$$\sum_n P_n^j(\hat{a}^i)\hat{w}^i(y_n|\hat{a}^i) \geq \sum_n P_n^j(\hat{a}^i)\hat{w}^j_\rho(y_n|\hat{a}^i). \tag{A6}$$

By changing the role of i and j in the above argument, we can show that $\hat{w}^j(y|\hat{a}^j)$ is feasible in the problem $(\text{CM}^i_\rho - \hat{a}^j)$. Thus we have

$$\sum_n P_n^i(\hat{a}^j)\hat{w}^j(y_n|\hat{a}^j) \geq \sum_n P_n^i(\hat{a}^j)\hat{w}^i_\rho(y_n|\hat{a}^j). \tag{A7}$$

Then we show Lemma A1 as follows: first, note that $\hat{w}^j_\rho(y|\hat{a}^i)$ is continuous in ρ and that $\lim_{\rho\to 0}\hat{w}^j_\rho(y|\hat{a}^i) = \hat{w}^j(y|\hat{a}^i) = \hat{w}^i(y|\hat{a}^i)$ for each $y \in Y$ because $\hat{U}^i = \hat{U}^j$, $P_n^j(a) = P_n^i(a)$ and $G_i(a) = G_j(a)$ are all satisfied when $\rho = 0$ so that $(\varepsilon, \delta) = (0, 0)$, and hence the optimal wage schemes which solve Problem $(\text{CM}^k - \hat{a}^i)$ for agents $k = i, j$ must be the same for implementation of a given action $\hat{a}^i$. Then, since $\hat{w}^i(y|\hat{a}^i)$ is well-defined, $\hat{w}^j(y|\hat{a}^i)$ does so when $\rho = 0$. Then, by letting ε and δ to be small enough (thus $\rho \to 0$), we obtain from (A6) that $\sum_n P_n^i(\hat{a}^i)\hat{w}^i(y_n|\hat{a}^i) \simeq \sum_n P_n^j(\hat{a}^i)\hat{w}^i(y_n|\hat{a}^i) \geq \sum_n P_n^j(\hat{a}^i)\hat{w}^j_\rho(y_n|\hat{a}^i) \simeq \sum_n P_n^j(\hat{a}^i)\hat{w}^j(y_n|\hat{a}^i)$. Also, from (A7) we obtain $\sum_n P_n^j(\hat{a}^j)\hat{w}^j(y_n|\hat{a}^j) \simeq \sum_n P_n^i(\hat{a}^j)\hat{w}^j(y_n|\hat{a}^j) \geq \sum_n P_n^i(\hat{a}^j)\hat{w}^i_\rho(y_n|\hat{a}^j) \simeq \sum_n P_n^i(\hat{a}^j)\hat{w}^i(y_n|\hat{a}^j)$. Thus, by taking small enough (ε, δ) so that ρ is close to zero, we can ensure the inequalities in Lemma A1. Q.E.D.

In what follows we fix such $\varepsilon > 0$ and $\delta > 0$ to ensure Lemma A1.

Lemma A2. *$\hat{a}^i = \hat{a}^j$ must hold for all risk averse agents $i, j \in I_r$ in any fair contract $\hat{C}$.*

Proof. Take any fair contract $\hat{C}$. Then first note that the action profile $\hat{\mathbf{a}}^r \in A^{N_r}$ of risk averse agents must maximize the expected residual surplus which the risk neutral agents totally receive after paying the total expected wages to all risk averse agents:

$$E_y[R(\mathbf{y})|\mathbf{a}^r, \mathbf{a}^n] - \sum_{i \in I_r} E_{y_i}[\hat{w}^i(y^i)|a_i].$$

Otherwise, $\hat{C}$ is not incentive efficient because some risk neutral agents can be strictly better off by implementing the action profile of risk averse agents $\mathbf{a}^r$ which maximizes the above expected residual surplus. In this proof we will omit the vector of the risk neutral agents' actions $\mathbf{a}^n$ from the argument of the above function because it does not play any role in the proof.

Here recall that, since the fair contract must solve Problem $(\text{CM}^i - \hat{a}^i)$ for each risk averse agent $i \in I_r$, we have $\hat{w}^i(y) = \hat{w}^i(y|\hat{a}^i)$ for each $i \in I_r$.

Now suppose that $\hat{a}^j \equiv a'' > \hat{a}^i \equiv a'$ for some risk averse agents $i, j \in I_r$. Then we have

$$\begin{aligned} &E_y[R(\mathbf{y})|\hat{a}^j, \hat{a}^i, \hat{a}^{-i-j}] - E_{y_j}[\hat{w}^j(y^j|\hat{a}^j)|\hat{a}^j] \\ &\geq E_y[R(\mathbf{y})|\hat{a}^i, \hat{a}^i, \hat{a}^{-i-j}] - E_{y^j}[\hat{w}^j(y^j|\hat{a}^i)|\hat{a}^i] \end{aligned}$$

and

$$\begin{aligned} &E_y[R(\mathbf{y})|\hat{a}^i, \hat{a}^j, \hat{a}^{-i-j}] - E_{y^i}[\hat{w}^i(y^i|\hat{a}^i)|\hat{a}^i] \\ &\geq E_y[R(\mathbf{y})|\hat{a}^j, \hat{a}^j, \hat{a}^{-i-j}] - E_{y^i}[\hat{w}^i(y^i|\hat{a}^j)|\hat{a}^j]. \end{aligned}$$

From these inequalities, we derive

$$\begin{aligned} &E_{y^i}[\hat{w}^i(y^i|a'')|a''] - E_{y^i}[\hat{w}^i(y^i|a')|a'] \\ &\geq \ \Delta_i(a''|a'', a') \\ &\geq \Delta_j(a''|a'', a') - \alpha \\ &\geq \ \Delta_j(a'|a'', a') + \tilde{\beta} - \alpha \ \ \text{(by Assumption 4)} \\ &\geq \ E_{y^j}[\hat{w}^j(y^j|a'')|a''] - E_{y^j}[\hat{w}^j(y^j|a')|a'] + \beta \end{aligned}$$

and hence by Lemma A1

$$\begin{aligned} \frac{1}{2}\beta \ &> \ E_{y^i}[\hat{w}^i(y^i|a'')|a''] - E_{y^j}[\hat{w}^j(y^j|a'')|a''] \\ &\geq \ E_{y^i}[\hat{w}^i(y^i|a')|a'] - E_{y^j}[\hat{w}^j(y^j|a')|a'] + \beta \\ &> \ -\frac{1}{2}\beta + \beta \\ &= \ \frac{1}{2}\beta, \end{aligned}$$

a contradiction. Thus we must have $\hat{a}^j = \hat{a}^i$. Q.E.D.

By Lemma A2, we must have $\hat{a}^j = \hat{a}^i$ for any $i, j \in I_r$, $i \neq j$, in any fair contract $\hat{C}$. Thus we will denote $\hat{a}^i = \hat{a}$ for all $i \in I_r$.

Now we pick any two risk averse agents $i, j \in I_r$ where $i > j$ in the sense of Assumption 5. Suppose then that $\hat{a} \equiv \hat{a}^i = \hat{a}^j > 0$. Then, since Assumption 3 is satisfied under Assumption 5 when ε is small enough given δ, we know from the proof of Theorem 1 that (IC) of Problem ($\text{CM}^j - \hat{a}$) must be slack for more efficient risk averse agent j at any action $a < \hat{a}^j = \hat{a}$ under the fair contract $\hat{w}^j(y)$. However, this contradicts to the fact that the optimal wage scheme $\hat{w}^j(y) = \hat{w}^j(y|\hat{a})$ which solves Problem ($\text{CM}^j - \hat{a}$) must have the property that (IC) is binding at some $a < \hat{a}$ (otherwise, the contract which solves Problem ($\text{CM}^j - \hat{a}$) must be the fixed wage contract but then it cannot satisfy (IC) for agent j to choose $\hat{a} > 0$). Thus the supposition that $\hat{a} > 0$ is false, so we must have $\hat{a} = 0$. Since $\hat{a}^l = \hat{a}^k$ for all $l, k \in I_r$ by Lemma A2, we then have $\hat{a}^i = 0$ for all $i \in I_r$. Then (IE*) implies that $\hat{w}^i(y)$ must be the fixed wage contract, i.e., $\hat{w}^i(y)$ does not depend on $y \in Y$ for all $i \in I_r$ (otherwise, $\hat{a}^j = 0$ but $\hat{w}^j(y)$ is not fixed wage for some $j \in I_r$. However this is not incentive efficient because offering a fixed wage contract to such agent $j \in I_r$ improves the welfare of some risk neutral agents while keeping all the actions and contracts of others unchanged.)

We define the set of the values ε and δ representing the heterogeneity of agents as follows:

$$D \equiv \{(\varepsilon, \delta) \in \Re_+ \mid (\varepsilon, \delta) \text{ satisfies (6), (7) and Lemma A1.}\}.$$

Then we obtain Theorem 2 for all $(\varepsilon, \delta) \in D$. Q.E.D.

8.5 Proof of Theorem 3

Given an action profile of risk averse agents in a fair contract $\mathbf{a}^r = (\hat{a}^i)_{i \in I_r}$, the maximum residual surplus which all risk neutral agents can share is

$$E[R(\mathbf{y})|\mathbf{a}^r, \mathbf{a}^n] - \sum_{i \in I_n} G_i(a^i) - \sum_{i \in I_r} E_{y^i}[\hat{w}^i(y^i)|\hat{a}^i]$$

where $\mathbf{a}^n = (a^i)_{i \in I_n}$ denotes the vector of actions of risk neutral agents.

By definition, $(\tilde{a}^i)_{i \in I_n}$ maximizes the above surplus. Let define the best response of action a^i for each risk neutral agent $i \in I_n$, given $a^{-i} \in A^{N_n - 1}$:

$$BR^i(a^{-i}) \equiv \arg\max_{a^i \in A} E_y[R(\mathbf{y})|\mathbf{a}^r, (a^i, a^{-i})] - \sum_{i \in I_n} G_i(a^i).$$

Then, the definition of $\tilde{a}^i$ means that $\tilde{a}^i \in BR^i(\tilde{a}^{-i})$ for each $i \in I_n$.

Now suppose contrary to the claim that the wage schemes of risk neutral agents in some fair contract $\hat{w}^i(\mathbf{y})$ do not implement the optimal actions $(\tilde{a}^i)_{i\in I_n}$ defined above. Let $(\hat{a}^i)_{i\in I_n}$ be the implemented action profile of risk neutral agents under $\{\hat{w}^i(\mathbf{y})\}_{i\in I_n}$. Since $(\hat{a}^i)$ does not maximize the residual surplus, there must exist some risk neutral agent k such that

$$\hat{a}^k \notin BR^k(\hat{a}^{-k}).$$

Then we define the alternative wage schemes as follows:

$$\tilde{w}^i(y^i) \equiv E_{y^{-i}}[\hat{w}^i(y^i, y^{-i})|\hat{a}^{-i}], \ \ \forall\, i \neq k \quad \forall\, i \in I_n,,$$

$$\tilde{w}^i(y^i) = \hat{w}^i(y^i) \ \ \forall\, i \in I_r,$$

and

$$\tilde{w}^k(\mathbf{y}) = R(\mathbf{y}) - \sum_{i\neq k, i\in I} \tilde{w}^i(y^i)$$

Note that the wage scheme of risk neutral agent k depends on y^{-k} only through $R(\mathbf{y})$ so that its expected wage depends on a^{-k} only through $E[R(\mathbf{y})|a^k, a^{-k}]$.

These new schemes satisfy (BB) by construction. Also, since $\{\hat{w}^i(\mathbf{y})\}_{i\in I_n}$ implements the action profile $(\hat{a}^i)_{i\in I_n}$ and $(y^i)_{i\in I_n}$ are statistically independent, the new scheme $\tilde{w}^i(y^i)$ still implements action $\hat{a}^i$ from agent $i \neq k$. Thus all agents but k chooses the same actions $\hat{a}^i$ as before. Finally, risk neutral agent k has the strict incentive to choose some action in $BR^k(\hat{a}^{-k})$ rather than $\hat{a}^k$ because of $\hat{a}^k \notin BR^k(\hat{a}^{-k})$. Then we have

$$\begin{aligned}
&\max_{a\in A} E_y[\tilde{w}^k(\mathbf{y})|\mathbf{a}^r, (a, \hat{a}^{-k})] - G_k(a) - \sum_{i\neq k} E_{y^i}[\tilde{w}^i(y^i)|\hat{a}^i] \\
&= \max_{a\in A} E_y[R(\mathbf{y})|\mathbf{a}^r, (a, \hat{a}^{-k})] - G_k(a) - \sum_{i\neq k} E_y[\hat{w}^i(\mathbf{y})|\hat{\mathbf{a}}] \\
&> E_y[R(\mathbf{y})|\mathbf{a}^r, \hat{a}^k, \hat{a}^{-k}] - G_k(a) - \sum_{i\neq k} E_{y^i}[\hat{w}^i(y^i)|\hat{\mathbf{a}}] \\
&= E_y[\hat{w}^k(\mathbf{y})|\mathbf{a}^r, (\hat{a}^k, \hat{a}^{-k})] - G_k(\hat{a}^k)
\end{aligned}$$

where the inequality follows from $\hat{a}^k \notin BR^k(\hat{a}^{-k})$ and the last equality from (BB): $R(\mathbf{y}) = \sum_{i\in I} \hat{w}^i(\mathbf{y})$.

However, then this contradicts the fact that $\hat{C}^i = \{\hat{w}^i, \hat{a}^i\}$ is incentive efficient. Q.E.D.

8.6 Proof of Lemma 4

Suppose that Assumption 4 and 6 hold. Suppose also contrary to the claim that there exists some asymmetric optimal action profile $\tilde{\mathbf{a}}^n$ which maximizes the residual surplus $RS(\mathbf{a}^n)$. Then there must exist some $l, k \in I_n$, $l \neq k$ such that $\tilde{a}^l \neq \tilde{a}^k$.

Let $a'' \equiv \tilde{a}^l > a' \equiv \tilde{a}^k$ without loss of generality. Then we have

$$\begin{aligned}&E_y[R(\mathbf{y})|\mathbf{0}, a'', a', \tilde{a}^{-l-k}] - G_l(a'')\\ \geq\ &E_y[R(\mathbf{y})|\mathbf{0}, a', a', \tilde{a}^{-l-k}] - G_l(a')\end{aligned}$$

and

$$\begin{aligned}&E_y[R(\mathbf{y})|\mathbf{0}, a', a'', \tilde{a}^{-l-k}] - G_k(a')\\ \geq\ &E_y[R(\mathbf{y})|\mathbf{0}, a'', a'', \tilde{a}^{-l-k}] - G_k(a'').\end{aligned}$$

By adding these inequalities, we have

$$\begin{aligned}G_k(a'') - G_k(a') &\geq \Delta_k(a''|a'', a')\\ &\geq \Delta_k(a'|a'', a') + \tilde{\beta}\\ &\geq \Delta_l(a'|a'', a') - \alpha + \tilde{\beta}\\ &= \Delta_l(a'|a'', a') + \beta\\ &\geq G_l(a'') - G_l(a') + \beta\end{aligned}$$

which implies that

$$\begin{aligned}\frac{1}{2}\beta &> G_l(a') - G_k(a')\\ &\geq G_l(a'') - G_k(a'') + \beta\\ &> -\frac{1}{2}\beta + \beta\\ &= \frac{1}{2}\beta,\end{aligned}$$

a contradiction. Q.E.D.

8.7 Proof of Theorem 4

By Lemma 4, we can ensure the existence of a symmetric optimal action $\tilde{a}$ which maximizes the residual surplus of organization $RS(\mathbf{a}^n)$ over the actions of risk neutral agents $\mathbf{a}^n$ after subtracting fixed wages of risk averse agents. In what follows we take such action $\tilde{a}$ and consider the implementation of $a^i = \tilde{a}$ from all risk neutral agents $i \in I_n$.

Now we will show that offering a fixed wage $\overline{w}$ to every risk averse agent $i \in I_r$ becomes incentive efficient. To this end, it suffices to show that the fixed wage contract $\overline{C} \equiv \{\overline{w}, a^i = 0\}$ maximizes the residual surplus of organization which belongs to all risk neutral agents, given any action profile of risk neutral agents $\mathbf{a}^n \in A^{N_n}$. This is because then there exist no other feasible contracts which Pareto dominate the fixed wage contract.

Formally, we show the following lemma:

Lemma A3. *When the fixed wage $\overline{w}$ is large enough, it becomes incentive efficient.*

Proof. We consider the following maximization problem which states that the expected residual surplus of risk neutral agents is maximized subject to (IC) and the acceptance constraint ensuring that risk averse agent cannot be worse off by switching to a new contract rather than sticking to the fixed wage $\overline{w}$:

Problem (OP)

$$\max_{\mathbf{a}^r \in A^{N_r}, w^i(y)} E_y[R(\mathbf{y})|\mathbf{a}^r, \mathbf{a}^n] - \sum_{i \in I_r} E_{y^i}[w^i(y^i)|a^i]$$

subject to

$$E_{y^i}[u(w^i(y^i))|a^i, a^{-i}] - G_i(a^i) \geq E_{y^i}[u(w^i(y^i))|a, a^{-i}] - G_i(a) \;\; \forall \, a \neq a^i \qquad \text{(IC)}$$

$$E_{y^i}[u(w^i(y^i))|a^i, a^{-i}] - G_i(a^i) \geq u(\overline{w}) \qquad \text{(PE)}$$

Here (PE) is the constraint to ensure that risk averse agent $i \in I_r$ can be better off by accepting a contract $C^i \neq \overline{C}$.

Then we show that the contract to solve the above problem becomes the fixed wage contract $\overline{C}$ for any given action profile of risk neutral agents $\mathbf{a}^n$ when $\overline{w}$ is large enough. To see this, note that the objective function in the above problem is bounded above by

$$E_y[R(\mathbf{y})|\mathbf{a}^r, \mathbf{a}^n] - \sum_{i \in I_r} \phi(G_i(a^i) + u(\overline{w}))$$

because, by concavity of u and (PE), we have

$$E_y[w^i(\mathbf{y})|a^i] \geq \phi(G_i(a^i) + u(\overline{w})).$$

Now consider the following maximization problem:

$$\max_{\mathbf{a}^r \in A^{N_r}} E_y[R(\mathbf{y})|\mathbf{a}^r, \mathbf{a}^n] - \sum_{i \in I_r} \phi(G_i(a^i) + u(\overline{w})).$$

Then we can verify that $\mathbf{a}^r = \mathbf{0}$ solves this maximization problem when

$$\phi(G_i(a^i) + u(\overline{w})) - \phi(G_i(0) + u(\overline{w})) \geq E_y[R(y)|\mathbf{0}, \mathbf{a}^n] - E_y[R(y)|\mathbf{a}^r, \mathbf{a}^n].$$

In fact we can find some $\overline{w}$ so that the above inequality holds because the right hand side is finite (by finiteness of A) but the left hand side goes to $+\infty$ as $\overline{w} \to +\infty$:

$$\begin{aligned} & \phi(G_i(a^i) + u(\overline{w})) - \phi(G_i(0) + u(\overline{w})) \\ > \;& \phi'(G_i(0) + u(\overline{w}))(G_i(a_i) - G_i(0)) \\ \to \;& +\infty \quad \text{as } \overline{w} \to +\infty \end{aligned}$$

due to the convexity of ϕ and $\lim_{w\to+\infty}\phi'(w)=+\infty$. Thus, when $\overline{w}$ is large enough, the optimal contract which solves Problem (OP) becomes the fixed wage $\overline{w}$ for all risk averse agents. Q.E.D.

Next we consider the incentives of risk neutral agents. We take the risk neutral agent $r \in I_n$ whose action cost is the largest among all risk neutral agents, i.e.,

$$r \in \arg\max_{i\in I_n} G_i(\tilde{a}).$$

We will define the wage scheme of the residual claimant later.

For the wage schemes of all other risk neutral agents than the residual claimant r, we will apply the following lemma.

Lemma A4. *Suppose that Assumption 8 is satisfied. Then there exists a non-decreasing individualistic wage scheme $\tilde{w}(y^l)$ for each risk neutral agent $l \in I_n$ with $l \neq r$ such that it implements the optimal action $\tilde{a}$:*

$$E_{y^l}[\tilde{w}(y^l)|\tilde{a}] - G_l(\tilde{a}) \geq E_{y^l}[\tilde{w}(y^l)|a] - G_l(a) + \delta$$

for all $a \neq \tilde{a}$ where δ is given in (6) in the main text.

Proof. The implementation problem above is reduced to find a $L-1$ dimensional vector $(\Delta w_n)_{n=1}^{L-1}$ such that

$$\sum_n (1 - F_n^l(\tilde{a}))\Delta w_n - G_l(\tilde{a}) \geq \sum_n (1 - F_n(a))\Delta w_n - G_l(a) + \delta, \ \ \forall\, a \neq \tilde{a}.$$

and

$$\Delta w_n \geq 0, \ \ n = 1, 2, ..., L-1.$$

The last condition is monotonicity requirement (MON) that the wage schedule w_n must be non-decreasing. By Avis and Kaluzny (Theorem 2, 2004), such non-negative vector (Δw_n) exists under Assumption 8. Q.E.D.

By Lemma A4, we obtain

$$\begin{aligned} E_z[\tilde{w}(z)|\tilde{a}] - E_z[\tilde{w}(z)|a] &\geq G_l(\tilde{a}) - G_l(a) + \delta \\ &\geq G_i(\tilde{a}) - G_i(a). \end{aligned}$$

for all $i \in I_n$ with $i \neq n$. Thus every risk neutral agent $i \in I_n$ other than the residual claimant r chooses the same optimal action $\tilde{a}$ under the same scheme $\tilde{w}$.

Define $\hat{w}(y^i) \equiv \tilde{w}(y^i) + f$ for $i \in I_n$, $i \neq r$, where f will be specified below, and offer such scheme to all $i \neq r$, $i \in I_n$. Then it is clear that every risk neutral agent $i \neq r$, $i \in I_n$, does not envy other risk neutral agent $j \neq i, r$, $j \in I_n$.

The wage scheme for residual claimant r is given by

$$w^r(y^r, y^{-r}) \equiv R(y^r, y^{-r}) - N_r\overline{w} - \sum_{i \in I_n} \hat{w}(y^i) + \hat{w}(y^r).$$

Since $\hat{w}(y^i)$ depends only on his performance y^i for each risk neutral agent i except the residual claimant r, the wage scheme of residual claimant $w^r(y^r, y^{-r})$ depends on his performance y^r only through $R(y^r, y^{-r})$. Since R is monotone increasing in y^r, w^r does so. Thus (MON) is satisfied for the residual claimant r as well.

We also define the fixed wage which is offered to risk averse agent:

$$\overline{w} = \frac{1}{N}\{E[R(y)|\mathbf{0}, \tilde{\mathbf{a}}] - N_n G_r(\tilde{a})\}.$$

Given such $\overline{w}$, we set f to satisfy

$$E[\tilde{w}(z)|\tilde{a}] + f = G_r(\tilde{a}) + \overline{w}.$$

Then, given such $(\overline{w}, f)$, we can show that

$$\begin{aligned}
& E_y[R(\mathbf{y})|\mathbf{0}, \tilde{\mathbf{a}}] - N_r\overline{w} - (N_n - 1)E_z[\hat{w}(z)|\tilde{a}] - G_r(\tilde{a}) \\
&= E_y[R(\mathbf{y})|\mathbf{0}, \tilde{\mathbf{a}}] - N_r\overline{w} - (N_n - 1)(E_z[\tilde{w}(z)|\tilde{a}] + f) - G_r(\tilde{a}) \\
&= E_z[\tilde{w}(z)|\tilde{a}] + f - G_r(\tilde{a}) \\
&= \overline{w}.
\end{aligned}$$

Then we show the following series of lemmas:

Lemma A5. *The residual claimant r does not envy other agents.*

Proof. By using the wage schemes defined above, we can verify that the residual claimant r does not envy other risk neutral agent $i \neq r$, $i \in I_n$, because

$$\begin{aligned}
& \max_{a^r \in A} E_y[R(\mathbf{y})|\mathbf{0}, (a^r, \tilde{a}^{-r})] - N_r\overline{w} - (N_n - 1)E[\hat{w}(z)|\tilde{a}] - G_r(a^r) \\
&\geq E_y[R(\mathbf{y})|\mathbf{0}, \tilde{\mathbf{a}}] - N_r\overline{w} - (N_n - 1)E_z[\hat{w}(z)|\tilde{a}] - G_r(\tilde{a}) \\
&= E_z[\hat{w}(z)|\tilde{a}] - G_r(\tilde{a}) \\
&= E_z[\tilde{w}(z)|\tilde{a}] + f - G_r(\tilde{a}) \\
&\geq E_z[\tilde{w}(z)|a] - G_r(a) + f \\
&= E_z[\hat{w}(z)|a] - G_r(a) \ \ \forall\, a \in A
\end{aligned}$$

Also the residual claimant r does not envy any risk averse agent $j \in I_r$ because

$$\begin{aligned}
& \max_{a^r \in A} E_y[R(\mathbf{y})|\mathbf{0}, a^r, \tilde{a}^{-r}] - N_r\overline{w} - (N_n - 1)E[\hat{w}(z)|\tilde{a}] - G_r(a^r) \\
&\geq E_y[R(\mathbf{y})|\mathbf{0}, \tilde{\mathbf{a}}] - N_r\overline{w} - (N_n - 1)E_z[\hat{w}(z)|\tilde{a}] - G_r(\tilde{a}) \\
&= \overline{w} \\
&\geq \overline{w} - G_r(a).
\end{aligned}$$

Q.E.D.

Lemma A6. *Each risk neutral agent other than the residual claimant r does not envy other agents.*

Proof. First we show that any other risk neutral agent $i \in I_n$ than the residual claimant r does not envy any risk averse agent. This is because

$$\begin{aligned} E_z[\hat{w}(z)|\tilde{a}] - G_i(\tilde{a}) &\geq E_z[\hat{w}(z)|\tilde{a}] - G_r(\tilde{a}) \\ &= \overline{w} \\ &\geq \overline{w} - G_i(a) \end{aligned}$$

for all $a \neq \tilde{a}$, where $G_i(\tilde{a}) \leq G_r(\tilde{a})$ for all $i \neq r$.

We also show that every risk neutral agent $i \in I_n$, $i \neq r$, does not envy the residual claimant r because the expected payoff of agent $i \neq r$ would be the following if he were offered the wage scheme of agent r (i.e., he were residual claimant):

$$E_y[R(\mathbf{y})|\mathbf{0}, a^i, \tilde{a}^{-i}] - N_r\overline{w} - \sum_{j \in I_n, j \neq i} E_{y^j}[\hat{w}(y^j)|\tilde{a}] - G_i(a^i)$$

where agent i was supposed to choose a_i. Since agent i obtains $E_{y^i}[\hat{w}(y^i)|\tilde{a}] - G_i(\tilde{a})$ under his own contract, he prefers his own contract $\hat{w}$ to the one offered to the residual claimant r if

$$\begin{aligned} &E_z[\hat{w}(z)|\tilde{a}] - G_i(\tilde{a}) \\ &\geq E_y[R(\mathbf{y})|\mathbf{0}, a_i, \tilde{a}_{-i}] - N_r\overline{w} - \sum_{j \in I_n, j \neq i} E_{y^i}[\hat{w}(y^j)|\tilde{a}] - G_i(a^i) \end{aligned}$$

where the right hand side cannot be greater than

$$E_y[R(\mathbf{y})|\mathbf{0}, \tilde{a}, \tilde{a}_{-i}] - N_r\overline{w} - (N_n - 1)E_z[\hat{w}(z)|\tilde{a}] - G_i(\tilde{a})$$

by definition of $\tilde{a}$ (which maximizes $E_y[R(\mathbf{y})|\mathbf{0}, a_i, \tilde{a}_{-i}] - G_i(a)$ over $a \in A$). Thus it suffices to show that

$$\begin{aligned} &E_z[\hat{w}(z)|\tilde{a}] - G_i(\tilde{a}) \\ &\geq E_y[R(\mathbf{y})|0, \tilde{\mathbf{a}}] - N_r\overline{w} - (N_n - 1)E_z[\hat{w}(z)|\tilde{a}] - G_i(\tilde{a}) \end{aligned}$$

which is equivalent to

$$E_z[\hat{w}(z)|\tilde{a}] \geq E_y[R(\mathbf{y})|\mathbf{0}, \tilde{\mathbf{a}}] - N_r\overline{w} - (N_n - 1)E_z[\hat{w}(z)|\tilde{a}].$$

This holds as equality due to the definitions of $\hat{w}(z) = \tilde{w}(z) + f$ and $\overline{w}$.

Finally, it is clear that any risk neutral agent $j \neq r$ does not envy other risk neutral agent $k \neq r$ because they are offered the same wage scheme. Q.E.D.

Lemma A7. *Each risk averse agent $i \in I_r$ does not envy other agents when δ is small enough and F is large enough.*

Proof. First we show that every risk averse agent does not envy any other risk neutral agent $j \in I_n$ than the residual claimant r when δ is small and $\overline{w}$ is large enough. To see this, we define η_i for each $i \in I$, which satisfies

$$E[\hat{w}(z)|\tilde{a}] - G_i(\tilde{a}) - \overline{w} = \eta_i.$$

Note here that $\eta_r = 0$ and $\eta_i \to 0$ as $\delta \to 0$. Take any risk averse agent $i \in I_r$. Then we can show that when $\delta > 0$ is sufficiently small and the fixed wage $\overline{w}$ can be large enough, we have

$$\begin{aligned}
& u(\overline{w}) - u(E_z[\hat{w}(z)|a]) \\
& \geq u'(\overline{w})\{\overline{w} - E_z[\hat{w}(z)|a]\} \\
& = u'(\overline{w})\{E_z[\hat{w}(z)|\tilde{a}] - G_i(\tilde{a}) - \eta_i - E_z[\hat{w}(z)|a]\} \\
& \geq u'(\overline{w})\{-G_i(a) - \eta_i\} \\
& \geq -G_i(a) - \eta_i
\end{aligned}$$

where the first inequality follows from concavity of u, the second inequality from the fact that $E_z[\hat{w}(z)|\tilde{a}] - G_i(\tilde{a}) \geq E_z[\hat{w}(z)|a] - G_i(a)$ for all $a \neq \tilde{a}$, and the third inequality from the fact that $u'(\overline{w}) \leq 1$ for large enough $\overline{w}$ (by using $u'(+\infty) = 0$) respectively. Thus we obtain

$$\begin{aligned}
u(\overline{w}) & \geq u(E[_z\hat{w}(z)|a]) - G_i(a) - \eta_i \\
& > E_z[u(\hat{w}(z))|a] - G_i(a)
\end{aligned}$$

for all $a \in A$ when δ is so small that η_i can be close to zero if $\overline{w}$ becomes large enough.

Next we show that every risk averse agent $i \in I_r$ does not envy the residual claimant r when ε is sufficiently small. To see this, note that

$$\begin{aligned}
& u(\overline{w}) \\
& = u\left(\frac{1}{N}\{E_y[R(\mathbf{y})|\mathbf{0}, \tilde{\mathbf{a}}] - N_n G_r(\tilde{a})\}\right) \\
& = u(E_y[R(\mathbf{y})|\mathbf{0}, \tilde{\mathbf{a}}] - N_r\overline{w} - (N_n - 1)E_z[\hat{w}(z)|\tilde{a}] - G_r(\tilde{a})).
\end{aligned}$$

Then we want to show

$$u(\overline{w}) \geq E\left[u\left(R(\mathbf{y}) - N_r\overline{w} - \sum_{j \in I_r} \hat{w}(y^j) + \hat{w}(y^i)\right) \middle| a_i, \hat{a}_{-i}\right] - G_i(a_i) \tag{A8}$$

for all $a_i \in A$. For this, by the Jensen's inequality, it suffices to show that

$$\begin{aligned} & u(E_y[R(\mathbf{y})|\mathbf{0},\tilde{\mathbf{a}}] - N_r\overline{w} - (N_n-1)E_z[\hat{w}(z)|\tilde{a}] - G_r(\tilde{a})) \\ \geq \;& u(E_y[R(\mathbf{y})|\mathbf{0},a^i,\tilde{a}^{-i}] - N_r\overline{w} - N_nE_z[\hat{w}(z)|\tilde{a}] + E_z[\hat{w}(z)|a^i]) - G_i(a^i) \end{aligned}$$

for all $a^i \in A$. To show this, note that by definition of $\overline{w}$,

$$\overline{w} = E_y[R(\mathbf{y})|\mathbf{0},\tilde{\mathbf{a}}] - N_r\overline{w} - (N_n-1)E_z[\hat{w}(z)|\tilde{a}] - G_r(\tilde{a}).$$

Let

$$B \equiv E_y[R(\mathbf{y})|\mathbf{0},a_i,\tilde{a}_{-i}] - N_r\overline{w} - N_nE_z[\hat{w}(z)|\tilde{a}] + E_z[\hat{w}(z)|a_i].$$

Then the above inequality (A8) can be written by

$$u(\overline{w}) \geq u(B) - G_i(a^i)$$

which can be further written by

$$\begin{aligned} & u(\overline{w}) - [u(B) - G_i(a^i)] \\ \geq \;& u'(\overline{w})(\overline{w} - B) + G_i(a^i) \end{aligned}$$

where

$$\begin{aligned} \overline{w} - B \;=\; & E_y[R(\mathbf{y})|\mathbf{0},\tilde{\mathbf{a}}] - E_y[R(\mathbf{y})|\mathbf{0},a^i,\tilde{a}^{-i}] \\ & + E_z[\hat{w}(z)|\tilde{a}] - G_r(\tilde{a}) - E_z[\hat{w}(z)|a^i]. \end{aligned}$$

Under Assumption 9 $E_y[R(\mathbf{y})|\mathbf{0},\tilde{\mathbf{a}}] - E_y[R(\mathbf{y})|\mathbf{0},a_i,\tilde{a}_{-i}]$ is independent of F. We also know that $E_z[\hat{w}(z)|\tilde{a}] - E_z[\hat{w}(z)|a] = E_z[\tilde{w}(z)|\tilde{a}] - E_z[\tilde{w}(z)|a]$ is independent of F (because $\tilde{w}(z)$ is independent of F). Thus $E_z[\hat{w}(z)|\tilde{a}] - E_z[\hat{w}(z)|a]$ is independent of F as well. Also the optimal action of risk neutral agents $\tilde{a}$ does not vary with F. Then the above term $\overline{w} - B$ is independent of F. We can also show that $\overline{w}$ increases in F (because the optimal action $\tilde{a}$ is independent of F so that it does not change with F). Thus we can ensure that $u'(\overline{w}) \to 0$ as $F \to +\infty$, and hence

$$\begin{aligned} u(\overline{w}) - [u(B) - G_i(a^i)] \;&\geq\; u'(\overline{w})(\overline{w} - B) + G_i(a^i) \\ &\simeq\; G_i(a^i) \\ &>\; 0 \end{aligned}$$

for all $a^i > 0$ when F is sufficiently large. Thus, since u is strictly concave and $R(\mathbf{y}) - N_r\overline{w} - \sum_{j\neq r}\hat{w}(y^j) + \hat{w}(y^r)$ is random, we can verify that when F is sufficiently large,

$$\begin{aligned} & u(\overline{w}) \\ \geq \;& u(B) - G_i(a^i) \\ > \;& E_y\left[u\left(R(\mathbf{y}) - N_r\overline{w} - \sum_{j\in I_n}\hat{w}(y_j) + \hat{w}(y_i)\right)\middle|0,a^i,\tilde{a}^{-i}\right] - G_i(a_i) \end{aligned}$$

for all $a_i \in A$. Thus agent $i \in I_r$ does not envy the residual claimant r. Q.E.D.

From Lemma A5-A7, we have established the result that any agent does not envy others under the constructed contract.

Finally, when F is large enough, we can take large f and $\overline{w}$ so that

$$E_z[\hat{w}(z)|\tilde{a}] - G_i(\tilde{a}) \geq \overline{U}^i$$

for all $i \in I_n$ but $i \neq r$. Thus we can satisfy (IR) for all $i \in I_n$. Also, since the expected payoff of the residual claimant r is same as $\overline{w}$, we can also ensure that the residual claimant r and all risk averse agents $i \in I_r$ obtain larger payoffs than their reservation payoffs $\overline{U}^i$ as well when F is sufficiently large (note that $\overline{w}$ can be large when F is large).

8.8 Proof of Theorem 6

We will modify the proof of Theorem 2 by replacing (MON) by (LL). The main modification of the proof is that we show that Assumption 3 is satisfied when we drop (MON) but impose (LL), if ε is sufficiently small given δ.

Take two risk averse agents i and j. Here agent j is more efficient than i in the sense that $\Delta G_i(\hat{a}, a) > \Delta G_j(\hat{a}, a)$ for all $\hat{a} > a$ (Assumption 5). Recall that these agents must choose the same action $\hat{a}$ due to Lemma 5, when their heterogeneity is sufficiently small.

Now suppose that $\hat{a} > 0$ as in the proof of Theorem 2. Then recall that we have reached the following conclusion that (IC) of more efficient risk averse agent j is slack at $\hat{a} > 0$ if

$$\Delta G_i(\hat{a}, a) - \Delta G_j(\hat{a}, a) + \sum_n (P_n^j(\hat{a}) - P_n^i(\hat{a}))\hat{u}_n^i - \sum_n (P_n^j(a) - P_n^i(a))\hat{u}_n^j \qquad \text{(A9)}$$

is strictly positive for all $a < \hat{a}$.

From now on we will show that this is actually the case when ε is sufficiently small, given (LL). Since $\overline{u} \geq \hat{u}_n^i \geq \underline{u}$ for all $y_n \in Y$ due to (BB) and (LL), the second term in (A9) is bounded below from

$$\Gamma^{ij}(\hat{a}) \equiv \sum_{y_n \in \overline{Y}(\hat{a})} (P_n^j(\hat{a}) - P_n^i(\hat{a}))\underline{u} - \sum_{y_n \in Y \setminus \overline{Y}(\hat{a})} (P_n^j(\hat{a}) - P_n^i(\hat{a}))\overline{u}$$

where $\overline{Y}(\hat{a})$ is defined as $\overline{Y}(\hat{a}) \equiv \{y_n \in Y \mid P_n^j(\hat{a}) \geq P_n^i(\hat{a})\}$. Also, the third term in (A9) is bounded above by

$$\Omega^{ij}(a) \equiv \sum_{y_n \in Y(a)} (P_n^j(a) - P_n^i(a))\overline{u} + \sum_{y_n \in Y \setminus \overline{Y}(a)} (P_n^j(a) - P_n^i(a))\underline{u}.$$

Thus (A9) is bounded below from

$$\Delta G_i(\hat{a}, a) - \Delta G_j(\hat{a}, a) + \Gamma^{ij}(\hat{a}) - \Omega^{ij}(a).$$

Here, if we take ε to be sufficiently small, both $\Gamma^{ij}(\hat{a})$ and $\Omega^{ij}(a)$ become small enough. Then, since $\Delta G_i(\hat{a}, a) > \Delta G_j(\hat{a}, a)$ for $i > j$ by Assumption 5, we can show that

$$\Delta G_i(\hat{a}, a) - \Delta G_j(\hat{a}, a) + \Gamma^{ij}(\hat{a}) - \Omega^{ij}(a) > 0$$

for all $\hat{a} > a$ when ε is sufficiently small, given δ. Thus, when we take ε to be sufficiently small given δ, we can establish the result that (IC) of more efficient risk averse agent j is not binding at the action $\hat{a} > 0$ at fair contract $\hat{C}^j$. However, this is not incentive efficient as we have argued in the proof of Theorem 2. Q.E.D.

References

[1] Avis, D. and B. Kaluzny (2004), "Solving Inequalities and Proving Farkas's Lemma Made Easy," *American Math Monthly*, 111, 152-157.

[2] Bartling, B. (2010), "Relative Performance or Team Evaluation? Optimal Contracts for Other-Regarding Agents," University of Zurich.

[3] Englmaier, F. and A. Wamback (2010), "Optimal Incentive Contracts under Inequity Aversion," *Games and Economic Behavior* 69, 312-28.

[4] Fehr, E. and K. M. Schmidt (2004), "Fairness and Incentives in a Multi-task Principal-Agent Model," *Scandinavian Journal of Economics*, 106, 453-474.

[5] Fehr, E. and K. M. Schmidt (2000), "Fairness, incentives, and contractual choices," *European Economic Review* 44, 1057-1068.

[6] Fehr, E. and K. M. Schmidt (1999), "A Theory Of Fairness, Competition, And Cooperation," *Quarterly Journal of Economics* 114, 817-868.

[7] Foley, D. (1967), "Resource allocation and the public sector", *Yale Economic Essays* 7, 45-98.

[8] Grossman, J. S. and O. D. Hart (1983), "An Analysis of the Principal-Agent Problem," *Econometrica* 51, 7-45.

[9] Hellwing, M. F. and K. M. Schmidt (2002), "Discrete-Time Approximation of the Holmström and Milgrom Brownian Motion Model of Intertemporal Incentive Provision," *Econometrica* 70, 2225-2264.

[10] Holmström, B. and R. B. Myerson (1983), "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica* 51, 1799-1819.

[11] Holmström, B. and P. Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization* 7, 24-52.

[12] Holmström, B. and P. Milgrom (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica* 55, 303-328.

[13] Innes, R. D. (1990), "Limited Liability and Incentive Contracting with Ex-Ante Action Choices," *Journal of Economic Theory* 52, 45-67.

[14] Itoh, H. (2004), "Moral Hazard and Other-Regarding Preferences," *Japanese Economic Review*, 55, 18-45.

[15] Jewitt, I. (1988), "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica* 56, 1177-1190.

[16] Matthews, S. (2001), "Renegotiation of Moral Hazard Contracts under Limited Liability and Monotonicity," *Journal of Economic Theory* 97, 1-29.

[17] Neilson, S. W. and J. Stone (2010), "Piece-Rate Contracts for Other-Regarding Workers," *Economic Inquiry*, 48, 575-586.

[18] Pazner, E. and D. Schmeidler (1974), "A difficulty in the concept of fairness," *Review of Economic Studies* 41, 441-443.

[19] Pazner, E. and D. Schmeidler (1978a), "Egalitarian equivalent allocations: A new concept of economic equity," *Quarterly Journal of Economics* 92, 671-687.

[20] Rey-Biel, P. (2008), "Inequity Aversion and Team Incentives," *Scandinaivian Journal of Economics*, 110, 297-320.

[21] Rogerson, P. W. (1985), "The First-Order Approach to Principal-Agent Problems," *Econometrica* 53, 1357-1367.

[22] Thomson, W. (2010), *Fair Allocation Rules*, in Handbook of Social Choice and Welfare (K. Arrow, A. Sen, and K. Suzumura, eds), North-Holland, Amsterdam, New York.

[23] Varian, H. (1974), "Equity, envy, and efficiency", *Journal of Economic Theory* 9, 63-91.