CESifo WORKING PAPERS

6631
2017

August 2017

Adverse Selection and Moral Hazard with Multidimensional Types

Suehyun Kwon

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6631*
Category 11: Industrial Organisation

Adverse Selection and Moral Hazard with Multidimensional Types

Abstract

This paper studies a contracting problem where agents' cost of actions is private information. With two actions, this leads to a two-dimensional screening problem with moral hazard. There is a natural one-dimensional ordering of types when there is both adverse selection and moral hazard. Regardless of the number of types, an optimal menu of contracts either pools every type together or offers a menu of two contracts. Any incentive-compatible menu of contracts has to satisfy pairwise single-crossing properties in incentivized actions and ex-ante utilities. The principal can no longer sell the firm to the agent.

Keywords: adverse selection, moral hazard, multidimensional types.

Suehyun Kwon
Department of Economics
University College London
30 Gordon Street
United Kingdom - London WC1H 0AX
suehyun.kwon@ucl.ac.uk

March 21, 2017
First draft: March 1, 2017

1 Introduction

There are many examples where agents' cost of actions is their private information. If there are different tasks to be performed, each agent can have comparative advantage in certain tasks over the others, and the comparative advantage can be their private information. In taxation, if the government observes income but not the hourly wage, then the government doesn't know the number of hours the agent worked, and therefore the government doesn't know the cost of working for the agent.

In addition, in many of these environments, once an agent is hired, the principal may not necessarily observe the agent's action. The agent may have comparative advantage in certain actions, but if the principal doesn't observe which action the agent takes, then there is both adverse selection and moral hazard. I study a contracting problem with adverse selection and moral hazard where the agent's cost of action is his private information and the principal doesn't observe the agent's action.

When the cost of action is the agent's private information, having more than one action leads to multi-dimensional types. The cost of each action corresponds to one dimension of the agent's type. In this paper, I consider the case where there are two actions; each agent has a two-dimensional type. However, two-dimensional types already capture many aspects of comparative advantage. I don't restrict the cost of actions in any way. Each type is allowed to have a comparative advantage in either of the two actions, but also the difference in costs, which can be interpreted as the degree of comparative advantage, is not restricted. Furthermore, the level of cost is not restricted either. When one type has a comparative advantage in the first action over the second action, there could be another type with exactly the same difference in costs but has different levels of costs.

The first main result is that there is a natural one-dimensional ordering of two-dimensional types if there is both adverse selection and moral hazard. The ordering is determined by the cost of action for each type, and it doesn't require any of the standard assumptions. It doesn't rely on single-crossing property, monotone likelihood ratio property or monotone hazard rate. It doesn't even rely on the outcome distribution induced by each action nor the set of outcomes. The cost of action also doesn't need to be restricted in any way; it probably makes sense to assume they are non-negative, but the result doesn't rely on any functional form assumption either. This ordering is in terms of difference in costs of two actions. Since the difference in costs captures the comparative advantage of each type, the ordering of types shows that when there is both adverse selection and moral hazard, the comparative advantage is the key factor in determining the optimal menu of contracts.

The second main result is that an optimal contract either pools every type together or offers a menu of two contracts. In any incentive-compatible menu of contracts, the types are partitioned into one or two partitions, and there is a cutoff type in the one-dimensional ordering from above. The result further shows that between two partitions, which action is incentivized for which partition only depends on the one-dimensional ordering. The location of the cutoff and whether to pool every type together or not depends on the outcome distribution. However, as long as the principal incentivizes both actions for some types, the lower types in the one-dimensional ordering is incentivized to take the same action. This indicates that if there is both adverse selection and moral hazard, the comparative advantage of each type is the dominant feature in deciding which action should be incentivized for the type; the outcome distribution induced by the action doesn't affect incentive-compatible menu of contracts qualitatively. And same as in the first result, this result also doesn't rely on any of the standard assumptions or functional form assumptions. In an optimal menu of contracts, any types incentivized for the same action are pooled together in the same contract.

When an optimal menu of contracts pools several types together, these types are offered exactly the same compensation scheme and in addition incentivized to take the same action. This implies that the principal must leave rent to some of the types and he can no longer sell the firm to the agent. My results hold with risk neutral agent as well, and in particular, the residual claimant argument breaks down even with a risk neutral agent with no limited liability.

In addition, I also show that any incentive-compatible menu of contracts has to satisfy a pairwise single-crossing property in actions and a pairwise single-crossing property in utilities. These are weaker than standard single-crossing property in a sense that it is only between each pair of types and properties only hold for the incentivized actions and the utilities from the contracts for their own types. However, these are necessary conditions for every incentive-compatible menu of contracts. And more importantly, these are endogenous properties that a menu of contracts has to satisfy; these don't rely on any primitives of the model apart from additive separability of the utility function.

Compared to multi-dimensional screening problems, the combination of adverse selection and moral hazard leads to a one-dimensional ordering that hasn't been shown in the literature. In Laffont-Tirole (1987), the firm has private information about its cost, but the private information is one-dimensional. Often times, when a multi-dimensional screening problem reduces to a one-dimensional problem, one needs to make assumptions about the underlying environment. In my model, the only relevant assumptions are the combination of adverse selection and moral hazard where the private information is the cost type. This relies on the additive separability of the contract and the cost in the agent's utility function, but since any comparative advantage in cost leads to this type of utility function, it is still widely applicable. Furthermore, many taxation models embed additive separability or more precisely, the cost of working being private information as a standard informational friction. I can also derive the optimal contract without assuming any single crossing property which is different from standard adverse selection problems.

In addition, the presence of moral hazard leads to more pooling than in standard adverse selection problems. No matter how many different cost types there are, an optimal menu of contracts either pools every type together or offers only two contracts. This already indicates that if there are at least three types, then there is pooling of some types in every optimal menu of contracts. I also don't need to restrict the number of types in any way either.

Compared to moral hazard problems, pooling implies that an optimal menu of contracts doesn't always reflect agent's cost of actions. When the agent's cost of action is common knowledge, an optimal contract with moral hazard typically depends on the cost of action, whereas in this case, if several types are pooled together, all of these types get exactly the same compensation scheme despite having different costs. As mentioned already, residual claimant argument breaks down with a risk-neutral agent with no limited liability if there is both adverse selection and moral hazard. However, payments in an optimal contract is still a function of the likelihood ratio, and whether to offer one or two contracts depends on the outcome distributions. I provide a three-step procedure for finding the optimal menu of contract which is an analogue of two-step approach in Grossman-Hart (1983).

The fact that with no functional form assumptions, any optimal menu of contracts takes a simple form suggests that simplicity of contracts observed in the real world doesn't always have to be assumed or attributed to behavioral traits. The combination of adverse selection and moral hazard implies that because of incentive compatibility, the principal cannot offer any more contracts and incentivize different types of agents to self-select into different contracts.

The most closely related paper in the literature is Gottlieb-Moreira (2015). They identify an environment where the agent has private information about the outcome distribution and/or the cost, and there exists an optimal menu of contracts that offers one or two compensation schemes. However, there are a few key differences between their paper and mine. First of all, since most of my results hold for any incentive-compatible menu of contracts and not just for optimal menu of contracts, my paper should be considered more as an implementation paper rather than optimal mechanism design. Second, they assume the agent is risk neutral and has limited liability. Their main results all depend on the linearity and limited liability whereas I don't assume limited liability and I allow the agent to be risk averse. I show that residual claimant argument breaks down and the principal cannot extract full surplus even if the agent is risk neutral and has no limited liability. Third, in their model, different types of agents might take different actions even if they choose the same contract (compensation scheme). In my model, I show that there is a one-dimensional ordering of types, and the incentivized action is pinned down by the ordering up to the cutoff. There is no equivalent result in their paper, and furthermore, this one-dimensional ordering highlights the economics behind this class of models. The difference in costs, which captures comparative advantage of each type, really is the key driving force in determining an optimal menu of contracts. However, they allow for bigger class of types and action space, and they also allow for private information about outcome distribution if it is multiplicatively separable.

The rest of the paper is organized as follows. Section 2 describes the model, and the results are in section 3. Section 4 concludes.

2 Model

A principal hires an agent. If the agent chooses a contract from the menu of contracts offered by the principal, the agent takes an action and an outcome is realized. The agent's action is unobserved by the principal, and the principal only observes the outcome. Each contract specifies outcome-contingent payments.

The key feature of this model is that in addition to moral hazard, the cost of action is the agent's private information; there is both adverse selection and moral hazard with multi-dimensional types. There are $N \geq 2$ types, $\theta_1, \cdots, \theta_N$, and two actions a_1, a_2. Types differ by the cost of action $c(a_k|\theta_i)$ for all $i = 1, \cdots, N,\ k = 1, 2$, and the principal's prior at the beginning of the game is that the agent is θ_i with probability π_i. Assume $\pi_i > 0$ for all i. Each action a_k has a probability distribution F_k over outcomes $y \in \mathcal{Y} \subseteq \mathbb{R}$. I assume $f_k(y) > 0$ for all k, y.

The agent first learns his type, then the principal offers a menu of contracts. Contract i consists of $w_i(y)$ for all y. The rest of the timing is as explained above.

The principal is risk neutral with utility $y - w$ when outcome y is realized and he pays w. The agent can be risk neutral or risk averse with $u(w) - c(a|\theta)$ when type θ takes action a and gets paid w. u is strictly increasing, weakly concave and $u(0) = 0$. I don't make any assumptions about $c(\cdot|\cdot)$. If the principal doesn't offer a menu of contracts or the agent rejects all contracts, they each get their outside options 0.

3 Results

Following revelation principle, I assume the principal offers contract i for θ_i.[1] Suppose type i is incentivized to take $a(\theta_i)$. The set of all deviations of the agent is to choose contract j and take a_k.

$$-c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq -c(a_k|\theta_i) + \mathbb{E}[u(w_j(y))|a_k], \ \forall j, k \tag{1}$$

In particular, since the agent can always choose $a_k = a(\theta_i)$ after deviating to contract j, the following IC constraint is necessary:

$$\begin{aligned} & -c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq -c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_j(y))|a(\theta_i)], \ \forall j \\ \Rightarrow & \ \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq \mathbb{E}[u(w_j(y))|a(\theta_i)], \forall j. \end{aligned} \tag{2}$$

For given action, the expected payment is maximized with the contract that incentivizes that action.

This already implies that if types i, j are incentivized to take the same action, then their expected utility from payment for the action should be the same, and from the principal's perspective, the expected payment from each contract should also be the same. Therefore, it is without loss of generality to offer identical contracts to both types and pool them together.

Proposition 1. *If two types are incentivized to take the same action in an optimal menu of contracts, it is without loss of generality to offer identical contracts to both types and pool them together.*

[1] See Laffont-Martimort (2002).

We also know that

$$-c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq -c(a(\theta_j)|\theta_i) + \mathbb{E}[u(w_j(y))|a(\theta_j)], \ \forall j \quad (3)$$

(3) is a necessary set of IC constraints.

If every action has full support (including the ones not incentivized for any type), then (2) together with (1) implies (3), but it is not sufficient. However, if the principal can restrict the set of actions to the ones that are incentivized for at least one type (for example, when there is at least one outcome that is perfectly informative about the action and no limited liability) then (2) and (3) will be necessary and sufficient.

Proposition 2. *Suppose the agent can only take an action that is incentivized for at least one type. Double deviations in (contract, action) are not binding IC constraints, and* (2) *and* (3) *are necessary and sufficient for all IC constraints.*

(3) in addition implies

$$\begin{aligned}
&-c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq -c(a(\theta_j)|\theta_i) + \mathbb{E}[u(w_j(y))|a(\theta_j)], \\
&-c(a(\theta_j)|\theta_j) + \mathbb{E}[u(w_j(y))|a(\theta_j)] \geq -c(a(\theta_i)|\theta_j) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \\
\Rightarrow\ & c(a(\theta_i)|\theta_j) - c(a(\theta_j)|\theta_j) \geq \mathbb{E}[u(w_i(y))|a(\theta_i)] - \mathbb{E}[u(w_j(y))|a(\theta_j)] \geq c(a(\theta_i)|\theta_i) - c(a(\theta_j)|\theta_i) \\
\Rightarrow\ & c(a(\theta_i)|\theta_j) - c(a(\theta_j)|\theta_j) \geq c(a(\theta_i)|\theta_i) - c(a(\theta_j)|\theta_i) \qquad (4)
\end{aligned}$$

for any pair (i, j). This requires that the cost of action satisfies single-crossing type of property in any pair, but because this is only between the incentivized actions, $a(\theta_i), a(\theta_j)$, it is weaker than the usual definition of single-crossing property. I call this property pairwise single-crossing property of actions.

Proposition 3. (4) *is necessary for any incentive-compatible menu of contracts. It holds with equality if and only if IC constraints between contracts i and j are binding for both types i and j. When it binds, it is without loss of generality to pool types i, j together in an optimal menu of contracts.*

The benefit of proposition 3 is that it is only in terms of incentivized actions, and we don't need to worry about payments. As long as the menu of contracts is incentive-compatible, then incentivized actions for each type have to satisfy (4). And it holds with equality if and only if both types are indifferent between two contracts in which case the principal can pool them together in any optimal menu of contracts.

We can also characterize payments that satisfy (2). (2) is equivalent to

$$\int (u(w_i(y)) - u(w_j(y)))dF_i \geq 0,\ \forall i, j$$

which implies that

$$\begin{aligned}
&\int (u(w_i(y)) - u(w_j(y)))dF_i \geq 0,\ \forall i, j \\
&\int (u(w_j(y)) - u(w_i(y)))dF_j \geq 0,\ \forall i, j \\
\Rightarrow &\int (u(w_i(y)) - u(w_j(y)))dF_i \geq 0 \geq \int (u(w_i(y)) - u(w_j(y)))dF_j \qquad (5) \\
\Rightarrow &\int (u(w_i(y)) - u(w_j(y)))(dF_i - dF_j) \geq 0.
\end{aligned}$$

Incentive compatibility requires that ex-ante utility of the agent and the distribution of the outcome again have to satisfy pairwise single-crossing type of property. In this case as well, it's only between the contracts intended for a particular type. It's weaker than the usual single-crossing property, and because there is a restriction on the sign of the difference in two contracts for the same action, there is further restriction on the level of payments. I call this property pairwise single-crossing property of utilities.

Proposition 4. (5) *is necessary in any incentive-compatible contract.*

For the rest of the paper, assume without loss of generality

Assumption 1.

$$c(a_1|\theta_1) - c(a_2|\theta_1) \leq c(a_1|\theta_2) - c(a_2|\theta_2) \leq \cdots \leq c(a_1|\theta_N) - c(a_2|\theta_N).$$

Proposition 5. *Suppose Assumption 1 holds. In any incentive-compatible menu of contracts, there exists k such that $a(\theta_i) = a_1$ if and only if $i \leq k$. In* *any optimal menu of contracts, it is without loss of generality that any* θ_i *such that* $c(a_1|\theta_i) - c(a_2|\theta_i) = c(a_1|\theta_k) - c(a_2|\theta_k)$ *is incentivized to take* a_1*.*

Proposition 5 shows that even though types are two-dimensional, there is a natural one-dimensional ordering of types, and the cost difference between two actions is the only determinant of ordering of types. The exact location of the cutoff k depends on the outcome distributions induced by a_1, a_2, but the structure of optimal menu of contracts is uniquely pinned down up to k. The proof follows from proposition 3 and assumption 1. The fact that any types with the same cost difference are incentivized to take exactly the same action in any optimal menu of contracts shows that the contract is completely driven by the comparative advantage and the level of costs doesn't matter at all for the structure of optimal menu of contracts.

This is already different from other two-dimensional screening models because there is no restriction on the cost function, utility function or outcome distributions. I don't assume single-crossing property, monotone likelihood ratio property or monotone hazard rate. I also don't make any assumptions on the outcome distribution except full support, and results up to now don't depend on full support, either. As long as there is both adverse selection and moral hazard at the same time, there is a natural one-dimensional ordering. The only relevant assumption in addition to having both adverse selection and moral hazard at the same time is the additively separable utility.

Proposition 1 and last sentences in propositions 3 and 5 are properties of optimal menu of contracts, but all other results so far hold for any incentive-compatible menu of contracts. As long as a menu of contracts is incentive compatible, pairwise single-crossing property has to hold between any incentivized actions for two types and any ex-ante utility levels for two types. Furthermore, the pairwise single-crossing property of actions implies that with two actions, there is a natural one-dimensional ordering of types.

Given results so far, we can find an optimal menu of contracts in the following three-step procedure. First of all, from proposition 5, an optimal menu of contracts either incentivizes the same action for every type or partitions the types with a cutoff in the one-dimensional ordering. Second, from proposition

1, the principal can offer without loss of generality a pooling contract for all types or a menu of two contracts.

Proposition 6. *In any optimal menu of contracts, it is without loss of generality to pool every type together or offer a menu of two contracts.*

Proposition 6 implies that the residual claimant argument breaks down even with a risk-neutral agent with no limited liability, and the principal can no longer sell the firm. Conditions in proposition 7 are sufficient but not necessary, and they can be relaxed.

Proposition 7. *Suppose there are at least three types participating and the cost of action, $c(a_k|\theta_i)$ is distinct for all types i and actions a_k. In any optimal menu of contracts, some types get strictly positive rent. Even if the agent is risk neutral and has no limited liability, the principal cannot extract full surplus.*

The three-step procedure is as follows.

First Step: In the first step, we have to choose a cutoff $k \in \{1, \cdots, N\}$ or a pooling contract for all types.

Second Step: Consider the case of two contracts. If we fix the location of cutoff k, the full optimization problem is

$$\begin{aligned} &\max_{\{\tilde{w}_1(y)\},\{\tilde{w}_2(y)\}} \sum_{i\le k} \pi_i \mathbb{E}[y - \tilde{w}_1(y)|a_1] + \sum_{i>k} \pi_i \mathbb{E}[y - \tilde{w}_2(y)|a_2] \\ s.t. \; & -c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \ge -c(a_l|\theta_i) + \mathbb{E}[u(w_j(y))|a_l], \;\forall j, l \\ & -c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \ge 0 \end{aligned}$$

where $\{\tilde{w}_k(y)\}$ denotes the contract incentivizing action a_k and $\{w_i(y)\}$ denotes the contract for θ_i. Since there are two actions, and at least one type is incentivized to take each action, it follows from proposition 2 that the following set of IC constraints is necessary and sufficient for all IC constraints. For all i, j,

$$\begin{aligned} \mathbb{E}[u(w_i(y))|a(\theta_i)] &\ge \mathbb{E}[u(w_j(y))|a(\theta_i)], \\ -c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] &\ge -c(a(\theta_j)|\theta_i) + \mathbb{E}[u(w_j(y))|a(\theta_j)]. \end{aligned}$$

Let $u_i = \mathbb{E}[u(\tilde{w}_i(y))|a_i]$, then the IC constraints are equivalent to

$$\begin{aligned}
&\mathbb{E}[u(\tilde{w}_i(y))|a_i] \geq \mathbb{E}[u(\tilde{w}_j(y))|a_i], \ \forall i, j \\
&c(a_1|\theta_k) - c(a_2|\theta_k) \leq u_1 - u_2 \leq c(a_1|\theta_{k+1}) - c(a_2|\theta_{k+1})
\end{aligned}$$

and the optimization problem for given k is

$$\begin{aligned}
&\max_{\{\tilde{w}_1(y)\},\{\tilde{w}_2(y)\}} \sum_{i \leq k} \pi_i \mathbb{E}[y - \tilde{w}_1(y)|a_1] + \sum_{i>k} \pi_i \mathbb{E}[y - \tilde{w}_2(y)|a_2] \\
s.t.\ &\mathbb{E}[u(\tilde{w}_i(y))|a_i] \geq \mathbb{E}[u(\tilde{w}_j(y))|a_i], \ \forall i, j \\
&c(a_1|\theta_k) - c(a_2|\theta_k) \leq u_1 - u_2 \leq c(a_1|\theta_{k+1}) - c(a_2|\theta_{k+1}) \\
&- c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq 0
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
&\min \sum_{i \leq k} \pi_i \mathbb{E}[\tilde{w}_1(y)|a_1] + \sum_{i>k} \pi_i \mathbb{E}[\tilde{w}_2(y)|a_2] \\
s.t.\ &\mathbb{E}[u(\tilde{w}_1(y))|a_1] \geq \mathbb{E}[u(\tilde{w}_2(y))|a_1] \\
&\mathbb{E}[u(\tilde{w}_2(y))|a_2] \geq \mathbb{E}[u(\tilde{w}_1(y))|a_2] \\
&c(a_1|\theta_k) - c(a_2|\theta_k) \leq u_1 - u_2 \leq c(a_1|\theta_{k+1}) - c(a_2|\theta_{k+1}) \\
&- c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq 0
\end{aligned}$$

We can further simplify it as

$$\begin{aligned}
&\min \sum_{i \leq k} \pi_i \mathbb{E}[\tilde{w}_1(y)|a_1] + \sum_{i>k} \pi_i \mathbb{E}[\tilde{w}_2(y)|a_2] \\
s.t.\ &u_1 \geq \mathbb{E}[u(\tilde{w}_2(y))|a_1] \\
&u_2 \geq \mathbb{E}[u(\tilde{w}_1(y))|a_2] \\
&c(a_1|\theta_k) - c(a_2|\theta_k) \leq u_1 - u_2 \leq c(a_1|\theta_{k+1}) - c(a_2|\theta_{k+1}) \\
&u_1 \geq c(a_1|\theta_i) \ i \leq k \\
&u_2 \geq c(a_2|\theta_i) \ i > k
\end{aligned}$$

Third Step: If we fix u_1, u_2 in addition to k, the optimization problem is

$$\begin{aligned}
&\min \sum_{i\leq k} \pi_i \mathbb{E}[\tilde{w}_1(y)|a_1] + \sum_{i>k} \pi_i \mathbb{E}[\tilde{w}_2(y)|a_2] \\
s.t.\ & u_1 \geq \mathbb{E}[u(\tilde{w}_2(y))|a_1] \\
& u_2 \geq \mathbb{E}[u(\tilde{w}_1(y))|a_2] \\
& \int u(\tilde{w}_1(y))dF_1 = u_1 \\
& \int u(\tilde{w}_2(y))dF_2 = u_2.
\end{aligned}$$

With a slight abuse of notation, let $u_i(y) = u(\tilde{w}_i(y)),\ u^{-1} = h,\ p_k = \sum_{i\leq k} \pi_k$, then we have

$$\begin{aligned}
&\min p_k \int h(u_1(y))dF_1 + (1-p_k)\int h(u_2(y))dF_2 \\
s.t.\ & \int u_1(y) - u_2(y)dF_1 \geq 0 \geq \int u_1(y) - u_2(y)dF_2 \\
& \int u_1(y)dF_1 = u_1 \\
& \int u_2(y)dF_2 = u_2
\end{aligned}$$

In the Lagrangian, the first-order conditions for $u_1(y), u_2(y)$ are

$$\begin{aligned}
\partial u_1(y) &: p_k \frac{1}{u'(\tilde{w}_1(y))} f_1(y) = \lambda_1 f_1(y) - \lambda_2 f_2(y) - \mu_1 f_1(y) \\
\partial u_2(y) &: (1-p_k) \frac{1}{u'(\tilde{w}_2(y))} f_2(y) = -\lambda_1 f_1(y) + \lambda_2 f_2(y) - \mu_2 f_2(y) \\
& \lambda_1, \lambda_2, \mu_1, \mu_2 \geq 0
\end{aligned}$$

Since we have full support, the first-order conditions can be rewritten as

$$\begin{aligned}
\partial u_1(y) &: \frac{p_k}{u'(\tilde{w}_1(y))} = \lambda_1 - \lambda_2 \frac{f_2(y)}{f_1(y)} - \mu_1 \\
\partial u_2(y) &: \frac{1-p_k}{u'(\tilde{w}_2(y))} = -\lambda_1 \frac{f_1(y)}{f_2(y)} + \lambda_2 - \mu_2
\end{aligned}$$

$$\lambda_1, \lambda_2, \mu_1, \mu_2 \geq 0$$

If $\int u_1(y) - u_2(y)dF_1 > 0$ then $\lambda_1 = 0$, and if $0 > \int u_1(y) - u_2(y)dF_2$, then $\lambda_2 = 0$. Since u is increasing and $p_k \in (0, 1)$, if $\lambda_1 = 0$, then the right-hand side of the first condition is weakly negative while the left-hand side is strictly positive which is a contradiction. Likewise, if $\lambda_2 = 0$, then the left-hand side of the second condition is strictly positive while the right-hand side is weakly negative. We must have $\int u_1(y) - u_2(y)dF_1 = 0 = \int u_1(y) - u_2(y)dF_2$. This implies that all contracts provide exactly the same expected utility from payments for the given action, and (2) binds in any optimal menu of contracts. However, this is equivalence in the expected utility from payments, and compared to proposition 3, proposition 8 doesn't automatically imply that IC constraints are binding between two contracts.

Proposition 8. *In any optimal menu of two contracts,* (2) *binds for all* i, j*.*

Given the three steps, one can compute the expected payment given u_1, u_2, k, then choose optimal u_1, u_2 for each k and finally pick the cutoff or choose to offer a pooling contract for all types. This is an analogue of Grossman-Hart two-step approach in standard moral hazard problems. Instead of optimizing for a given action, I fixed the cutoff type, thereby fixing the incentivized action for every type, then optimized over payments.

4 Conclusion

I study a model of adverse selection and moral hazard where the agent's cost of action is his private information. When there is both adverse selection and moral hazard at the same time, the additively separable utility alone leads to a number of results that distinguish this class of models from any adverse selection or moral hazard problem on its own. I already mentioned in the introduction that any comparative advantage in cost has additively separable utility and cost of action being private information is a standard informational friction in many taxation models.

One of the main results is that there is a natural one-dimensional ordering of types despite agents having two-dimensional types. The ordering only depends on the cost of action and highlights that the difference in cost for two actions, which I interpret as the comparative advantage of each type, is the determining factor of an optimal menu of contracts. In fact, incentive compatibility alone already requires this one-dimensional ordering, and any incentive-compatible menu of contracts has to partition the one-dimensional ordering into one or two. The incentivized action for each partition is also pinned down by the one-dimensional ordering. Every incentive-compatible menu of contracts has exactly the same qualitative property, and nothing else about the model changes it.

Once we require optimality, then in any optimal menu of contracts, the principal can either pool every type together or offer a menu of two contracts. This is "without loss of generality" statement, and if the principal wanted to, he could offer more contracts incentivizing exactly the same action. However, types incentivized for the same action must be indifferent between all of these contracts, and there is no benefit for the principal to offer more contracts.

The main implication of the optimal menu of contracts is that offering a menu of one or two contracts is optimal regardless of the number of types. No matter how many types there are, what are their costs of action, utility functions, outcome distributions or any other aspect of the model, it is optimal to offer only one or two contracts. This is another illustration that simple contracts can be optimal with a fully rational agent, and simple contracts don't always have to be assumed in the contract space or attributed to behavioral traits. There are recent papers showing optimality of simple contracts with a behavioral agent, and I won't repeat the difference with Gottlieb-Moreira (2015), but my paper is closer to Holmström-Milgrom (1987) in spirit.

Furthermore, one might ask why offering one or two contracts is surprising when there are two actions. But in any moral hazard problem where the agent's cost of action is common knowledge, an optimal contract depends on the agent's cost of actions. Pooling several types together requires that these types are incentivized to take the same action, but they are also offered exactly the same compensation scheme despite having different costs. It follows that the principal can no longer sell the firm if there is both adverse selection and moral hazard. This holds even if the agent is risk neutral and there is no limited liability. In any standard moral hazard problem, if the agent is risk neutral and has no limited liability, the principal can make the agent residual claimant and extract all the surplus. This is no longer the case if the agent has private information.

Lastly, I also identify pairwise single-crossing properties between any pair for incentivized actions and ex-ante utilities from payments. These are necessary conditions for any incentive-compatible menu of contracts, but there are two key differences from standard single crossing property. First of all, it's only between a pair and not for all actions or contracts. But second, more crucially, this is an endogenous property that has to be satisfied by a menu of contracts, and it has nothing to do with the primitives of the model. The usual single crossing property involves the utility of the agent, but in my case, as long as the utility is additively separable, there is no restriction on the model, and a menu of contracts, when the principal designs it, has to satisfy these propreties to be incentive compatible.

References

[1] Asseyer, Andreas. "Optimal monitoring in dynamic procurement contracts." (2016) https://sites.google.com/site/andreasasseyer/research

[2] Carroll, Gabriel. "Robustness and linear contracts." *American Economic Review* 105.2 (2015): 536-563.

[3] Chassang, Sylvain. "Calibrated incentive contracts." *Econometrica* 81.5 (2013): 1935-1971.

[4] Gottlieb, Daniel, and Humberto Moreira. "Simple Contracts with Adverse Selection and Moral Hazard." (2015) https://ssrn.com/abstract=2568271

[5] Grossman, Sanford J., and Oliver D. Hart. "An analysis of the principal-agent problem." *Econometrica* 51.1 (1983): 7-45.

[6] Holmström, Bengt, and Paul Milgrom. "Aggregation and linearity in the provision of intertemporal incentives." *Econometrica* 55.2 (1987): 303-328.

[7] Laffont, Jean-Jacques, and David Martimort. "The Theory of Incentives: The Principal-Agent Model." (2002)

[8] Laffont, Jean-Jacques, and Jean Tirole. "Auctioning Incentive Contracts." *Journal of Political Economy* 95.5 (1987): 921-937.

[9] Milgrom, Paul, and Chris Shannon. "Monotone comparative statics." *Econometrica* 62.1 (1994): 157-180.

[10] Stantcheva, Stefanie. "Optimal taxation and human capital policies over the life cycle." *forthcoming in Journal of Political Economy*

A Proofs

Proof of Proposition 1. Denote $a(\theta_i)$ to be the incentivized action for type i. The set of all deviations of type i is to choose contract j and take a_k.

$$-c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq -c(a_k|\theta_i) + \mathbb{E}[u(w_j(y))|a_k], \ \forall j, k \quad (6)$$

In particular, since the agent can always choose $a_k = a(\theta_i)$ after deviating to contract j, the following IC constraint is necessary:

$$-c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq -c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_j(y))|a(\theta_i)], \ \forall j$$
$$\Rightarrow \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq \mathbb{E}[u(w_j(y))|a(\theta_i)], \forall j.$$

If types i, j are incentivized to take the same action a^*, then we have

$$\mathbb{E}[u(w_i(y))|a^*] \geq \mathbb{E}[u(w_j(y))|a^*]$$

$$
\begin{aligned}
&\mathbb{E}[u(w_j(y))|a^*] \geq \mathbb{E}[u(w_i(y))|a^*] \\
\Rightarrow\ &\mathbb{E}[u(w_i(y))|a^*] = \mathbb{E}[u(w_j(y))|a^*].
\end{aligned}
$$

The principal can choose the one that minimizes his expected payment. Since in the original menu of contracts, all IC constraints were already satisfied and types i, j are indifferent between two contracts incentivized for them, it is without loss of generality to pool types i, j together. □

Proof of Proposition 2. Consider

$$
-c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq -c(a_k|\theta_i) + \mathbb{E}[u(w_j(y))|a_k],\ \forall j, k \tag{7}
$$

If every action is incentivized for at least one type, when the agent deviates to a_k, there exists some type l with $a_k = a(\theta_l)$. We already know that

$$
\mathbb{E}[u(w_l(y))|a(\theta_l)] \geq \mathbb{E}[u(w_j(y))|a(\theta_l)]\ \forall j
$$

is necessary and the right-hand side of the agent's IC constraint is bounded from above by

$$
\begin{aligned}
& -c(a_k|\theta_i) + \mathbb{E}[u(w_j(y))|a_k] \\
= & -c(a(\theta_l)|\theta_i) + \mathbb{E}[u(w_j(y))|a(\theta_l)] \\
\leq & -c(a(\theta_l)|\theta_i) + \mathbb{E}[u(w_l(y))|a(\theta_l)].
\end{aligned}
$$

Therefore, the maximum deviation payoff the agent can get when he deviates to $a(\theta_l)$ is also to choose contract for θ_l, and double deviations are not binding IC constraints. This already shows that (2) and (3) are necessary. Sufficiency of (2) and (3) for all IC constraints follows from

$$
\begin{aligned}
-c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] &\geq -c(a(\theta_l)|\theta_i) + \mathbb{E}[u(w_l(y))|a(\theta_l)],\ \forall l \\
&\geq -c(a(\theta_l)|\theta_i) + \mathbb{E}[u(w_j(y))|a(\theta_l)],\ \forall j, l
\end{aligned}
$$

and for any a_k, there exists l such that $a_k = a(\theta_l)$. □

Proof of Proposition 3. (3) implies

$$
\begin{aligned}
& -c(a(\theta_i)|\theta_i) + \mathbb{E}[u(w_i(y))|a(\theta_i)] \geq -c(a(\theta_j)|\theta_i) + \mathbb{E}[u(w_j(y))|a(\theta_j)], \\
& -c(a(\theta_j)|\theta_j) + \mathbb{E}[u(w_j(y))|a(\theta_j)] \geq -c(a(\theta_i)|\theta_j) + \mathbb{E}[u(w_i(y))|a(\theta_i) \\
\Rightarrow\ & c(a(\theta_i)|\theta_j) - c(a(\theta_j)|\theta_j) \geq \mathbb{E}[u(w_i(y))|a(\theta_i)] - \mathbb{E}[u(w_j(y))|a(\theta_j)] \geq c(a(\theta_i)|\theta_i) - c(a(\theta_j)|\theta_i) \\
\Rightarrow\ & c(a(\theta_i)|\theta_j) - c(a(\theta_j)|\theta_j) \geq c(a(\theta_i)|\theta_i) - c(a(\theta_j)|\theta_i)
\end{aligned}
$$

for any pair (i, j). The last inequality binds if and only if all inequality bind, and types i, j are indifferent between two contracts. Since IC constraints of all other types or any IC constraints for i, j into different actions and contracts have to be satisfied anyway, the principal can offer only one of contracts i, j without violating IC constraints. In any optimal menu of contracts, the principal can choose the contract that maximizes $\mathbb{E}[y - w(y)|a]$. □

Proof of Proposition 4. (2) is necessary for incentive compatibility and is equivalent to

$$\int (u(w_i(y)) - u(w_j(y)))dF_i \geq 0,\ \forall i, j$$

which implies that

$$
\begin{aligned}
& \int (u(w_i(y)) - u(w_j(y)))dF_i \geq 0,\ \forall i, j \\
& \int (u(w_j(y)) - u(w_i(y)))dF_j \geq 0,\ \forall i, j \\
\Rightarrow\ & \int (u(w_i(y)) - u(w_j(y)))dF_i \geq 0 \geq \int (u(w_i(y)) - u(w_j(y)))dF_j \\
\Rightarrow\ & \int (u(w_i(y)) - u(w_j(y)))(dF_i - dF_j) \geq 0.
\end{aligned}
$$

□

Proof of Proposition 5. (4) requires

$$c(a(\theta_i)|\theta_j) - c(a(\theta_j)|\theta_j) \geq c(a(\theta_i)|\theta_i) - c(a(\theta_j)|\theta_i)\ \forall i, j.$$

If $a(\theta_i) = a_1,\ a(\theta_j) = a_2$, then we need

$$c(a_1|\theta_j) - c(a_2|\theta_j) \geq c(a_1|\theta_i) - c(a_2|\theta_i)\ \forall i, j.$$

Together with assumption 1, we must have $j \geq i$. Therefore, in any incentive-compatible menu of contracts, types are partitioned into two by the one-dimensional ordering in assumption 1. From proposition 3, any type with the same cost difference as the cutoff type is incentivized for the same action. □

Proof of Proposition 6. Propositions 1 and 5 imply that it is without loss of generality to offer one or two contracts in any optimal menu of contracts. □

Proof of Proposition 7. When there are at least three types, there exist two types incentivized for the same contract and action. Suppose types i, j take contract $w_1(y)$ and action a_1. The ex-ante utility of type i is $\mathbb{E}[u(w_1(y))|a_1] - c(a_1|\theta_i) \neq \mathbb{E}[u(w_1(y))|a_1] - c(a_1|\theta_j)$. Since the IR constraints for both types have to be satisfied, at least one of the two has to get strictly positive rent. If they take contract $w_2(y)$ and action a_2, the ex-ante utility of type i is $\mathbb{E}[u(w_2(y))|a_2] - c(a_2|\theta_i) \neq \mathbb{E}[u(w_2(y))|a_2] - c(a_2|\theta_j)$. Since the IR constraints for both types have to be satisfied, at least one of the two has to get strictly positive rent. Therefore, the principal can never extract full surplus, and this doesn't depend on risk preferences of the agent or limited liability. □

Proof of Proposition 8. This is already derived in the main body of the text but replicated here for completeness.

In the third step of the three-step procedure, if we fix u_1, u_2 in addition to k, the optimization problem is

$$\begin{aligned}
&\min \sum_{i \leq k} \pi_i \mathbb{E}[\tilde{w}_1(y)|a_1] + \sum_{i > k} \pi_i \mathbb{E}[\tilde{w}_2(y)|a_2] \\
s.t.\ & u_1 \geq \mathbb{E}[u(\tilde{w}_2(y))|a_1] \\
& u_2 \geq \mathbb{E}[u(\tilde{w}_1(y))|a_2] \\
& \int u(\tilde{w}_1(y)) dF_1 = u_1
\end{aligned}$$

$$\int u(\tilde{w}_2(y))dF_2 = u_2.$$

With a slight abuse of notation, let $u_i(y) = u(\tilde{w}_i(y)),\ u^{-1} = h,\ p_k = \sum_{i\leq k} \pi_k$, then we have

$$\begin{aligned}
&\min p_k \int h(u_1(y))dF_1 + (1-p_k)\int h(u_2(y))dF_2 \\
s.t.\ &\int u_1(y) - u_2(y)dF_1 \geq 0 \geq \int u_1(y) - u_2(y)dF_2 \\
&\int u_1(y)dF_1 = u_1 \\
&\int u_2(y)dF_2 = u_2
\end{aligned}$$

In the Lagrangian, the first-order conditions for $u_1(y), u_2(y)$ are

$$\begin{aligned}
&\partial u_1(y) : p_k \frac{1}{u'(\tilde{w}_1(y))} f_1(y) = \lambda_1 f_1(y) - \lambda_2 f_2(y) - \mu_1 f_1(y) \\
&\partial u_2(y) : (1-p_k)\frac{1}{u'(\tilde{w}_2(y))} f_2(y) = -\lambda_1 f_1(y) + \lambda_2 f_2(y) - \mu_2 f_2(y) \\
&\lambda_1, \lambda_2, \mu_1, \mu_2 \geq 0
\end{aligned}$$

Since we have full support, the first-order conditions can be rewritten as

$$\begin{aligned}
&\partial u_1(y) : \frac{p_k}{u'(\tilde{w}_1(y))} = \lambda_1 - \lambda_2 \frac{f_2(y)}{f_1(y)} - \mu_1 \\
&\partial u_2(y) : \frac{1-p_k}{u'(\tilde{w}_2(y))} = -\lambda_1 \frac{f_1(y)}{f_2(y)} + \lambda_2 - \mu_2 \\
&\lambda_1, \lambda_2, \mu_1, \mu_2 \geq 0
\end{aligned}$$

If $\int u_1(y) - u_2(y)dF_1 > 0$ then $\lambda_1 = 0$, and if $0 > \int u_1(y) - u_2(y)dF_2$, then $\lambda_2 = 0$. Since u is increasing and $p_k \in (0,1)$, if $\lambda_1 = 0$, then the right-hand side of the first condition is weakly negative while the left-hand side is strictly positive which is a contradiction. Likewise, if $\lambda_2 = 0$, then the left-hand side of the second condition is strictly positive while the right-hand side is weakly negative. We must have $\int u_1(y) - u_2(y)dF_1 = 0 = \int u_1(y) -$

$u_2(y)dF_2$. This implies that all contracts provide exactly the same expected utility from payments for the given action, and (2) binds in any optimal menu of contracts. $\square$