The Mathematics of Principal-Agent Problems

by

Bibo Liu

BSc, Huazhong Normal University, 1999

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

University of Victoria

The Mathematics of Principal-Agent Problems

by

Bibo Liu
B.Sc., Huazhong Normal University, 1999.

Supervisory Committee

Dr. Jane Ye, Supervisor
(Department of Mathematics & Statistics)

Dr. Martial Agueh, Departmental Member
(Department of Mathematics & Statistics)

Dr. Bill Reed, Departmental Member
(Department of Mathematics & Statistics)

Dr. Paul Schure, Outside Examiner
(Department of Economics)

Dr. David Scoones, External Examiner
(Department of Economics)

Supervisory Committee
Dr. Jane Ye, Supervisor
(Department of Mathematics & Statistics)

Dr. Martial Agueh, Departmental Member
(Department of Mathematics & Statistics)

Dr. Bill Reed, Departmental Member
(Department of Mathematics & Statistics)

Dr. Paul Schure, Outside Examiner
(Department of Economics)

Dr. David Scoones, External Examiner
(Department of Economics)

ABSTRACT

The principal-agent problem is an important model in the field of Economics of Information. In this thesis we study only a particular type of principal-agent problem which is called moral hazard model and by the principal-agent problem we mean it is moral hazard model. The moral hazard model actually belongs to the class of bilevel programming problems in Mathematics. In Economics, the first order approach is used to reduce the principal-agent problem to a single level optimization problem. However, this approach is only valid under some strong conditions. Moreover the approach can only be used under the assumption that the optimal action of the principal-agent problem and its relaxed problem appears only at an interior point. In this thesis, we consider a new relaxed problem. Under more general assumptions, we can solve the principal-agent problem without restricting the optimal action of the agent to be in the interior.

Table of Contents

List of Figures

List of Abbreviations

KKT conditions	Karush-Kuhn-Tucker necessary optimality conditions
LICQ	linear independence constraint qualification
MFCQ	Mangasarian-Fromovitz constraint qualification
MPEC	Mathematical program with equilibrium constraints
MPEC LICQ	MPEC linear independence constraint qualification
MPEC GMFCQ	MPEC generalized Mangasarian-Fromovitz constraint qualification
PA problem	Principal-agent problem
RPA	Relaxed principal-agent problem
MLRC	Monotone likelihood ratio condition
SDC	Stochastic dominance condition
CDFC	Convexity of the distribution function condition
DRPA	Double-relaxed principal-agent program
PA-MPEC	Principal-agent MPEC
W-stationary	Weakly stationary
M-stationary	Mordukhovich stationary
S-stationary	Strong stationary
PA-MPEC LICQ	Principal-agent MPEC linear independence constraint qualification
PA-MPEC GMFCQ	Principal-agent MPEC generalized Mangasarian-Fromovitz constraint qualification
k-MLRC	k-monotone likelihood ratio condition
k-CDFC	k-convexity of the distribution function condition
k-OCDFC	k-opposite convexity of the distribution function condition

Acknowledgment

I would like to thank the University of Victoria, Department of Mathematics and Statistics for providing a place of employment and inspiration.

I wish to thank my supervisor Dr. Jane Ye for her patience, helpful advice and supervision.

I also would like to thank Dr. Paul Schure for his help, support and patience.

Bibo Liu

Chapter 1

Introduction

In this thesis, we study the moral hazard model in the principal-agent problems. This problem arises when a principal hires an agent. The principal offers a contract to the agent. The agent will accept the offer if and only if it maximizes his payoff and gives him a payoff that is not smaller than the minimal acceptable payoff as well. The agent takes an unobservable action from a set A to maximize his payoff that affects the principal's payoff as well. The principal now chooses a contract which is acceptable to the agent so as to maximize his payoff. We call this problem the principal-agent problem and we only study this type of principal-agent problem in this thesis.

The mathematical format of the principal-agent problem is a bilevel programming problem which is difficult to treat. The economists relax the principal-agent problem by replacing the set of maximum efforts by the set of stationary points of the agent's problem. Since stationary points may be maximum, minimum or saddle points, the relaxed problem is not equivalent to the original problem. Finding the conditions under which the relaxed problem is equivalent to the original problem is an interesting topic in Economics. A proper way is to prove that the agent's payoff function is concave in his action with the so-called monotone likelihood ratio condition (MLRC) and the convexity of the distribution function condition (CDFC). Under the concavity of the agent's payoff function in the action variable, a stationary point must be a maximum point. Since a stationary point in the first order approach must be an interior point, it can not handle the case when the maximum effort appears at the boundary of A.

In order to solve the principal-agent problem when the optimal action lies either in the interior or at the boundary of A, we use the Karush-Kuhn-Tucker necessary optimality conditions to replace the agent's maximization problem. If the agent's payoff function is concave in his action, then a optimal solution to the new relaxed problem is a optimal solution of the principal-agent problem. To get the concavity of the agent's payoff function, we find new conditions which are more general than the monotone likelihood ratio condition and the convexity of the distribution function condition. Under these conditions, we can obtain the equivalence between the principal-agent problem and the new relaxed problem in both the cases when the optimal actions are in the interior and at the boundary of A.

Chapter 2 provides theories from nonlinear programming and mathematical program with equilibrium constraints. Chapter 3 introduces the principal-agent problem that we study in this thesis. Chapter 4 discusses when the first-order approach can be valid in Economics and rewrites the proof for easier reading. Chapter 5 uses our new approach to study the principal-agent problem and finds the sufficient conditions for our approach to be valid. Chapter 6 reviews all the results in the thesis.

Chapter 2

Preliminaries

The aim of this Chapter is to provide some definitions for later use in Chapter 5.

2.1 Standard nonlinear programming

In this section, we review the Karush-Kuhn-Tucker necessary optimality conditions and give the definitions of the constraint qualifications which can be found in standard references for nonlinear programming such as [2, 11]. Two constraint qualifications are introduced here, namely the linear independence constraint qualification and the Mangasarian-Fromovitz constraint qualification. The linear independence constraint qualification is simple and used very often, but the Mangasarian-Fromovitz constraint qualification is a weaker condition than the linear independence constraint qualification.

Let $f : R^N \to R$, $g_i : R^N \to R, i = 1, \cdots, m$, $h_j : R^N \to R, j = 1, \cdots, q$. We consider the following nonlinear programming problem P:

$$
\begin{array}{llll}
\text{(P)} & \min & f(z) & \\
& s.t. & g_i(z) \leq 0, & i = 1, \cdots, m, \\
& & h_j(z) = 0, & j = 1, \cdots, q.
\end{array}
$$

Let z^* be a feasible solution and $I(z^*) = \{i : g_i(z^*) = 0\}$ be the set of binding constraints. We assume that f, g_i and h_j are continuously differentiable at z^*.

Definition 2.1 (KKT conditions) *Let the point z^* be a feasible solution of the problem P. We say that Karush-Kuhn-Tucker (KKT) necessary optimality conditions hold at z^* if there exist Lagrange multipliers $\lambda_1, \cdots, \lambda_m$, $\mu_1, \cdots, \mu_q$ such that*

$$
\begin{aligned}
0 &= \nabla f(z^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(z^*) + \sum_{j=1}^{q} \mu_j \nabla h_j(z^*), \\
& \lambda_i \geq 0, \lambda_i g_i(z^*) = 0, i = 1, 2, \cdots, m,
\end{aligned}
$$

where ∇f denotes the gradient of f.

In 1951, Kuhn-Tucker gave the following example in [8] to point out that KKT conditions may not hold at a local optimal solution.

Example 2.1

$$
\begin{aligned}
\min \quad & -x \\
s.t. \quad & (x-1)^3 + y \leq 0, \\
& y \geq 0.
\end{aligned}
$$

Solution. The optimal solution is $(x^*, y^*) = (1, 0)$ showing on Figure 2.1.

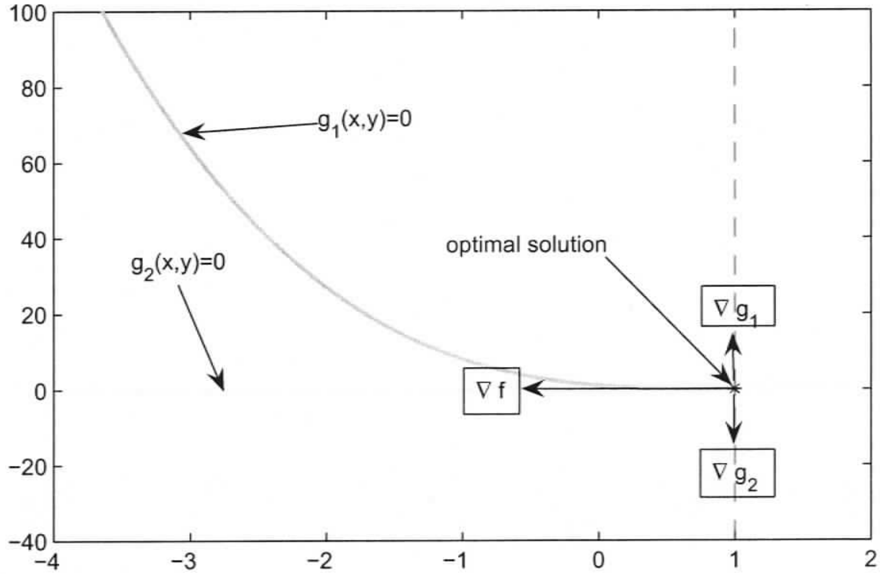

Figure 2.1: The optimal solution of Example 2.1

Here is the details about why KKT conditions can not be satisfied at the optimal solution for the above example.

Because $\nabla g_1(x^*, y^*) = (0, 1)$ and $\nabla g_2(x^*, y^*) = (0, -1)$, $\nabla g_1(x^*, y^*) + \nabla g_2(x^*, y^*) = 0$. Since $\nabla f(x^*, y^*) = (-1, 0)$, there is no λ_1, λ_2 satisfying $\nabla f(x^*, y^*) + \lambda_1 \nabla g_1(x^*, y^*) + \lambda_2 \nabla g_2(x^*, y^*) = 0$.

We now state sufficient conditions for the Karush-Kuhn-Tucker necessary optimality conditions to hold.

Definition 2.2 (LICQ) *Let the point z^* be a feasible solution of the problem P. We say that the linear independence constraint qualification (LICQ) holds at z^* if the set of gradients of the active constraints $\nabla g_i(z^*), i \in I(z^*)$ and $\nabla h_j(z^*), j = 1, \cdots, q$ are linearly independent.*

Definition 2.3 (MFCQ) *Let the point z^* be a feasible solution of the problem P. We say that the Mangasarian-Fromovitz constraint qualification (MFCQ) holds at z^* if $\nabla h_j(z^*), j = 1, \cdots, q$ are linearly independent and there exists a vector $d \in R^N$ such that*

$$\nabla g_i(z^*)^\top d < 0, \forall i \in I(z^*),$$

$$\nabla h_j(z^*)^\top d = 0, j = 1, \cdots, q,$$

where $^\top$ indicates the transpose. It can be shown that MFCQ is equivalent to the following condition:

$$\left.\begin{array}{l} 0 = \sum_{i=1}^m \lambda_i \nabla g_i(z^*) + \sum_{j=1}^q \mu_j \nabla h_j(z^*) \\ \lambda_i \geq 0, \lambda_i g_i(z^*) = 0 \end{array}\right\} \Rightarrow \lambda_i = 0, \mu_j = 0, \forall i, j.$$

When there is no inequality constraint, MFCQ is equivalent to saying that the gradient vectors are linearly independent. Actually, MFCQ is equivalent to saying that the Fritz-John conditions (see [2, 11]) do not hold if the multiplies for the objective function is zero.

Theorem 2.1 (see e.g. [2]) *Let the point z^* be an optimal solution of the problem P. If all binding constraint functions are linearly independent or MFCQ holds at z^*, then the Karush-Kuhn-Tucker necessary optimality conditions hold at z^*.*

Now we explain the relation between LICQ and MFCQ.

LICQ means that we can not find scalars $\lambda_i (i \in I(z^*)), \mu_j (j = 1, \cdots, q)$ which are not all zero such that $0 = \sum_{i \in I(z^*)} \lambda_i \nabla g_i(z^*) + \sum_{j=1}^q \mu_j \nabla h_j(z^*)$. MFCQ means that we can not find nonnegative scalars $\lambda_i (i \in I(z^*))$ and scalars $\mu_j (j = 1, \cdots, q)$ which are not all zero such that

$$0 = \sum_{i \in I(z^*)} \lambda_i \nabla g_i(z^*) + \sum_{j=1}^{q} \mu_j \nabla h_j(z^*).$$

Since scalars $\lambda_i (i \in I(z^*))$ are only required to be nonnegative, MFCQ is weaker then LICQ. We can see this from the following example.

Example 2.2 *Let* $z^* = (x^*, y^*) = (0, 0), g_1(z) = x + y, g_2(z) = 2x + 2y.$

Solution. Since $\nabla g_1(z^*) = (1, 1), \nabla g_2(z^*) = (2, 2)$, there exist $\lambda_1 = -2, \lambda_2 = 1$ such that $\nabla g_2(z^*) - 2\nabla g_1(z^*) = 0$. LICQ can not be satisfied here, but MFCQ can be satisfied because $\lambda_1 = -2 < 0$.

2.2 Mathematical program with equilibrium constraints

2.2.1 Introduction

The agent's maximization problem can be written as follows:

$$\begin{aligned} \max \quad & V(w, a) \\ \text{s.t.} \quad & g_1(w, a) = a - \overline{a} \le 0, \\ & g_2(w, a) = \underline{a} - a \le 0, \end{aligned}$$

where $V(w, a)$ is the payoff of the agent when the principal offers a contract w and the action a is taken.

Since at most one of the two constraints is binding, LICQ holds at the optimal solution a^* of the agent's maximization problem when the principal offers a contract w^*. By Theorem 2.1, there exist $b_1^* \geq 0, b_2^* \geq 0$ s.t.

$$
\begin{aligned}
&V_a(w^*, a^*) - b_1^* + b_2^* = 0,\\
&b_1^*(a^* - \underline{a}) = 0,\\
&b_2^*(\overline{a} - a^*) = 0.
\end{aligned}
$$

Therefore, we may consider the principal-agent problem as a special case of the following mathematical program with equilibrium constraints (MPEC) in the standard form:

$$
\begin{aligned}
\text{(MPEC)} \quad \min \quad & f(z)\\
\text{s.t.} \quad & g(z) \leq 0, \quad h(z) = 0,\\
& B(z) \geq 0, \quad A(z) \geq 0, \quad B(z)^\top A(z) = 0,
\end{aligned}
$$

where $f : R^n \to R$, $B : R^n \to R^m$, $A : R^n \to R^m$, $g : R^n \to R^p$, $h : R^n \to R^q$ and $^\top$ indicates the transpose. The reader is referred to [10, 19] for applications and recent developments for the standard MPEC. Most part of this section is quoted from the comprehensive reference [27]. The reader can also find some different formulations of MPEC which are equivalent to the standard MPEC in Ye [27]. Since there are several different approaches to reformulate MPEC, various stationary concepts arise (see e.g. [22], [27]). There are three kinds of stationary points in Definition 2.4

which are weakly stationary (W-stationary), Mordukhovich stationary (M-stationary) and strong stationary (S-stationary). The name "M-stationary condition" was first used in [22]. However, the definition of M-stationary was first introduced in Ye [28, Theorem 3.2] by using Mordukhovich coderivative of set-valued maps (see [16]) and further studied by Ye in [26] and Outrata in [18]. More information about S-stationary can be found in [10, 25]. We also introduced two different constraint qualifications for stationary conditions to hold for the MPEC problems. One is the MPEC linear independence constraint qualification (MPEC LICQ) under which all three stationary conditions hold at a local optimal solution of MPEC. The other one is the MPEC generalized Mangasarian-Fromovitz constraint qualification (MPEC GMFCQ) under which both M-stationary and W-stationary conditions hold at a local optimal solution of MPEC. We add more examples here to explain why we need the constraint qualifications and the relation between the two constraint qualifications.

Given a feasible vector z^* of MPEC, we define the following index sets:

$$
\begin{aligned}
&I_g := \{i : g_i(z^*) = 0\},\\
&\alpha := \alpha(z^*) := \{i : B_i(z^*) = 0, A_i(z^*) > 0\},\\
&\beta := \beta(z^*) := \{i : B_i(z^*) = 0, A_i(z^*) = 0\},\\
&\gamma := \gamma(z^*) := \{i : B_i(z^*) > 0, A_i(z^*) = 0\}.
\end{aligned}
$$

In deriving Kuhn-Tucker-type necessary optimality conditions, however, we need to find the constraint qualifications. Unfortunately, the standard constraint qualifications such as the MFCQ never hold for the MPEC problems. A proposition quoted from the reference [29] is a precise statement of this fact. For completeness, we write the proposition in a different format as follows:

Proposition 2.2 *Let $z^* = (w^*, a^*, b_1^*, b_2^*)$ be a feasible point of MPEC. Then MFCQ does not hold at z^*.*

Proof. For simplicity, we omit the constraints $g(z) \leq 0, \;\; h(z) = 0$, in the proof. Set

$$H(z) := \langle A(z), B(z) \rangle,$$

then

$$A_i(z) = 0, \;\; i \in \gamma \cup \beta,$$

$$B_i(z) = 0, \;\; i \in \alpha \cup \beta,$$

$$H_i(z) = 0$$

are active constraints for MPEC at z^*.

If $\nabla H(z^*) = 0$, then $\nabla H(z^*)$ can not satisfy the linear independence condition. If $\nabla H(z^*) = \sum_{i \in \gamma} \nabla A_i(z^*) B_i(z^*) + \sum_{i \in \alpha} A_i(z^*) \nabla B_i(z^*) \neq 0$ with $B_i(z^*) > 0, i \in \gamma, A_i(z^*) > 0, i \in \alpha$ and there exists a vector $v \in R^n$ such that

$$\langle v, \nabla A_i(z^*) \rangle > 0, \;\; \forall i \in \gamma \cup \beta$$

and

$$\langle v, \nabla B_i(z^*) \rangle > 0, \;\; \forall i \in \alpha \cup \beta,$$

then

$$\begin{aligned}\langle v, \nabla H(z^*)\rangle &= \langle v, \sum_{i=1}^{m}(\nabla A_i(z^*)B_i(z^*) + A_i(z^*)\nabla B_i(z^*))\rangle \\ &= \sum_{i=1}^{m}\Big[\langle v, \nabla A_i(z^*)\rangle B_i(z^*) + \langle v, \nabla B_i(z^*)\rangle A_i(z^*)\Big] \\ &= \sum_{i\in\gamma}\langle v, \nabla A_i(z^*)\rangle B_i(z^*) + \sum_{i\in\alpha}\langle v, \nabla B_i(z^*)\rangle A_i(z^*) > 0.\end{aligned}$$

MFCQ requires that $\langle v, \nabla H(z^*)\rangle = 0$, so MFCQ cannot hold at z^*.

□

The following example shows that MFCQ does not hold when $\nabla H(z^*) = 0$.

Example 2.3 ([9])

$$\begin{aligned} \min \quad & -y \\ s.t. \quad & x - y = 0, \\ & x \geq 0, y \geq 0, \\ & xy = 0. \end{aligned}$$

Solution. The optimal solution is $(x^*, y^*) = (0, 0)$ showing on Figure 2.2 which is the only feasible point.

LICQ and MFCQ can not be satisfied for the above example. The details are as follows:

Because $x \geq 0, y \geq 0$ are both binding, LICQ requires that the gradients of the four functions $g_1 = -x$, $g_2 = y$, $h_1 = x - y$, $h_2 = xy$ are linear independent.

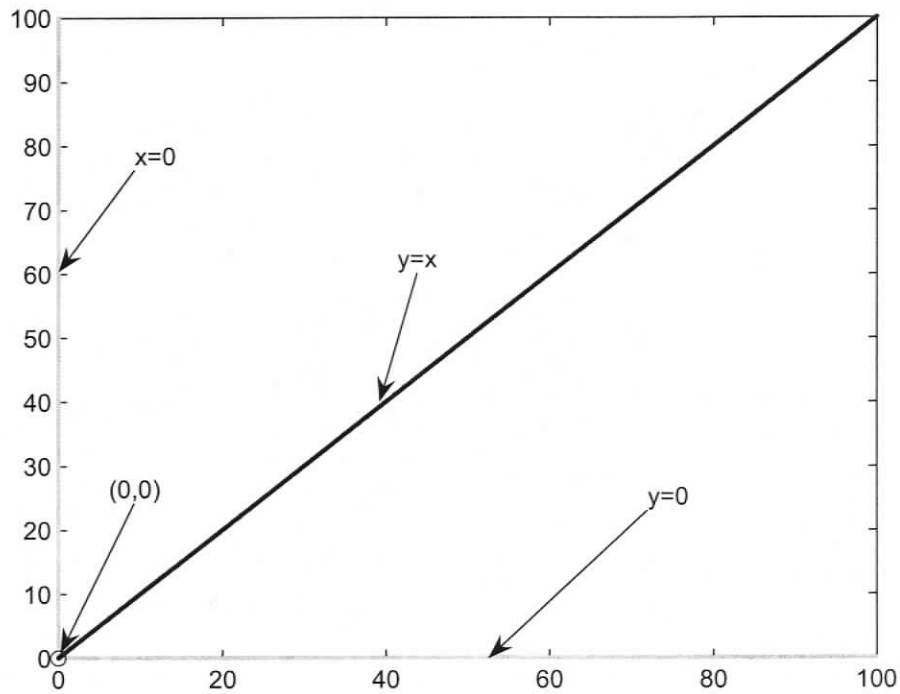

Figure 2.2: The optimal solution of Example 2.3

MFCQ requires that the gradients of the two functions $h_1 = x - y$, $h_2 = xy$ are linear independent. Since $\nabla h_2(x^*, y^*) = (0,0)^T$, LICQ and MFCQ can not be satisfied here.

The following example shows that MFCQ does not hold when $\nabla H(z^*) \neq 0$.

Example 2.4

$$
\begin{aligned}
\min \quad & -y \\
s.t. \quad & x - y - 10 = 0, \\
& x \geq 0, y \geq 0, \\
& xy = 0.
\end{aligned}
$$

Solution. The optimal solution is $(x^*, y^*) = (10, 0)$ showing on Figure 2.3 which is the only feasible point.

(a) Because $y \geq 0$ is binding, LICQ requires that the gradients of the three functions $g_2 = y$, $h_1 = x - y - 10$, $h_2 = xy$ are linear independent. Since $\nabla g_2(x^*, y^*) = (0,1)^T$ and $\nabla h_2(x^*, y^*) = (0, 10)^T$, LICQ can not be satisfied here. (b) MFCQ requires that there exists $v = (v_1, v_2) \in R^2$ such that

$$\langle v, \nabla g_2(x^*, y^*)\rangle = v_2 > 0, \;\; \langle v, \nabla h_2(x^*, y^*)\rangle = v_2 = 0,$$

which contradicts to each other.

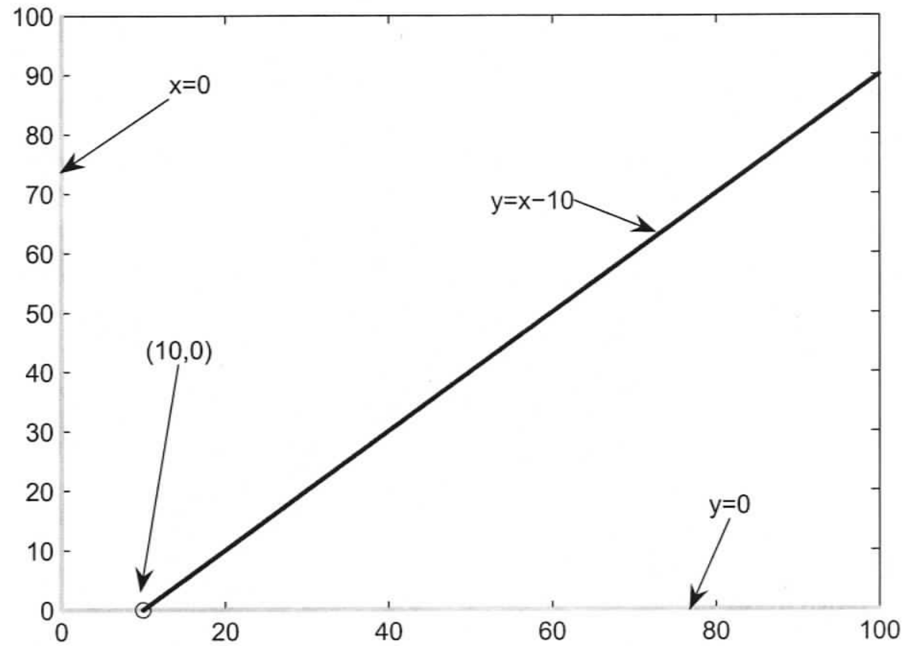

Figure 2.3: The optimal solution of Example 2.4

2.2.2 Stationary conditions

Unlike the standard nonlinear programming which has only one stationary condition, i.e. the KKT conditions, there are various stationary conditions for MPEC. We now summarize them.

Definition 2.4 (Stationary points for MPEC) *Let z^* be a feasible point of MPEC. We say that z^* is weakly stationary (W-stationary) if there exists $\lambda = (\lambda^g, \lambda^h, \lambda^B, \lambda^A) \in R^{p+q+2m}$ such that the following conditions hold:*

$$\begin{aligned} 0 \;=\;& \nabla f(z^*) + \sum_{i \in I_g} \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(z^*) - \sum_{i=1}^{m} [\lambda_i^B \nabla B_i(z^*) \\ & + \lambda_i^A \nabla A_i(z^*)], \end{aligned} \tag{2.1}$$

$$\lambda_{I_g}^g \geq 0, \quad \lambda_\gamma^B = 0, \quad \lambda_\alpha^A = 0. \tag{2.2}$$

We say that z^* is strong stationary (S-stationary) and Mordukhovich stationary (M-stationary) respectively if it is weakly stationary and the following additional conditions are satisfied respectively:

$$\forall i \in \beta, \; \lambda_i^B \geq 0, \; \lambda_i^A \geq 0.$$

$$\forall i \in \beta, \quad \textit{either } \lambda_i^B > 0, \; \lambda_i^A > 0 \quad \textit{or } \lambda_i^B \lambda_i^A = 0.$$

From the above definition, we can see that W-stationary condition or M-stationary condition is weaker than the S-stationary condition. Actually, the S-stationary condition is equivalent to the standard KKT conditions of MPEC. Thus, the W-stationary condition or M-stationary condition is weaker than the standard KKT conditions. Actually, the weakness is in the signs of the multipliers. For the inequality constraints, some of the multipliers can be negative.

2.2.3 Constraint qualifications

Because the standard constraint qualifications such as the MFCQ never hold for the MPEC problems, we need the following new constraint qualifications.

Definition 2.5 (MPEC LICQ) *Let z^* be a feasible point of MPEC where all functions are continuously differentiable at z^*. We say that MPEC linear independence constraint qualification (MPEC LICQ) is satisfied at z^* if the following vectors*

$$\begin{array}{ll} \nabla g_i(z^*) & \forall i \in I_g, \\ \nabla h_i(z^*) & \forall i = 1, 2, \ldots, q, \\ \nabla B_i(z^*) & \forall i \in \alpha \cup \beta, \\ \nabla A_i(z^*) & \forall i \in \gamma \cup \beta \end{array}$$

are linearly independent.

Definition 2.6 (MPEC GMFCQ) *Let z^* be a feasible point of MPEC where all functions are continuously differentiable at z^*. We say that the MPEC generalized Mangasarian-Fromovitz constraint qualification (MPEC GMFCQ) is satisfied at z^* if* *(i) for every partition of β into sets P, Q, R with $R \neq \emptyset$, there exists d such that*

$$\begin{aligned}
&\nabla g_i(z^*)^\top d \leq 0 \quad \forall i \in I_g,\\
&\nabla h_i(z^*)^\top d = 0 \quad \forall i = 1, 2, \ldots, q,\\
&\nabla B_i(z^*)^\top d = 0 \quad \forall i \in \alpha \cup Q,\\
&\nabla A_i(z^*)^\top d = 0 \quad \forall i \in \gamma \cup P,\\
&\nabla B_i(z^*)^\top d \geq 0, \nabla A_i(z^*)^\top d \geq 0 \quad \forall i \in R
\end{aligned}$$

and for some $i \in R$ either $\nabla B_i(z^*)^\top d > 0$ or $\nabla A_i(z^*)^\top d > 0$;

(ii) for every partition of β into sets P, Q, the gradient vectors

$$\begin{aligned}
&\nabla h_i(z^*) \quad \forall i = 1, 2, \ldots, q,\\
&\nabla B_i(z^*) \quad \forall i \in \alpha \cup Q,\\
&\nabla A_i(z^*) \quad \forall i \in \gamma \cup P
\end{aligned}$$

are linearly independent and there exists $d \in R^n$ such that

$$\begin{aligned}
&\nabla g_i(z^*)^\top d < 0 \quad \forall i \in I_g,\\
&\nabla h_i(z^*)^\top d = 0 \quad \forall i = 1, 2, \ldots, q,\\
&\nabla B_i(z^*)^\top d = 0 \quad \forall i \in \alpha \cup Q,\\
&\nabla A_i(z^*)^\top d = 0 \quad \forall i \in \gamma \cup P.
\end{aligned}$$

The following theory tells us that with different constraint qualifications we can get different stationary points.

Theorem 2.3 (Necessary Optimality conditions for MPEC) *Let z^* be a local optimal solution of MPEC. If MPEC GMFCQ holds at z^*, then the M-stationary condition holds at z^*. If MPEC LICQ holds at z^*, then z^* satisfies the S-stationary condition.*

Chapter 3

The principal-agent problem

In this thesis, we study the moral hazard model of the principal-agent problems in the following form as in the references [20, 21]. This problem arises when a principal hires an agent. For example, the agent is an employee and the principal is an employer.

In the moral hazard model, the principal offers a contract to the agent. If the agent's payoff from the contract is lower than his reservation payoff V^*, the agent will reject this contract. The reservation payoff represents the payoff the agent can obtain by breaking his relationship with the principal. If his payoff is higher than V^*, the agent will accept this contract and choose an action from a real interval $A = [\underline{a}, \overline{a}]$ that will maximize his payoff.

The outcome can be one of the N given alternatives, $x_1, x_2, \ldots, x_N$ which are ordered as $x_1 < x_2 < \ldots < x_N$.

If the agent chooses action $a \in A$, the probability for him to generate the outcome x_j is $p_j(a)$.

A contract between the principal and agent consists of an agreement about what wage will be paid to the agent as a function of the outcomes $x_1, x_2, \ldots, x_N$. Let

$$w = (w_1, \ldots, w_N) \in R^N$$

denote a contract, where w_i is the wage paid to the agent if the outcome x_i occurs.

Let $u(x_i - w_i)$ be utility of the principal and $v(w_j)$ be utility of the agent.

The function $c(a)$ is the cost function of the agent's action a.

The expected payoffs to the principal and agent of the contract w when the agent chooses action a are as follows:

$$U(w,a) = \sum_{j=1}^{N} p_j(a)u(x_j - w_j),$$

$$V(w,a) = \sum_{j=1}^{N} p_j(a)v(w_j) - c(a).$$

The relationship between the principal (P) and the agent (A) can be described as follows:

$$\text{P offers } w \text{ to A} \begin{cases} \text{if } V < V^*, \text{ reject,} \left\langle \begin{array}{l} \text{P: } 0, \\ \text{A: } V^*, \end{array} \right. \\ \\ \text{if } V \geq V^*, \text{ accept,} \\ a \in \arg\max_{a' \in A} V(w,a') \text{ s.t. } \left\langle \begin{array}{l} \text{P: } \sum_{j=1}^{N} p_j(a)u(x_j - w_j), \\ \text{A: } \sum_{j=1}^{N} p_j(a)v(w_j) - c(a). \end{array} \right. \end{cases}$$

The following assumptions will be made about the above functions.

Assumption (A.1): $p_j(a) > 0$ for $\forall j \in \{1, 2, \ldots, N\}$ and $\forall a \in A$.

Assumption (A.2): $p_j(a)$ is twice continuously differentiable for $\forall j \in \{1, 2, \ldots, N\}$.

Assumption (A.3): u is increasing.

Assumption (A.4): u is concave and continuously differentiable.

Assumption (A.5): $u(x) = x$.

Assumption (A.6): v is increasing.

Assumption (A.7): v is strictly concave and continuously differentiable.

Assumption (A.8): $c(a)$ is a convex function.

As in most of the literature, the principal is assumed to be risk-neutral which means that u is a linear function of x. Because he hires a large number of agents, the principal can tolerate risks. For simplicity, we assume (A.5) holds. In general, however, the principal is assumed to be risk-adverse which means that he has a concave utility function as in (A.4).

The principal looks for a contract w^* which will influence the agent to choose an action a^* such that (w^*, a^*) solves the following principal-agent problem:

$$\text{(PA)} \quad \max_{w,a} \quad U(w, a) \tag{3.1}$$

$$\text{s.t.} \quad V(w, a) \geq V^*, \tag{3.2}$$

$$a \in \arg\max_{a' \in A} V(w, a'). \tag{3.3}$$

Example 3.1 *BC Firm is a profit-maximizing firm that is currently without a CEO ("manager"). The shareholders are looking for a new CEO. Once hired, the CEO* *can decide to run the firm with effort $a \in A$. The effort cannot be observed directly by the shareholders, but they observe that BC Firm's profits are one of the outcomes: $(x_1, x_2) = (150, 300)$. The relationship between effort and profits is random. With effort a, profits turn out to be x_1 with probability $p_1(a) = (\frac{1}{2})^a$ and x_2 with probability $p_2(a) = 1 - (\frac{1}{2})^a$. Finally, shareholders have a utility function $u(w) = w$ and CEO has a utility function $v(w) = \sqrt{w}$, a reservation payoff $V^* = 5$ and the cost function of his action $c(a) = a$.*

The expected payoffs to the shareholder and CEO of the contract (w_1, w_2) when the CEO chooses action a can be determined as follows:

$$U(w, a) = (\frac{1}{2})^a(150 - w_1) + [1 - (\frac{1}{2})^a](300 - w_2),$$

$$V(w, a) = (\frac{1}{2})^a\sqrt{w_1} + [1 - (\frac{1}{2})^a]\sqrt{w_2} - a.$$

The shareholder looks for a contract w^* which will influence the CEO to choose an action a^* such that (w^*, a^*) solves the following principal-agent problem:

$$\begin{aligned}
\max_{w,a} \quad & (\frac{1}{2})^a(150 - w_1) + [1 - (\frac{1}{2})^a](300 - w_2) \\
\text{s.t.} \quad & (\frac{1}{2})^a\sqrt{w_1} + [1 - (\frac{1}{2})^a]\sqrt{w_2} - a \geq 5, \\
& a \in \arg\max_{a' \in A} \left\{ (\frac{1}{2})^{a'}\sqrt{w_1} + [1 - (\frac{1}{2})^{a'}]\sqrt{w_2} - a' \right\}.
\end{aligned}$$

Chapter 4

Discussion of the first-order approach in Economics

In this section, we introduce most of the economical works about the sufficient conditions for the first order approach to be valid for the PA problem.

The difficulty of the PA problem lies in the incentive constraint (3.3):

$$a \in \arg\max_{a' \in A} V(w, a'),$$

which is a set of countless constraints in the following form

$$V(w, a) \geq V(w, a'), \forall a' \in A.$$

The first-order approach is a technical shortcut widely used in Economics which is to replace the PA problem by the following relaxed principal-agent (RPA) problem:

$$
\begin{aligned}
\text{(RPA)} \quad \max_{w,a} \quad & U(w,a) \\
\text{s.t.} \quad & V(w,a) \geq V^*, \\
& V_a(w,a) = 0, \\
& a \in A.
\end{aligned}
$$

Suppose that (w^*, a^*) is a solution to RPA. If (w^*, a^*) is also a solution of PA, then we say the first order approach is valid. In 1975, Mirrlees [15] pointed out that the first-order approach is not generally correct. In order to find the sufficient conditions for the first order approach to be valid, we need the following conditions.

The first condition is that the probability function $\{p_j(a)\}_{j=1}^N$ of the output x_j with the agent's action a as the parameter satisfies the monotone likelihood ratio condition (MLRC):

Definition 4.1 (MLRC) *The functions $\{p_j(a)\}_{j=1}^N$ are said to satisfy the monotone likelihood ratio condition (MLRC) if $a_1 \leq a_2$ implies that $p_j(a_2)/p_j(a_1)$ is nondecreasing in j.*

The following example satisfies MLRC.

Example 4.1 *In the University of Victoria, there are many courses for students to choose. For each course, the students will get different grades after every exam. Suppose that there are three grades A, B and C which are outcomes $x_1 = C, x_2 = B, x_3 = A$. The student can choose to work hard or not which is higher effort a_H or lower effort a_L, with $a_H > a_L$. The probability functions satisfy the following conditions*

$p_C(a_H)/p_C(a_L) = \frac{1}{2}$, $p_B(a_H)/p_B(a_L) = 1$, $p_A(a_H)/p_A(a_L) = 2$.

By Definition 4.1, MLRC is satisfied here. If the student chooses a_H, the probability to get A is twice of the probability if the student chooses a_L. If the student chooses a_H, the probability to get B is equivalent to the probability if the student chooses a_L. If the student chooses a_H, the probability to get C is half of the probability if the student chooses a_L.

The following proposition provides us an equivalent condition to MLRC.

Proposition 4.1 (Milgrom [12]) *Suppose that (A.1) and (A.2) hold, the functions* $\{p_j(a)\}_{j=1}^N$ *satisfy MLRC if and only if* $p_j'(a)/p_j(a)$ *is nondecreasing in* j *for every* a.

Proof. Notice that $p_j'(a)/p_j(a) = d\log p_j/da$. It follows that for any $a_1, a_2 \in A$ with $a_1 \leq a_2$,

$$p_j(a_2)/p_j(a_1) = exp\Big\{ \int_{a_1}^{a_2} [d\log p_j/da]da \Big\} = exp\Big\{ \int_{a_1}^{a_2} [p_j'(a)/p_j(a)]da \Big\}.$$

The conclusion follows easily.

□

By Proposition 4.2, the following condition can be implied by MLRC.

Definition 4.2 (SDC) *The functions* $\{p_j(a)\}_{j=1}^N$ *are said to satisfy the stochastic dominance condition (SDC) if* $F_j'(a)$ *is nonpositive for every* $j \in \{1, \ldots, N\}$ *and* $a \in A$, *where* $F_j(a)$ *denotes the corresponding distribution function:*

$$F_j(a) = \sum_{i=1}^{j} p_i(a).$$

Proposition 4.2 (Whitt[23]) *Suppose that (A.1) and (A.2) hold. MLRC implies SDC.*

Proof. By MLRC, we know that $p_j(a_2)/p_j(a_1)$ is nondecreasing in j if $a_1 \leq a_2$ and $a_1, a_2 \in A$. Since $p_j(a_1)$ and $p_j(a_2)$ are probability functions, there exists $j \in \{1, \ldots, N\}$ such that $p_j(a_2)/p_j(a_1) \leq 1 \leq p_i(a_2)/p_i(a_1)$ for all $i \geq j$. For any $k \geq j+1$,

$$\sum_{i=k}^{N} p_i(a_1) \leq \sum_{i=k}^{N} p_i(a_1)[p_i(a_2)/p_i(a_1)] = \sum_{i=k}^{N} p_i(a_2)$$

and for any $k \leq j$,

$$\sum_{i=1}^{k} p_i(a_1) \geq \sum_{i=1}^{k} p_i(a_1)[p_i(a_2)/p_i(a_1)] = \sum_{i=1}^{k} p_i(a_2)$$

which implies that $F_j(a_2) \leq F_j(a_1)$, for every $j \in \{1, \ldots, N\}$ and $a_1, a_2 \in A$ with $a_1 \leq a_2$. Hence, $F_j'(a)$ is nonpositive for every $j \in \{1, \ldots, N\}$ and $a \in A$.

□

MLRC means that increase in effort causes expected outcome to increase in the sense of stochastic dominance. The details are as follows:

$$\sum_{i=1}^{N} p_i(a)x_i = \sum_{i=1}^{N-1} F_i(a)(x_i - x_{i+1}) + x_N.$$

Since $x_i < x_{i+1}$ and $F_i(a)$ decreases in a from SDC condition, the expected outcome $\sum_{i=1}^{N} p_i(a)x_i$ increases in a.

The second condition is the convexity of the distribution function condition (CDFC):

Definition 4.3 (CDFC) *The functions $\{p_j(a)\}_{j=1}^N$ are said to satisfy the convexity of the distribution function condition (CDFC) if $F_j''(a)$ is nonnegative for every $j \in \{1, \ldots, N\}$ and $a \in A$.*

CDFC really has no clear economic interpretation, but it allows us to obtain the concavity of $V(w, \cdot)$. The follows are two examples of a family of density functions satisfying CDFC:

Example 4.2 *Let*

$$F_j(a) = (\frac{x_j}{x_N})^{a-\underline{a}},$$

then we have

$$F_j''(a) = (\frac{x_j}{x_N})^{a-\underline{a}} \log^2(\frac{x_j}{x_N}) \geq 0.$$

Example 4.3 *Let*

$$F_j(a) = x_j^{k_1(x_j - x_N)} e^{(a-\underline{a})(x_j - x_N)^{k_2}},$$

where $k_1 \in R, k_2$ is an odd number,

then we have

$$F_j''(a) = x_j^{k_1(x_j - x_N)} e^{(a-\underline{a})(x_j - x_N)^{k_2}} (x_j - x_N)^{2k_2} \geq 0.$$

We can take Example 3.1 as an example here to show that MLRC and CDFC are satisfied.

Example 4.4 *The probability functions in Example 3.1 are as follows:*

$p_1(a) = (\frac{1}{2})^a$ *and* $p_2(a) = 1 - (\frac{1}{2})^a$.

Solution. Since $p_1(a_2)/p_1(a_1) = \frac{(\frac{1}{2})^{a_2}}{(\frac{1}{2})^{a_1}} = (\frac{1}{2})^{a_2-a_1}$ and $p_2(a_2)/p_2(a_2) = \frac{1-(\frac{1}{2})^{a_2}}{1-(\frac{1}{2})^{a_1}}$, we have that for all $a_1 \leq a_2$,

$$\begin{aligned} p_2(a_2)/p_2(a_1) - p_1(a_2)/p_1(a_1) &= \frac{1-(\frac{1}{2})^{a_2}}{1-(\frac{1}{2})^{a_1}} - (\frac{1}{2})^{a_2-a_1} \\ &= \frac{1-(\frac{1}{2})^{a_2-a_1}}{1-(\frac{1}{2})^{a_1}} \\ &\geq 0, \end{aligned}$$

i.e., MLRC holds.

Since $F_1(a) = p_1(a) = (\frac{1}{2})^a$, $F_1''(a) = (\log\frac{1}{2})^2(\frac{1}{2})^a \geq 0$. Since $F_2(a) = 1$, $F_2''(a) = 0$. Thus CDFC is satisfied.

We can use the first order approach to solve Example 3.1 when $a \in A = [0.1, 3]$ as follows:

Solution. By the Optimization Toolbox in Matlab, we can find the optimal solution $w^* = (3.0416, 75.9576)$, $a^* = 2.2727$, $V(w^*, a^*) = V^* = 5$ and $U(w^*, a^*) = 208.0901$.

Mirrlees tried to give a proof in [14] that the first-order approach is valid when MLRC and CDFC hold, but there was some error.

Later, Rogerson [20] found a method by using a further relaxation of the PA program by weakening the incentive constraint (3.3) even further, which is called a double-relaxed principal-agent (DRPA) program:

$$\begin{array}{lll} \text{(DRPA)} & \max\limits_{w,a} & U(w,a) \\ & \text{s.t.} & V(w,a) \geq V^*, \\ & & V_a(w,a) \geq 0, \\ & & a \in A. \end{array}$$

Three more assumptions are required for the results.

Assumption (A.9): A solution to program DRPA exists with $a < \overline{a}$.

Assumption (A.10): A solution to program PA exists with $a > \underline{a}$.

Assumption (A.11): $c'(a) > 0$.

The main idea of the proof in Rogerson [20] is to show that an optimal solution (w^*, a^*) of DRPA is also an optimal solution of PA under MLRC and CDFC. The method is to prove that under MLRC and CDFC the agent's payoff function $V(w^*, a)$ is concave in a when (w^*, a^*) solves DRPA . Thus the stationary point a^* can only be the global maximum of $V(w^*, a)$ on A. Note that, with the Assumptions (A.9), (A10), the global maximum a^* can not appear at the boundary of A.

In order to make the proof easier to read, we change the structure of the proof and add more details. The original proof provides the references for the proof of the nondecreasing property of the w_i^* and $x_i^* - w_i^*$ in i, but we can only find the proof for the nondecreasing property from the references when the w and x are continuous variables. In this model, the w_i and x_i are discrete. That is why we add a different proof for the nondecreasing property of the w_i^* and $x_i^* - w_i^*$ in i into the following proof.

Proposition 4.3 *Suppose that (A.1)-(A.4), (A.6)-(A.11), MLRC and CDFC hold,* *then (w^*, a^*) solves PA if and only if it solves RPA for $a^* \in (\underline{a}, \overline{a})$.*

Proof.

Step 1. We prove that for $a^* < \overline{a}$, (w^*, a^*) solves DRPA if and only if it solves RPA.

Step 1 (A): DRPA$\Rightarrow$ RPA.

Let the Lagrange function of DRPA be

$$\begin{aligned}\mathcal{L}(w, a, \lambda^V, \lambda^{V_a}) &= -U + \lambda^V(V^* - V) + \lambda^{V_a}V_a \\ &= -\sum_{i=1}^{N} p_i(a)u(x_i - w_i) - \lambda^V[\sum_{i=1}^{N} p_i(a)v(w_i) - c(a) - V^*] \\ &\quad +\lambda^{V_a}[\sum_{i=1}^{N} p_i'(a)v(w_i) - c'(a)].\end{aligned} \tag{4.1}$$

Let (w^*, a^*) solve DRPA, then we can find the fact that the solution of DRPA satisfies the KKT conditions from the reference [24] with (A.11). That is, there exist nonnegative real number λ^V and nonpositive real number λ^{V_a} such that:

$$\frac{\partial \mathcal{L}}{\partial w_i}(w^*, a^*) = 0, \forall i \in \{1, \ldots, N\}, \tag{4.2}$$

$$\frac{\partial \mathcal{L}}{\partial a}(w^*, a^*) \begin{cases} \geq & 0 \text{ if } a^* = \underline{a}, \\ = & 0 \text{ if } a^* \in (\underline{a}, \overline{a}), \\ \leq & 0 \text{ if } a^* = \overline{a}. \end{cases} \tag{4.3}$$

$$\lambda^V(V(w^*, a^*) - V^*) = 0, \lambda^{V_a} V_a(w^*, a^*) = 0. \tag{4.4}$$

Equations (4.2) and (4.3) are

$$\frac{u'(x_i - w_i^*)}{v'(w_i^*)} = \lambda^V - \lambda^{V_a} \frac{p_i'(a^*)}{p_i(a^*)}, \forall i \in \{1, \ldots, N\}, \tag{4.5}$$

$$U_a + \lambda^V V_a - \lambda^{V_a} V_{aa} \begin{cases} \leq & 0 \text{ if } a^* = \underline{a}, \\ = & 0 \text{ if } a^* \in (\underline{a}, \overline{a}), \\ \geq & 0 \text{ if } a^* = \overline{a} \end{cases} \tag{4.6}$$

respectively.

By Proposition 4.1, we know that $p_i'(a)/p_i(a)$ is nondecreasing in i. From the fact that $\lambda^V \geq 0, \lambda^{V_a} \leq 0$, it follows from (4.5) that $\frac{u'(x_i - w_i^*)}{v'(w_i^*)}$ is also nondecreasing in i, i.e.,

$$\frac{u'(x_{i+1} - w_{i+1}^*)}{v'(w_{i+1}^*)} \geq \frac{u'(x_i - w_i^*)}{v'(w_i^*)}, \ \forall i \in \{1, \ldots, N-1\}.$$

If there exits an $i \in \{1, \ldots, N-1\}$ such that $w_i^* > w_{i+1}^*$, then $x_i - w_i^* < x_{i+1} - w_{i+1}^*$ for $x_i < x_{i+1}$. Since v is strictly concave, v' is decreasing. Because u is concave, u' is nonincreasing. Therefore, $u'(x_i - w_i^*) \geq u'(x_{i+1} - w_{i+1}^*)$ and $v'(w_i^*) < v'(w_{i+1}^*)$. Thus

$$\frac{u'(x_{i+1} - w_{i+1}^*)}{v'(w_{i+1}^*)} < \frac{u'(x_i - w_i^*)}{v'(w_i^*)}$$

which contradicts to (4.5). Thus we have $w_i^* \leq w_{i+1}^*$, $\forall i \in \{1, \ldots, N-1\}$.

If $\lambda^{V_a} \neq 0$, (4.4) implies that $V_a(w^*, a^*) = 0$. Now suppose, instead, that $\lambda^{V_a} = 0$. It will be shown that $V_a(w^*, a^*)$ still equals to zero. By (4.5),

$$\frac{u'(x_{i+1} - w^*_{i+1})}{v'(w^*_{i+1})} = \frac{u'(x_i - w^*_i)}{v'(w^*_i)}, \forall i \in \{1, \ldots, N-1\}.$$

Since u is concave, u' is nonincreasing. Now we prove that

$$x_i - w^*_i \leq x_{i+1} - w^*_{i+1}, \forall i \in \{1, \ldots, N-1\}.$$

Suppose that there exists an $i \in \{1, \ldots, N-1\}$ such that $x_i - w^*_i > x_{i+1} - w^*_{i+1}$, then $u'(x_i - w^*_i) \leq u'(x_{i+1} - w^*_{i+1})$ which implies that $v'(w^*_i) \leq v'(w^*_{i+1})$. Since v is strictly concave, v' is decreasing. Hence, $w^*_i \geq w^*_{i+1}$. However, we just proved that $w^*_i \leq w^*_{i+1}$, $\forall i \in \{1, \ldots, N-1\}$. Thus $w^*_i = w^*_{i+1}$. We know that $x_i < x_{i+1}$, then $x_i - w^*_i < x_{i+1} - w^*_{i+1}$. This is a contradiction to the assumption that there exists an $i \in \{1, \ldots, N-1\}$ such that $x_i - w^*_i > x_{i+1} - w^*_{i+1}$. Thus $x_i - w^*_i \leq x_{i+1} - w^*_{i+1}$, $\forall i \in \{1, \ldots, N-1\}$.

The principal's expected payoff can be rewritten as

$$U = \sum_{i=1}^{N} p_i(a) u(x_i - w_i) = \sum_{i=1}^{N} \Delta_i [\sum_{j=i}^{N} p_j(a)],$$

where $\Delta_i = \begin{cases} u(x_i - w_i) - u(x_{i-1} - w_{i-1}), & i > 1 \\ u(x_1 - w_1), & i = 1. \end{cases}$

Therefore, $U_a = \sum_{i=2}^{N} \Delta_i [\sum_{j=i}^{N} p'_j(a)]$. By SDC which is implied by MLRC, each of the terms in square brackets $\sum_{j=i}^{N} p'_j(a) = [1 - F_{i-1}(a)]'$ is nonnegative. Because $x_i - w_i^*$ is nondecreasing in i, each of the Δ_i terms is also nonnegative. Hence, U_a is nonnegative. Since $\lambda^{V_a} = 0$, λ^V must be positive for increasing property of u, v by (4.5). Because $\lambda^{V_a} = 0, \lambda^V > 0$ and $U_a \geq 0$, it follows from (4.6) that $V_a(w^*, a^*) \leq 0$ when $a^* < \overline{a}$. However, $V_a(w, a) \geq 0$ in DRPA . Thus $V_a(w^*, a^*) = 0$, then $(w^*, a^*) \in \mathcal{F}_{RPA}$. The fact that (w^*, a^*) solves DRPA means that $U(w^*, a^*) \geq U(w, a)$, $\forall (w, a) \in \mathcal{F}_{DRPA}$. Because $\mathcal{F}_{RPA} \subset \mathcal{F}_{DRPA}$, $U(w^*, a^*) \geq U(w, a)$, $\forall (w, a) \in \mathcal{F}_{RPA}$. Thus (w^*, a^*) solves RPA.

Step 1 (B): RPA $\Rightarrow$ DRPA.

If (w^*, a^*) solves RPA, then $V_a(w^*, a^*) = 0$. Assume, for contradiction, that (w^*, a^*) does not solve DRPA. Let (w, a) be a solution to DRPA with $a < \overline{a}$ which exists from (A.9). Because $\mathcal{F}_{RPA} \subset \mathcal{F}_{DRPA}$, $(w^*, a^*) \in \mathcal{F}_{DRPA}$. Since (w^*, a^*) does not solve DRPA, $U(w, a) > U(w^*, a^*)$. Because (w, a) is a solution to DRPA and $a < \overline{a}$, it is also a solution to RPA from Step 1 (A). Thus $(w^*, a^*), (w, a) \in \mathcal{F}_{RPA}$, when $a < \overline{a}$. For $U(w, a) > U(w^*, a^*)$, (w^*, a^*) cannot solve RPA. This is a contradiction.

Step 2. We prove that (w^*, a^*) solves DRPA if and only if it solves PA.

Step 2 (A): If (w^*, a^*) is a solution to DRPA, then (w^*, a^*) is also a solution to PA.

The agent's expected payoff can be rewritten as

$$V = \sum_{i=1}^{N} v(w_i)p_i(a) - c(a) = \sum_{i=1}^{N} \Delta_i [\sum_{j=i}^{N} p_j(a)] - c(a),$$

$$\text{where } \Delta_i = \begin{cases} v(w_i) - v(w_{i-1}), & i > 1, \\ v(w_1), & i = 1. \end{cases}$$

Therefore, $V_{aa} = \sum_{i=2}^{N} \Delta_i [\sum_{j=i}^{N} p_j''(a)] - c''(a)$. Because w_i^* is nondecreasing (from Step 1 (A)) and v is increasing, each of the Δ_i terms is nonnegative. The function $c(a)$ is convex in a, so $c''(a) \geq 0$. By CDFC, $\sum_{j=i}^{N} p_j''(a) = [1 - F_{i-1}(a)]'' \leq 0$. Hence, $V_{aa}(w^*, a) \leq 0, \forall a \in A$.

If (w^*, a^*) solves DRPA, then V satisfies

$$V_{aa}(w^*, a) \leq 0, \ \forall a \in A.$$

Moreover by Step 1 (A),

$$V_a(w^*, a^*) = 0 \text{ if } a^* \in [\underline{a}, \overline{a})$$

and

$$V_a(w^*, a^*) \geq 0 \text{ if } a^* = \overline{a}.$$

These conditions imply that the agent's action choice is a global maximum, so (w^*, a^*) is also an element of the PA solution set. The PA problem and DRPA problem have the same objective function (principal's expected payoff): $U(w, a)$ and (w^*, a^*) maximizes the objective function over the double-relaxed constrain set. That means

$$U(w^*, a^*) \geq U(w, a), \ \forall (w, a) \in \mathcal{F}_{DRPA},$$

where $\mathcal{F}_{DRPA} = \{(w,a)|V_a(w,a) \geq 0, a \in A\}$.

$$\mathcal{F}_{PA} = \{(w,a)|a \in \arg\max V(w,a)\} = \{(w,a)|V_a(w,a) \begin{cases} \geq 0, & a = \overline{a}, \\ = 0, & \underline{a} < a < \overline{a}, \end{cases} \}$$

when (A.10) holds, so $\mathcal{F}_{PA} \subset \mathcal{F}_{DRPA}$. Now we have

$$U(w^*, a^*) \geq U(w,a),\ \forall (w,a) \in \mathcal{F}_{PA},$$

so (w^*, a^*) solves the PA.

Step 2 (B). If (w^*, a^*) is a solution to the PA, then (w^*, a^*) is a solution to DRPA.

Assume, for contradiction, that (w^*, a^*) does not solve DRPA. Since $a^* > \underline{a}$ for (A.10), it must be true that $V_a(w^*, a^*) \geq 0$. Therefore, $(w^*, a^*) \in \mathcal{F}_{DRPA}$. Let (w,a) be a solution to DRPA with $a < \overline{a}$. Since (w^*, a^*) does not solve DRPA, $U(w,a) > U(w^*, a^*)$. Because (w,a) is a solution to DRPA, (w,a) is also a solution to PA from Step 2 (A). Then $(w^*, a^*), (w,a) \in \mathcal{F}_{PA}$. Thus (w^*, a^*) cannot solve PA for $U(w,a) > U(w^*, a^*)$. This is a contradiction.

In all, for $a^* \in (\underline{a}, \overline{a})$, (w^*, a^*) solves PA if and only if it solves RPA.

□

Chapter 5

PA-MPEC approach

Most of the contributions in this thesis are presented in this chapter. By rewriting the PA problem as a MPEC problem named PA-MPEC, we can use PA-MPEC to solve the PA problem. We call this approach the PA-MPEC approach. This approach is a more general first order approach which allows the optimal action of the agent to be at both interior and boundary points while the first order approach can only be used when the optimal action of the agent is in interior. By the concavity of $V(w, \cdot)$, we can prove that the PA-MPEC approach is valid, which means that the optimal solution of PA-MPEC is also an optimal solution of PA as the first order approach in Economics. By using the methods of MPEC reviewed in Section 2.2, we can find the stationary conditions and constraint qualifications for the PA-MPEC problem. With different constraint qualifications, we can get different stationary conditions at the optimal solution of PA-MPEC. If we assume the proper constraint qualifications, the Weakly-stationary condition holds which is weaker than the KKT conditions. We only need to use the Weakly-stationary condition instead of the KKT conditions to prove the concavity of $V(w,\cdot)$. In order to find the concavity of $V(w,\cdot)$, we also need MLRC and CDFC, when the principal is risk-averse. When the principal is risk-neutral, we just need to assume more general conditions named k-MLRC, k-CDFC and k-OCDFC instead of MLRC and CDFC. Now we get the conditions such that PA-MPEC is equivalent to PA. Under these conditions, we obtain some corollaries to get the optimal solutions of PA both in the interior and at the boundary. We will use the notation $\Theta_k = (0,\ldots,0) \in R^k$ in this chapter.

5.1 Composing the PA-MPEC approach

In this section, we use the methods of MPEC in section 2.2 to set up the PA-MPEC approach.

By replacing (3.3) in the PA problem by its KKT conditions, we obtain the following principal-agent MPEC (PA-MPEC):

$$
\begin{array}{llll}
\text{(PA-MPEC)} & \min\limits_{w,a,b_1,b_2} & -U(w,a) & \\
& \text{s.t.} & V(w,a) \geq V^*, & (5.1)\\
& & V_a(w,a) - b_1 + b_2 = 0, & (5.2)\\
& & (b_1, b_2) \geq 0, & \\
& & (a-\underline{a}, \overline{a}-a) \geq 0, & \\
& & (b_1(a-\underline{a}), b_2(\overline{a}-a)) = 0. &
\end{array}
$$

We call this approach PA-MPEC approach. Suppose that (w^*, a^*, b_1^*, b_2^*) is an optimal solution to PA-MPEC. If (w^*, a^*) is also an optimal solution of PA, then we say the PA-MPEC approach is valid as the first order approach in Economics.

In order to treat the PA-MPEC program as a MPEC problem, we need to define the functions in MPEC as follows:

$$
\begin{array}{lrl}
\min\limits_{w,a,b_1,b_2} & -U(w,a) & f(z) = -U(w,a), \\
\text{s.t.} & V(w,a) \geq V^* & g(z) = V^* - V(w,a), \\
 & V_a(w,a) - b_1 + b_2 = 0 & h(z) = V_a(w,a) - b_1 + b_2, \\
 & (b_1, b_2) \geq 0 & B(z) = (b_1, b_2), \\
 & (a - \underline{a}, \overline{a} - a) \geq 0 & A(z) = (a - \underline{a}, \overline{a} - a), \\
 & (b_1(a - \underline{a}), b_2(\overline{a} - a)) = 0 & B(z)^\top A(z) = 0.
\end{array}
$$

Given a feasible vector $z^* = (w^*, a^*, b_1^*, b_2^*)$ of PA-MPEC, we define the following index sets:

$$
\begin{aligned}
&\alpha := \alpha(z^*) := \{i : B_i(z^*) = 0, A_i(z^*) > 0\}, \\
&i.e.\ 1 \in \alpha \quad if\, b_1^* = 0,\ a^* - \underline{a} > 0, \\
&\qquad 2 \in \alpha \quad if\, b_2^* = 0,\ \overline{a} - a^* > 0. \\
&\beta := \beta(z^*) := \{i : B_i(z^*) = 0, A_i(z^*) = 0\}, \\
&i.e.\ 1 \in \beta \quad if\, b_1^* = 0,\ a^* - \underline{a} = 0, \\
&\qquad 2 \in \beta \quad if\, b_2^* = 0,\ \overline{a} - a^* = 0. \\
&\gamma := \gamma(z^*) := \{i : B_i(z^*) > 0, A_i(z^*) = 0\}, \\
&i.e.\ 1 \in \gamma \quad if\, b_1^* > 0,\ a^* - \underline{a} = 0, \\
&\qquad 2 \in \gamma \quad if\, b_2^* > 0,\ \overline{a} - a^* = 0.
\end{aligned}
$$

In all, there are five cases:

$$
\begin{array}{lll}
Case\,1 & \alpha = \{1,2\}, \beta = \phi, \gamma = \phi & if\, \underline{a} < a^* < \overline{a},\ b_1^* = b_2^* = 0. \\
Case\,2 & \alpha = \{2\}, \beta = \{1\}, \gamma = \phi & if\, a^* = \underline{a},\ b_1^* = 0,\ a^* < \overline{a},\ b_2^* = 0. \\
Case\,3 & \alpha = \{2\}, \beta = \phi, \gamma = \{1\} & if\, a^* = \underline{a},\ b_1^* > 0,\ a^* < \overline{a},\ b_2^* = 0. \\
Case\,4 & \alpha = \{1\}, \beta = \{2\}, \gamma = \phi & if\, a^* = \overline{a},\ b_2^* = 0,\ \underline{a} < a^*,\ b_1^* = 0. \\
Case\,5 & \alpha = \{1\}, \beta = \phi, \gamma = \{2\} & if\, a^* = \overline{a},\ b_2^* > 0,\ \underline{a} < a^*,\ b_1^* = 0.
\end{array}
$$

Actually for each case, we have:

Case 1: (w^*, a^*) is a solution to the following problem:

$$\begin{aligned} \min_{w,a} \quad & -U(w,a) \\ \text{s.t.} \quad & V(w,a) \geq V^*, \\ & V_a(w,a) = 0, \\ & (a - \underline{a}, \overline{a} - a) > 0. \end{aligned}$$

Case 2: $(w^*, \underline{a}, b_1^*)$ is a solution to the following problem:

$$\begin{aligned} \min_{w,a,b_1} \quad & -U(w,a) \\ \text{s.t.} \quad & V(w,a) \geq V^*, \\ & V_a(w,a) \leq 0, \\ & b_1 \geq 0, \\ & a \geq \underline{a}, \\ & b_1(a - \underline{a}) = 0. \end{aligned}$$

Case 3: $(w^*, \underline{a}, b_1^*)$ is a solution to the following problem:

$$\begin{aligned} \min_{w,a,b_1} \quad & -U(w,a) \\ \text{s.t.} \quad & V(w,a) \geq V^*, \\ & V_a(w,a) < 0, \\ & b_1 > 0, \\ & a \geq \underline{a}. \end{aligned}$$

Case 4: $(w^*, \overline{a}, b_2^*)$ is a solution to the following problem:

$$\begin{aligned} \min_{w,a,b_2} \quad & -U(w,a) \\ \text{s.t.} \quad & V(w,a) \geq V^*, \\ & V_a(w,a) \geq 0, \\ & b_2 \geq 0, \\ & a \leq \overline{a}, \\ & b_2(\overline{a} - a) = 0. \end{aligned}$$

Case 5: $(w^*, \overline{a}, b_2^*)$ is a solution to the following problem:

$$\begin{aligned} \min_{w,a,b_2} \quad & -U(w,a) \\ \text{s.t.} \quad & V(w,a) \geq V^*, \\ & V_a(w,a) > 0, \\ & b_2 > 0, \\ & a \leq \overline{a}. \end{aligned}$$

In the following proposition we find that the constraint $V(w,a) \geq V^*$ is always binding for PA-MPEC.

Proposition 5.1 *For the PA-MPEC problem, the constraint (5.1)*

$$V(w,a) \geq V^*$$

is always binding at an optimal solution (w^*, a^*, b_1^*, b_2^*).

Proof. If the constraint (5.1) is not binding at an optimal solution (w^*, a^*, b_1^*, b_2^*), then we can change w^* a little bit such that the constraint (5.1) is binding, i.e. there exists $dw = (dw_1, \cdots, dw_N)$ with $dw_i \leq 0, i = 1, \ldots, N$ such that

$$V(w^* + dw, a^*) = \sum_{i=1}^{N} v(w_i^* + dw_i)p_i(a^*) - c(a^*) = V^*$$

and there exist b_1, b_2 such that

$$V_a(w^* + dw, a^*) = \sum_{i=1}^{N} v(w_i^* + dw_i)p_i'(a^*) - c'(a^*) - b_1 + b_2 = 0.$$

The objective function

$$U(w,a) = \sum_{i=1}^{N} p_i(a)u(x_i - w_i)$$

is increased by $\sum_{i=1}^{N} p_i(a^*)u(x_i - w_i^* - dw_i) - \sum_{i=1}^{N} p_i(a^*)u(x_i - w_i^*) \geq 0$. Thus $(w^* + dw, a^*, b_1, b_2)$ is better than (w^*, a^*, b_1^*, b_2^*). This contradicts to the fact that (w^*, a^*, b_1^*, b_2^*) is an optimal solution.

□

5.2 Definitions for PA-MPEC problem

By the use of Definition 2.4 of stationary conditions for MPEC, we get the following stationary conditions for the PA-MPEC problem.

Definition 5.1 (Stationary points for PA-MPEC) *A feasible point* (w^*, a^*, b_1^*, b_2^*) *of PA-MPEC problem is called weakly stationary (W-stationary) if there exists* $\lambda = (\lambda^V, \lambda^{V_a}, \lambda^A) \in R^4$ *such that the following conditions hold:*

$$\Theta_N = \nabla_w U(w^*, a^*) + \lambda^V \nabla_w V(w^*, a^*) - \lambda^{V_a} \nabla_w V_a(w^*, a^*), \tag{5.3}$$

$$0 = U_a(w^*, a^*) + \lambda^V V_a(w^*, a^*) - \lambda^{V_a} V_{aa}(w^*, a^*) - \lambda_1^A + \lambda_2^A, \tag{5.4}$$

$$\lambda^V \geq 0, \tag{5.5}$$

and

$$
\begin{aligned}
&Case\,1: \lambda_1^A = \lambda_2^A = 0,\\
&Case\,2: \lambda_2^A = 0,\\
&Case\,3: \lambda_2^A = \lambda^{V_a} = 0,\\
&Case\,4: \lambda_1^A = 0,\\
&Case\,5: \lambda_1^A = \lambda^{V_a} = 0.
\end{aligned}
$$

We say that (w^*, a^*, b_1^*, b_2^*) *is Mordukhovich stationary (M-stationary) if it satisfies (5.3)-(5.5) and*

$$
\begin{aligned}
&Case\,1: \lambda_1^A = \lambda_2^A = 0,\\
&Case\,2: \lambda_2^A = 0, \lambda^{V_a} > 0, \lambda_1^A > 0 \quad or \quad \lambda^{V_a}\lambda_1^A = 0,\\
&Case\,3: \lambda_2^A = \lambda^{V_a} = 0,\\
&Case\,4: \lambda_1^A = 0, \lambda^{V_a} < 0, \lambda_2^A > 0 \quad or \quad \lambda^{V_a}\lambda_2^A = 0,\\
&Case\,5: \lambda_1^A = \lambda^{V_a} = 0.
\end{aligned}
$$

We say that (w^*, a^*, b_1^*, b_2^*) *is strong stationary (S-stationary) if it satisfies (5.3)-*

(5.5) and

$$Case\,1: \lambda_1^A = \lambda_2^A = 0,$$

$$Case\,2: \lambda_2^A = 0, \lambda^{V_a} \geq 0, \lambda_1^A \geq 0,$$

$$Case\,3: \lambda_2^A = \lambda^{V_a} = 0,$$

$$Case\,4: \lambda_1^A = 0, \lambda^{V_a} \leq 0, \lambda_2^A \geq 0,$$

$$Case\,5: \lambda_1^A = \lambda^{V_a} = 0.$$

From the above definition, we have the following relation between the stationary conditions:

The S-stationary condition implies the M-stationary condition and the M-stationary condition implies the W-stationary condition. Actually, the S-stationary condition, M-stationary condition and W-stationary condition coincide in Case 1, Case 3 and Case 5 from the above definition.

As explained in the next proposition, the S-stationary condition is equivalent to the standard KKT conditions.

Proposition 5.2 *PA-MPEC satisfies the standard KKT conditions at (w^*, a^*, b_1^*, b_2^*) if and only if the S-stationary condition is satisfied at (w^*, a^*, b_1^*, b_2^*) for PA-MPEC.*

Proof. Let

$$\begin{aligned} \mathcal{L}(w, a, b_1, b_2) \quad = \quad & -U(w,a) + \lambda^V(V^* - V(w,a)) + \lambda^{V_a}(V_a(w,a) + b_1 - b_2) - \lambda_1^{'B} b_1 \\ & -\lambda_2^{'B} b_2 - \lambda_1^{'A}(a - \underline{a}) - \lambda_2^{'A}(\overline{a} - a) + \mu_1 b_1^*(a - \underline{a}) + \mu_2 b_2(\overline{a} - a). \end{aligned}$$

(I) PA-MPEC satisfies the standard KKT conditions at (w^*, a^*, b_1^*, b_2^*) if and only if there exist $\lambda^{V_a}, \mu_1, \mu_2 \in R$ and $\lambda^V, \lambda_1'^A, \lambda_2'^A, \lambda_1'^B, \lambda_2'^B \in [0, \infty)$ such that

$$\nabla \mathcal{L}(w^*, a^*, b_1^*, b_2^*) = 0,$$

$$\lambda^V (V^* - V(w^*, a^*)) = 0,$$

$$\lambda_1'^B b_1^* = 0, \lambda_2'^B b_2^* = 0,$$

$$\lambda_1'^A (a^* - \underline{a}) = 0,$$

$$\lambda_2'^A (\overline{a} - a^*) = 0,$$

i.e.,

$$\begin{aligned}
\Theta_N &= \nabla_w U(w^*, a^*) + \lambda^V \nabla_w V(w^*, a^*) - \lambda^{V_a} \nabla_w V_a(w^*, a^*), \\
0 &= U_a(w^*, a^*) + \lambda^V V_a(w^*, a^*) - \lambda^{V_a} V_{aa}(w^*, a^*) + \lambda_1'^A - \lambda_2'^A \\
&\quad -\mu_1 b_1^* + \mu_2 b_2^*,
\end{aligned}$$

$$\begin{aligned}
0 &= \lambda^{V_a} - \lambda_1'^B + \mu_1 (a^* - \underline{a}), \\
0 &= -\lambda^{V_a} - \lambda_2'^B + \mu_2 (\overline{a} - a^*)
\end{aligned}$$

and $\lambda^V \geq 0$ because constraint (5.1) is always binding by Proposition 5.1;

$\lambda_1'^A \geq 0$, when $a^* = \underline{a}$; $\lambda_1'^A = 0$, when $a^* > \underline{a}$;

$\lambda_2'^A \geq 0$, when $a^* = \overline{a}$; $\lambda_2'^A = 0$, when $a^* < \overline{a}$;

$\lambda_1'^B \geq 0$, when $b_1^* = 0$; $\lambda_1'^B = 0$, when $b_1^* > 0$;

$\lambda_2^{'B} \geq 0$, when $b_2^* = 0$; $\lambda_2^{'B} = 0$, when $b_2^* > 0$.

Set $\lambda_1^A = \lambda_1^{'A} - \mu_1 b_1^*$, $\lambda_2^A = \lambda_2^{'A} - \mu_2 b_2^*$,

$\lambda_1^B = \lambda_1^{'B} - \mu_1(a^* - \underline{a})$, $\lambda_2^B = \lambda_2^{'B} - \mu_2(\overline{a} - a^*)$.

Since Case 1: $\underline{a} < a^* < \overline{a}$, $b_1^* = b_2^* = 0$,

Case 2: $a^* = \underline{a}$, $b_1^* = 0$, $a^* < \overline{a}$, $b_2^* = 0$,

Case 3: $a^* = \underline{a}$, $b_1^* > 0$, $a^* < \overline{a}$, $b_2^* = 0$,

Case 4: $a^* = \overline{a}$, $b_2^* = 0$, $\underline{a} < a^*$, $b_1^* = 0$,

Case 5: $a^* = \overline{a}$, $b_2^* > 0$, $\underline{a} < a^*$, $b_1^* = 0$,

we have the following conditions:

$$\Theta_N = \nabla_w U(w^*, a^*) + \lambda^V \nabla_w V(w^*, a^*) - \lambda^{V_a} \nabla_w V_a(w^*, a^*),$$

$$0 = U_a(w^*, a^*) + \lambda^V V_a(w^*, a^*) - \lambda^{V_a} V_{aa}(w^*, a^*) + \lambda_1^A - \lambda_2^A,$$

$$\lambda^V \geq 0$$

and

Case 1: $\lambda_1^A = \lambda_1^{'A} = \lambda_2^A = \lambda_2^{'A} = 0$,

Case 2 : $\lambda^{V_a} = \lambda_1^B = \lambda_1^{'B} \geq 0, \lambda_2^A = \lambda_2^{'A} = 0, \lambda_1^A = \lambda_1^{'A} \geq 0$,

Case 3: $\lambda^{V_a} = \lambda_1^B = \lambda_1^{'B} = \lambda_2^A = \lambda_2^{'A} = 0$,

Case 4: $\lambda^{V_a} = -\lambda_2^B = -\lambda_2^{'B} \leq 0, \lambda_1^A = \lambda_1^{'A} = 0, \lambda_2^A = \lambda_2^{'A} \geq 0$,

Case 5: $\lambda^{V_a} = -\lambda_2^B = -\lambda_2^{'B} = \lambda_2^A = \lambda_2^{'A} = 0$,

i.e., the S-stationary condition is satisfied at (w^*, a^*, b_1^*, b_2^*) for PA-MPEC.

(II) If the S-stationary condition is satisfied at (w^*, a^*, b_1^*, b_2^*) for PA-MPEC, there exist $\lambda^V, \lambda^{V_a}, \lambda_1^B, \lambda_2^B, \lambda_1^A, \lambda_2^A \in R$ such that the following conditions hold:

$$\Theta_N = \nabla_w U(w^*, a^*) + \lambda^V \nabla_w V(w^*, a^*) - \lambda^{V_a} \nabla_w V_a(w^*, a^*),$$

$$0 = U_a(w^*, a^*) + \lambda^V V_a(w^*, a^*) - \lambda^{V_a} V_{aa}(w^*, a^*) - \lambda_1^A + \lambda_2^A,$$

$$\lambda^V \geq 0$$

and

$$Case\,1 : \lambda_1^A = \lambda_2^A = 0,$$

$$Case\,2 : \lambda_2^A = 0, \lambda^{V_a} \geq 0, \lambda_1^A \geq 0,$$

$$Case\,3 : \lambda_2^A = \lambda^{V_a} = 0,$$

$$Case\,4 : \lambda_1^A = 0, \lambda^{V_a} \leq 0, \lambda_2^A \geq 0,$$

$$Case\,5 : \lambda_1^A = \lambda^{V_a} = 0.$$

Set $\lambda_1{}'A = \lambda_1^A - \mu_1 b_1^*$, $\lambda_2{}'A = \lambda_2^A - \mu_2 b_2^*$,

$\lambda_1'^B = \lambda^{V_a} - \mu_1(a^* - \underline{a})$, $\lambda_2'^B = \lambda^{V_a} - \mu_2(\overline{a} - a^*)$.

Since

Case 1: $\underline{a} < a^* < \overline{a}$, $b_1^* = b_2^* = 0$,

Case 2: $a^* = \underline{a}$, $b_1^* = 0$, $a^* < \overline{a}$, $b_2^* = 0$,

Case 3: $a^* = \underline{a}$, $b_1^* > 0$, $a^* < \overline{a}$, $b_2^* = 0$,

Case 4: $a^* = \overline{a}$, $b_2^* = 0$, $\underline{a} < a^*$, $b_1^* = 0$,

Case 5: $a^* = \overline{a}$, $b_2^* > 0$, $\underline{a} < a^*$, $b_1^* = 0$, if there exist $\mu_1, \mu_2 \in R$ such that

$$\lambda^{V_a} = -\mu_1(a^* - \underline{a}) = \mu_2(\overline{a} - a^*),$$

$$\lambda_1^A = -\mu_1 b_1^*, \lambda_2^A = -\mu_2 b_2^*,$$

then we have $\lambda_1^{'A}, \lambda_2^{'A}, \lambda_1^{'B}, \lambda_2^{'B} \in [0, \infty)$ and

$$\nabla \mathcal{L}(w^*, a^*, b_1^*, b_2^*) = 0,$$

$$\lambda_1^{'B} b_1^* = 0, \lambda_2^{'B} b_2^* = 0,$$

$$\lambda_1^{'A}(a^* - \underline{a}) = 0,$$

$$\lambda_2^{'A}(\overline{a} - a^*) = 0.$$

We know that $\lambda^V \geq 0$ and $\lambda^V(V^* - V(w^*, a^*)) = 0$ because of Proposition 5.1. Thus the standard KKT conditions of MPEC hold at z^*.

□

From the equivalence of the S-stationary condition and the standard KKT conditions of MPEC, we can see that the W-stationary condition or M-stationary condition is weaker than the standard KKT conditions. Actually, the weakness is in the signs of the multipliers. For the inequality constraints, some of the multipliers can be negative. Later, we will discuss about the signs of λ^{V_a} and assume different conditions for different signs of λ^{V_a} in Theorem 5.8.

The standard constraint qualifications such as LICQ and MFCQ never hold for the PA-MPEC problem in Case 2 and Case 4. We can take Case 4 as an example here to show this fact.

Example 5.1

$$\begin{aligned} \min_{w,a,b_2} \quad & -U(w,a) \\ s.t. \quad & V(w,a) \geq V^*, \\ & V_a(w,a) \geq 0, \\ & b_2 \geq 0, \\ & a \leq \overline{a}, \\ & b_2(\overline{a} - a) = 0. \end{aligned}$$

Solution. The optimal solution is $z^* = (w^*, \overline{a}, 0)$. In this example, $g_1 = V^* - V(w,a)$, $g_2 = -V_a(w,a)$, $g_3 = -b_2$, $g_4 = a - \overline{a}$ are all binding and $h_1 = b_2(\overline{a} - a)$. LICQ requires the gradients of the five functions g_1, g_2, g_3, g_4, h_1 are linearly independent and MFCQ requires $\nabla h_1(z^*) \neq 0$. Because $\nabla h_1(z^*) = (0,0,0)$, both LICQ and MFCQ can not be satisfied here.

In order to discuss the PA-MPEC problem, we need the following constraint qualifications which are deducted from the Definitions 2.5, 2.6 for Problem PA:

Definition 5.2 (PA-MPEC LICQ) *Let (w^*, a^*, b_1^*, b_2^*) be a feasible point of PA-MPEC problem where all functions are continuously differentiable at (w^*, a^*, b_1^*, b_2^*). We say that the principal-agent problem linear independence constraint qualification (PA-MPEC LICQ) is satisfied at (w^*, a^*, b_1^*, b_2^*) if the following conditions hold for each case:*

Case 1: $\nabla V(w^*, a^*), \nabla V_a(w^*, a^*)$ are linearly independent,

Case 2: $\nabla_w V(w^*, a^*)$, $\nabla_w V_a(w^*, a^*)$ are linearly independent,

Case 3: $\nabla_w V(w^*, a^*) \neq \Theta_N$,

Case 4: $\nabla_w V(w^*, a^*)$, $\nabla_w V_a(w^*, a^*)$ are linearly independent, *Case 5:* $\nabla_w V(w^*, a^*) \neq \Theta_N$.

Definition 5.3 (PA-MPEC GMFCQ) *Let* (w^*, a^*, b_1^*, b_2^*) *be a feasible point of PA-MPEC problem where all functions are continuously differentiable at* (w^*, a^*, b_1^*, b_2^*). *We say that the principal-agent problem generalized Mangasarian-Fromovitz constraint qualification (PA-MPEC GMFCQ) is satisfied at* (w^*, a^*, b_1^*, b_2^*) *if the following conditions hold for each case.*

Case 1: $\alpha = \{1, 2\}, \beta = \phi, \gamma = \phi$.

$\nabla V_a(w^*, a^*) \neq \Theta_{N+1}$ *and* $\exists\, d \in R^{N+1}$ *such that*

$$\nabla V(w^*, a^*)^\top d > 0,$$

$$\nabla V_a(w^*, a^*)^\top d = 0.$$

Case 2: $\alpha = \{2\}, \beta = \{1\}, \gamma = \phi$.

(I) $\exists\, d \in R^N, d_{N+1}, d_{N+2} \in R$ *such that*

$$\nabla_w V(w^*, a^*)^\top d + V_a(w^*, a^*) d_{N+1} \geq 0,$$

$$\nabla_w V_a(w^*, a^*)^\top d + V_{aa}(w^*, a^*) d_{N+1} + d_{N+2} = 0$$

and either $d_{N+2} > 0, d_{N+1} \geq 0$ *or* $d_{N+1} > 0, d_{N+2} \geq 0$.

(II) i) $\nabla_w V(w^*, a^*) \neq \Theta_N$.

ii) $\nabla V_a(w^*, a^*) \neq \Theta_{N+1}$ *and* $\exists\, d \in R^{N+1}$ *such that*

$$\nabla V(w^*, a^*)^\top d > 0,$$

$$\nabla V_a(w^*, a^*)^\top d = 0.$$

Case 3: $\alpha = \{2\}, \beta = \phi, \gamma = \{1\}$.

$$\nabla_w V(w^*, a^*) \neq \Theta_N.$$

Case 4: $\alpha = \{1\}, \beta = \{2\}, \gamma = \phi$.

(I) $\exists\, d \in R^N, d_{N+1}, d_{N+3} \in R$ *such that*

$$\nabla_w V(w^*, a^*)^\top d + V_a(w^*, a^*) d_{N+1} \geq 0,$$

$$\nabla_w V_a(w^*, a^*)^\top d + V_{aa}(w^*, a^*) d_{N+1} - d_{N+3} = 0$$

and either $d_{N+3} > 0, d_{N+1} \leq 0$ *or* $d_{N+1} < 0, d_{N+3} \geq 0$.

(II) i) $\nabla_w V(w^*, a^*) \neq \Theta_N$.

ii) $\nabla V_a(w^*, a^*)^T \neq \Theta_{N+1}$ *and* $\exists\, d \in R^{N+1}$ *such that*

$$\nabla V(w^*, a^*)^\top d > 0,$$

$$\nabla V_a(w^*, a^*)^\top d = 0.$$

Case 5: $\alpha = \{1\}, \beta = \phi, \gamma = \{2\}$.

$$\nabla_w V(w^*, a^*) \neq \Theta_N.$$

The following theorem is a PA-MPEC version of Theorem 2.3.

Theorem 5.3 (Necessary optimality conditions for PA-MPECs) *Let z^* be a local (global) optimal solution of PA-MPEC. If the PA-MPEC GMFCQ holds at z^*, then M-stationary condition holds at z^*. If PA-MPEC LICQ holds at z^*, then z^* satisfies the S-stationary condition.*

□

5.3 New conditions

If we want to prove the relation between PA and PA-MPEC, we need the concavity of $V(w, \cdot)$. In order to prove that $V(w, \cdot)$ is concave, we introduce the following new conditions.

Definition 5.4 (k-MLRC) *Let $1 \leq k \leq N$. The density functions $\{p_i(a)\}_{i=1}^N$ are said to satisfy the k-monotone likelihood ratio condition (k-MLRC) if $a_1 \leq a_2$ implies that $p_i(a_2)/p_i(a_1)$ is nondecreasing in i, for $i = 1, \cdots, k$, $p_i(a_2)/p_i(a_1)$ is nonincreasing in i, for $i = k, \cdots, N$.*

From x_1 to x_k, the ratio of the probability $p_i(a_2)/p_i(a_1)$ with the high action over low action increases when the outcome increases. From x_{k+1} to x_N, the ratio of the probability $p_i(a_2)/p_i(a_1)$ with the high action over low action decreases when outcome increases, which means the risk is higher for the higher outcome. For example, the agent may work really hard to produce more old products with low profits. In this case, the probability to earn money is very high, but the outcome is not extremely high. Instead of working on old products, the worker may spend more time to do research on finding new products. He may find a better product which can be sold for extremely high profits, but the probability to find a new product is little.

Proposition 5.4 *Suppose that (A.1) and (A.2) hold, if the density functions* $\{p_i(a)\}_{i=1}^N$ *satisfies k-MLRC if and only if* $p_i'(a)/p_i(a)$ *is nondecreasing in* i*, for* $i = 1, \cdots, k$ *and* $p_i'(a)/p_i(a)$ *is nonincreasing in* i*, for* $i = k, \cdots, N$*,* $\forall a \in A$*.*

The proof is the same as Proposition 4.1.

Definition 5.5 (k-CDFC) *The functions* $\{p_j(a)\}_{j=1}^N$ *are said to satisfy the* k*-convexity of the distribution function condition (*k*-CDFC) if* $F_i''(a)$ *is nonnegative for every* $i \in \{1, \ldots, k-1\}$*,* $F_i''(a)$ *is nonpositive for every* $i \in \{k, \ldots, N-1\}$*.*

The following example contains a technology that satisfies the k-CDFC condition.

Example 5.2 *(a) When* N *is odd, let* $k = \frac{N+1}{2}$ *and*

$$p_j(a) = \begin{cases} \frac{m}{n}(\frac{x_1}{x_k})^{a-\underline{a}}, & j = 1, \\ \frac{m}{n}[(\frac{x_j}{x_k})^{a-\underline{a}} - (\frac{x_{j-1}}{x_k})^{a-\underline{a}}], & \forall j \in \{2, \ldots, k-1\}, \\ 1 - (\frac{x_{k-1}}{x_k})^{a-\underline{a}}, & j = k, \\ \frac{n-m}{n}[(\frac{x_{N+1-j}}{x_k})^{a-\underline{a}} - (\frac{x_{N-j}}{x_k})^{a-\underline{a}}], & \forall j \in \{k+1, \ldots, N-1\}, \\ \frac{n-m}{n}(\frac{x_1}{x_k})^{a-\underline{a}}, & j = N, \end{cases}$$

where $m, n \in R$ *and* $m \leq n$*.*

(b) When N is even, let $k = \frac{N}{2}$ and

$$p_j(a) = \begin{cases} \frac{m}{n}(\frac{x_1}{x_k})^{a-\underline{a}}, & j = 1, \\ \frac{m}{n}[(\frac{x_j}{x_k})^{a-\underline{a}} - (\frac{x_{j-1}}{x_k})^{a-\underline{a}}], & \forall j \in \{2, \ldots, k\}, \\ \frac{n-m}{n}[(\frac{x_{N+1-j}}{x_k})^{a-\underline{a}} - (\frac{x_{N-j}}{x_k})^{a-\underline{a}}], & \forall j \in \{k+1, \ldots, N-1\}, \\ \frac{n-m}{n}(\frac{x_1}{x_k})^{a-\underline{a}}, & j = N, \end{cases}$$

where $m, n \in R$ and $2m \geq n \geq m$.

Hence, we have

(a)

$$F_j(a) = \begin{cases} \frac{m}{n}(\frac{x_j}{x_k})^{a-\underline{a}}, & \forall j \in \{1, \ldots, k-1\}, \\ 1 - \frac{n-m}{n}(\frac{x_{N-j}}{x_k})^{a-\underline{a}}, & \forall j \in \{k, \ldots, N-1\}, \\ 1, & j = N \end{cases}$$

and

$$F_j''(a) = \begin{cases} \frac{m}{n}(\frac{x_j}{x_k})^{a-\underline{a}} \log^2(\frac{x_j}{x_k}) \geq 0, & \forall j \in \{1, \ldots, k-1\}, \\ -\frac{n-m}{n}(\frac{x_{N-j}}{x_k})^{a-\underline{a}} \log^2(\frac{x_{N-j}}{x_k}) \leq 0, & \forall j \in \{k, \ldots, N-1\}, \\ 0, & j = N. \end{cases}$$

(b)

$$F_j(a) = \begin{cases} \frac{m}{n}(\frac{x_j}{x_k})^{a-\underline{a}}, & \forall j \in \{1, \ldots, k-1\}, \\ 1 - \frac{n-m}{n}(\frac{x_{N-j}}{x_k})^{a-\underline{a}}, & \forall j \in \{k, \ldots, N-1\}, \\ 1, & j = N \end{cases}$$

and

$$F_j''(a) = \begin{cases} \frac{m}{n}(\frac{x_j}{x_k})^{a-\underline{a}} \log^2(\frac{x_j}{x_k}) \geq 0, & \forall j \in \{1, \ldots, k-1\}, \\ -\frac{n-m}{n}(\frac{x_{N-j}}{x_k})^{a-\underline{a}} \log^2(\frac{x_{N-j}}{x_k}) \leq 0, & \forall j \in \{k+1, \ldots, N-1\}, \\ 0, & j = N. \end{cases}$$

The following example satisfies 2-MLRC and 2-CDFC conditions.

Example 5.3 *BC Firm is a profit-maximizing firm that is currently without a CEO ("manager"). The shareholders are looking for a new CEO. Once hired, the CEO can decide to run the firm with effort $a \in A = [0.1, 2]$. The effort cannot be observed directly by the shareholders, but they observe that BC Firm's profits are one of the outcomes: $(x_1, x_2, x_3) = (10, 100, 1000)$. The relationship between effort and profits is random. With effort a, profits turn out to be x_1 with probability $p_1(a) = 0.99(\frac{1}{10})^a$, x_2 with probability $p_2(a) = 1 - (\frac{1}{10})^a$ and x_3 with probability $p_3(a) = 0.01(\frac{1}{10})^a$. Finally, shareholders have a utility function $u(w) = w$ and CEO has a utility function $v(w) = \sqrt{w}$, a reservation payoff $V^* = 4$ and the cost function of his action $c(a) = a^2$.*

The expected payoffs to the shareholder and CEO of the contract (w_1, w_2, w_3) when the CEO chooses action a can be determined as follows:

$$U(w, a) = 0.99(\frac{1}{10})^a(10 - w_1) + [1 - (\frac{1}{10})^a](100 - w_2) + 0.01(\frac{1}{10})^a(1000 - w_3),$$

$$V(w, a) = 0.99(\frac{1}{10})^a\sqrt{w_1} + [1 - (\frac{1}{10})^a]\sqrt{w_2} + 0.01(\frac{1}{10})^a\sqrt{w_3} - a^2.$$

The shareholder looks for a contract w^* which will influence the CEO to choose an action a^* such that (w^*, a^*) solves the following principal-agent problem:

$$\begin{array}{ll}\max\limits_{w,a} & 0.99(\frac{1}{10})^a(10-w_1)+[1-(\frac{1}{10})^a](100-w_2)+0.01(\frac{1}{10})^a(1000-w_3)\\ \text{s.t.} & 0.99(\frac{1}{10})^a\sqrt{w_1}+[1-(\frac{1}{10})^a]\sqrt{w_2}+0.01(\frac{1}{10})^a\sqrt{w_3}-a^2\geq 4,\\ & a\in\arg\max\limits_{a'\in A}\left\{0.99(\frac{1}{10})^{a'}\sqrt{w_1}+[1-(\frac{1}{10})^{a'}]\sqrt{w_2}+0.01(\frac{1}{10})^{a'}\sqrt{w_3}-a'^2\right\}.\end{array}$$

Solution. By the PA-MPEC approach, the optimal solution is $w^* = (1.0271, 30.4993, 1.0272)$, $a^* = 0.8738$, $V(w^*, a^*) = V^* = 4$ and $U(w^*, a^*) = 60.9786$.

The contract $w^* = (1.0271, 30.4993, 1.0272)$ is a risk sharing contract which has a different distribution as the outcomes.

We can take the probability functions in Example 5.3 as an example here to show that 2-MLRC and 2-CDFC are satisfied.

Example 5.4 $p_1(a) = 0.99(\frac{1}{10})^a$, $p_2(a) = 1-(\frac{1}{10})^a$ *and* $p_3(a) = 0.01(\frac{1}{10})^a$.

Proof. For $a_1 \leq a_2 \in A$, $p_1(a_2)/p_1(a_1) = \frac{0.99(\frac{1}{10})^{a_2}}{0.99(\frac{1}{10})^{a_1}} = (\frac{1}{10})^{a_2-a_1}$, $p_2(a_2)/p_2(a_1) = \frac{1-(\frac{1}{10})^{a_2}}{1-(\frac{1}{10})^{a_1}}$ and $p_3(a_2)/p_3(a_1) = \frac{0.01(\frac{1}{10})^{a_2}}{0.01(\frac{1}{10})^{a_1}} = (\frac{1}{10})^{a_2-a_1}$.

$$\begin{aligned} p_2(a_2)/p_2(a_1) - p_1(a_2)/p_1(a_1) &= \frac{1-(\frac{1}{10})^{a_2}}{1-(\frac{1}{10})^{a_1}} - (\frac{1}{10})^{a_2-a_1} \\ &= \frac{1-(\frac{1}{10})^{a_2-a_1}}{1-(\frac{1}{10})^{a_1}} \\ &\geq 0 \end{aligned}$$

with $a_1 \leq a_2 \in A$. Then $p_2(a_1)/p_2(a_2) \geq p_3(a_1)/p_3(a_2) = p_1(a_1)/p_1(a_2)$. Thus 2-MLRC is satisfied.

$F_1(a) = p_1(a) = (\frac{1}{10})^a$, then $F_1''(a) = (log\frac{1}{10})^2(\frac{1}{10})^a \geq 0$. $F_2(a) = 1 - 0.01(\frac{1}{10})^a$, then $F_2''(a) = -0.01(log\frac{1}{10})^2(\frac{1}{10})^a \leq 0$. $F_3(a) = 1$, then $F_3''(a) = 0$. Thus 2-CDFC is satisfied.

In order to compare the examples under 2-MLRC and MLRC, we need the following example.

Example 5.5 *BC Firm is a profit-maximizing firm that is currently without a CEO ("manager"). The shareholders are looking for a new CEO. Once hired, the CEO can decide to run the firm with effort $a \in A = [0.1, 2]$. The effort cannot be observed directly by the shareholders, but they observe that BC Firm's profits are one of the outcomes: $(x_1, x_2) = (19.9, 100)$. The relationship between effort and profits is random. With effort a, profits turn out to be x_1 with probability $p_1(a) = (\frac{19.9}{100})^a$ and x_2 with probability $p_2(a) = 1 - (\frac{19.9}{100})^a$. Finally, shareholders have a utility function $u(w) = w$ and CEO has a utility function $v(w) = \sqrt{w}$, a reservation payoff $V^* = 4$ and the cost function of his action $c(a) = a^2$.*

The expected payoffs to the shareholder and CEO of the contract (w_1, w_2) when the CEO chooses action a can be determined as follows:

$$U(w, a) = (\frac{19.9}{100})^a(19.9 - w_1) + [1 - (\frac{19.9}{100})^a](100 - w_2),$$

$$V(w, a) = (\frac{19.9}{100})^a\sqrt{w_1} + [1 - (\frac{19.9}{100})^a]\sqrt{w_2} - a^2.$$

The shareholder looks for a contract w^* which will influence the CEO to choose an action a^* such that (w^*, a^*) solves the following principal-agent problem:

$$\begin{array}{ll} \max\limits_{w,a} & (\frac{19.9}{100})^a(19.9 - w_1) + [1 - (\frac{19.9}{100})^a](100 - w_2) \\ \text{s.t.} & (\frac{19.9}{100})^a\sqrt{w_1} + [1 - (\frac{19.9}{100})^a]\sqrt{w_2} - a^2 \geq 4, \\ & a \in \arg\max\limits_{a' \in A}\left\{(\frac{19.9}{100})^{a'}\sqrt{w_1} + [1 - (\frac{19.9}{100})^{a'}]\sqrt{w_2} - a'^2\right\}. \end{array}$$

Solution. By the PA-MPEC approach, the optimal solution is

$w^* = (2.7110, 36.0780)$, $a^* = 0.9266$, $V(w^*, a^*) = V^* = 4$ and $U(w^*, a^*) = 51.6178$.

The contract $w^* = (2.7110, 36.0780)$ is an incentive contract which is increasing in the outcomes.

For Examples 5.3, 5.5, the expected outcome $\sum_{i=1}^N p_i(a)x_i$ is lower in Example 5.5. The payoff of the principal U in Example 5.3 is higher, but the agent in Example 5.5 needs to work harder and the principal needs to pay more wage. The difference between Examples 5.3, 5.5 is that there is a risky outcome in Example 5.3. This outcome is very high, but the chance to get that outcome is very low. In this situation, the risky project is good to the principal because he can save money on wages when the agent does not work too hard.

Definition 5.6 (k-OCDFC) *The functions $\{p_j(a)\}_{j=1}^N$ are said to satisfy the k-opposite convexity of the distribution function condition (k-OCDFC) if $F_i''(a)$ is non-positive for every $i \in \{1, \dots, k-1\}$, $F_i''(a)$ is nonnegative for every $i \in \{k, \dots, N-1\}$.*

There is no clear economic interpretation about k-CDFC and k-OCDFC, but it allows us to obtain the concavity of $V(w, \cdot)$.

When $k < N$, we discuss the concavity of $V(w, \cdot)$ in the model with the risk-neutral principal, so (A.5) holds. We call this model Risk Neutral Principal Model. We prove the property of w_i^* by weaker condition which is W-stationary condition instead of the KKT conditions.

Proposition 5.5 (Risk Neutral Principal Model) *Let z^* be a feasible point of PA-MPEC. Suppose that (A.5) and (A.7) hold. Assume that the W-stationary condition holds at z^* and there is $k \in \{1, \ldots, N\}$ such that k-MLRC holds when $\lambda^{V_a} \neq 0$. Then*

- *when $\lambda^{V_a} < 0$, w_i^* is nondecreasing in i for $\forall i \in \{1, \ldots, k\}$ and nonincreasing in i for $\forall i \in \{k, \ldots, N\}$;*
- *when $\lambda^{V_a} > 0$, w_i^* is nonincreasing in i for $\forall i \in \{1, \ldots, k\}$ and nondecreasing in i for $\forall i \in \{k, \ldots, N\}$;*
- *when $\lambda^{V_a} = 0$, $w_i^* = w_{i+1}^*$, $\forall i \in \{1, \ldots, N-1\}$.*

Proof. For the definition of W-stationary condition, the reader can refer to Definition 5.1. By (5.3),

$$-p_i(a^*) + \lambda^V p_i(a^*) v'(w_i^*) - \lambda^{V_a} p_i'(a^*) v'(w_i^*) = 0, \forall i \in \{1, \ldots, N\},$$

which implies that

$$\frac{1}{v'(w_i^*)} = \lambda^V - \lambda^{V_a} \frac{p_i'(a^*)}{p_i(a^*)}, \forall i \in \{1, \dots, N\}. \tag{5.6}$$

By Proposition 5.4 and Definition 5.4 for $\lambda^{V_a} < 0$, it follows from (5.6) that

$$\frac{1}{v'(w_{i+1}^*)} \geq \frac{1}{v'(w_i^*)}, \forall i \in \{1, \dots, k-1\}.$$

Then

$$v'(w_{i+1}^*) \leq v'(w_i^*), \forall i \in \{1, \dots, k-1\}.$$

Since v is concave, v' is decreasing. Then we have $w_i^* \leq w_{i+1}^*$, $\forall i \in \{1, \dots, k-1\}$. Similar to above proof, we have $w_i^* \geq w_{i+1}^*$, $\forall i \in \{k, \dots, N-1\}$, which is in the opposite way.

Similarly, if k-MLRC holds for $\lambda^{V_a} > 0$, then we have $w_i^* \geq w_{i+1}^*$, $\forall i \in \{1, \dots, k-1\}$, $w_i^* \leq w_{i+1}^*$, $\forall i \in \{k, \dots, N-1\}$.

If $\lambda^{V_a} = 0$, it follows from (5.6) that

$$\frac{u'(x_{i+1} - w_{i+1}^*)}{v'(w_{i+1}^*)} = \frac{u'(x_i - w_i^*)}{v'(w_i^*)}, \forall i \in \{1, \dots, N-1\}.$$

Similar to the proof in (I), we can prove that $w_i^* = w_{i+1}^*$, $\forall i \in \{1, \dots, N-1\}$.

$\square$

In the following, we will find the sign of λ^{V_a} for Examples 5.3, 5.5.

By Equation (5.6), we have following equations for Example 5.3.

$$\frac{1}{v'(w_1^*)} = \lambda^V - \lambda^{V_a}\frac{p_1'(a^*)}{p_1(a^*)}, \tag{5.7}$$
$$\frac{1}{v'(w_2^*)} = \lambda^V - \lambda^{V_a}\frac{p_2'(a^*)}{p_2(a^*)}, \tag{5.8}$$
$$\frac{1}{v'(w_3^*)} = \lambda^V - \lambda^{V_a}\frac{p_3'(a^*)}{p_3(a^*)}. \tag{5.9}$$

Because $\frac{p_1'(a^*)}{p_1(a^*)} = \frac{p_3'(a^*)}{p_3(a^*)} = \log\frac{1}{10} = -2.3026$, $w_1^* = w_3^*$ from equations (5.7), (5.9) and it is true in Example 5.3. By using the value of the optimal solutions, we have the equations (5.7), (5.8) as follows:

$$2.0270 = \lambda^V + 2.3026\lambda^{V_a},$$

$$11.0452 = \lambda^V - 0.4659\lambda^{V_a}.$$

Thus $\lambda^{V_a} = -3.2574 < 0$. Similarly, $\lambda^{V_a} = -3.9790 < 0$ for Example 5.5, therefore the signs of Examples 5.3, 5.5 can satisfy Theorem 5.5.

When $k = N$ and $k = N$, we discuss the concavity of $V(w, \cdot)$ in the model with the risk-averse principal, so (A.4) holds. We call this model Risk Averse Principal Model. N-MLRC is MLRC and N-CDFC is CDFC. The property that w_i^* is nondecreasing in i is very important in Economics, which means with the higher outcome the agent can get higher wage. We prove this property by a weaker condition which is W-stationary condition instead of the KKT conditions.

Proposition 5.6 (Risk Averse Principal Model) *Let $z^* = (w^*, a^*, w_1^*, w_2^*)$ be a feasible point of PA-MPEC. Suppose that (A.4) and (A.7) hold. Assume that the*

W-stationary condition holds at z^* and MLRC holds when $\lambda^{V_a} < 0$. Then

- *when $\lambda^{V_a} < 0$, w_i^* is nondecreasing in i for $\forall\, i \in \{1, \ldots, N\}$;*
- *when $\lambda^{V_a} = 0$, w_i^* and $x_i - w_i^*$ are nondecreasing in i.*

Proof. For the definition of the W-stationary conditions, the reader can refer to Definition 5.1. By (5.3),

$$-p_i(a^*)u'(x_i - w_i^*) + \lambda^V p_i(a^*)v'(w_i^*) - \lambda^{V_a} p_i'(a^*)v'(w_i^*) = 0, \forall i \in \{1, \ldots, N\},$$

which implies that

$$\frac{u'(x_i - w_i^*)}{v'(w_i^*)} = \lambda^V - \lambda^{V_a} \frac{p_i'(a^*)}{p_i(a^*)}, \forall i \in \{1, \ldots, N\}. \tag{5.10}$$

(a) By Propositions 4.1, MLRC for $\lambda^{V_a} < 0$, it follows from (5.10) that

$$\frac{u'(x_{i+1} - w_{i+1}^*)}{v'(w_{i+1}^*)} \geq \frac{u'(x_i - w_i^*)}{v'(w_i^*)}, \forall i \in \{1, \ldots, N-1\}.$$

If there exits $i \in \{1, \ldots, N-1\}$ such that $w_i^* > w_{i+1}^*$, then $x_i - w_i^* < x_{i+1} - w_{i+1}^*$ for $x_i < x_{i+1}$. Since v is strictly concave, v' is decreasing. Because u is concave, u' is decreasing. $u'(x_i - w_i^*) \geq u'(x_{i+1} - w_{i+1}^*)$ and $v'(w_i^*) < v'(w_{i+1}^*)$, so

$$\frac{u'(x_{i+1} - w_{i+1}^*)}{v'(w_{i+1}^*)} < \frac{u'(x_i - w_i^*)}{v'(w_i^*)}$$

which contradicts to (5.10). Then we have $w_i^* \leq w_{i+1}^*$, $\forall i \in \{1, \ldots, N-1\}$. Hence, w_i^* is nondecreasing in i.

(b) If $\lambda^{V_a} = 0$, it follows from (5.10) that

$$\frac{u'(x_{i+1} - w_{i+1}^*)}{v'(w_{i+1}^*)} = \frac{u'(x_i - w_i^*)}{v'(w_i^*)}, \forall i \in \{1, \ldots, N-1\}.$$

Similar to the proof in (a), we can prove that $w_i^* \leq w_{i+1}^*$, $\forall i \in \{1, \ldots, N-1\}$. Since v is strictly concave, v' is decreasing. $\frac{1}{v'(w_i^*)} \leq \frac{1}{v'(w_{i+1}^*)}$. Now we have $u'(x_i - w_i^*) \geq u'(x_{i+1} - w_{i+1}^*)$. Since u is concave, u' is decreasing. Thus $x_i - w_i^* \leq x_{i+1} - w_{i+1}^*$.

□

In order to prove the equivalence of PA and PA-MPEC, we need to get the concavity of $V(w^*, \cdot)$.

Lemma 5.7 *Suppose that (A.1),(A.2) and (A.8) hold.*

(a) Suppose that (A.6) holds.

- *If w_i^* is nondecreasing in i for $\forall\, i \in \{1, \ldots, k\}$ and nonincreasing in i for $\forall\, i \in \{k, \ldots, N\}$, k-CDFC is satisfied, then $V(w^*, \cdot)$ is concave in a.*
- *If w_i^* is nonincreasing in i for $\forall\, i \in \{1, \ldots, k\}$ and nondecreasing in i for $\forall\, i \in \{k, \ldots, N\}$, k-OCDFC is satisfied, then $V(w^*, \cdot)$ is concave in a.*

(b) If $w_i^* = w_{i+1}^*$, $\forall i \in \{1, \ldots, N-1\}$, then $V(w^*, \cdot)$ is concave in a.

Proof. The agent's expected payoff can be rewritten as

$$V = \sum_{i=1}^{N} v(w_i)p_i(a) - c(a) = \sum_{i=1}^{N} \Delta_i [\sum_{j=i}^{N} p_j(a)] - c(a),$$

where

$$\Delta_i = \begin{cases} v(w_i) - v(w_{i-1}), & i > 1, \\ v(w_1), & i = 1. \end{cases}$$

Therefore, $V_{aa} = \sum_{i=2}^{N} \Delta_i [\sum_{j=i}^{N} p_j''(a)] - c''(a)$.

(a) Because $w_i^* \leq w_{i+1}^*$, $\forall i \in \{1, \ldots, k-1\}$, $w_i^* \geq w_{i+1}^*$, $\forall i \in \{k, \ldots, N-1\}$ and v is increasing, each of the Δ_i terms is nonnegative for $i \in \{2, \ldots, k\}$ and nonpositive for $i \in \{k+1, \ldots, N\}$. The function $c(a)$ is convex in a, so $c''(a) \geq 0$. By k-CDFC,

$$\sum_{j=i}^{N} p_j''(a) = [1 - F_{i-1}(a)]'' = \begin{cases} \leq 0, \ \forall i \in \{2, \ldots, k\}, \\ \geq 0, \ \forall i \in \{k+1, \ldots, N\}. \end{cases}$$

Therefore, $V_{aa}(w^*, a) \leq 0, \forall a \in A$. Similarly, if $w_i^* \geq w_{i+1}^*$, $\forall i \in \{1, \ldots, k-1\}$, $w_i^* \leq w_{i+1}^*$, $\forall i \in \{k, \ldots, N-1\}$ and k-OCDFC holds, then $V_{aa}(w^*, a) \leq 0, \forall a \in A$. Thus $V(w^*, \cdot)$ is concave in a.

(b) Because $w_i^* = w_{i+1}^*$, $\forall i \in \{1, \ldots, N-1\}$,

$$\Delta_i = \begin{cases} 0, & i > 1, \\ v(w_1), & i = 1. \end{cases}$$

Then $V = v(w_1)[\sum_{j=1}^{N} p_j(a)] - c(a) = v(w_1) - c(a)$.

Therefore, $V_{aa} = -c''(a)$. The function $c(a)$ is convex in a, so $c''(a) \geq 0$. Hence, $V_{aa}(w^*, a) \leq 0, \forall a \in A$. Thus $V(w^*, \cdot)$ is concave in a.

□

Now we can prove the equivalence of PA and PA-MPEC in the following theorem for both Risk Neutral Principal Model and Risk Averse Principal Model.

Theorem 5.8 *Suppose that (A.1), (A.2), (A.6)-(A.8) hold, (A.5) holds for Risk Neutral Principal Model and (A.4) holds for Risk Adverse Principal Model.*

(I) Let $z^* = (w^*, a^*, b_1^*, b_2^*)$ be a local (global) optimal solution of PA-MPEC. We assume that PA-MPEC LICQ or PA-MPEC GMFCQ holds at z^*. For Risk Neutral Principal Model, assume that there exists $k \in \{1, \ldots, N\}$ such that k-MLRC and k-OCDFC hold when $\lambda^{V_a} > 0$; k-MLRC and k-CDFC hold when $\lambda^{V_a} < 0$. For Risk Adverse Principal Model, assume that MLRC and CDFC hold when $\lambda^{V_a} < 0$; CDFC holds when $\lambda^{V_a} = 0$. Then (w^*, a^*) is a local (global) optimal solution to PA.

(II) Let (w^*, a^*) be a local (global) optimal solution of PA. Suppose that the W-stationary condition holds at any feasible solution (w', a', b_1', b_2') to PA-MPEC. For Risk Neutral Principal Model, assume that there exists $k \in \{1, \ldots, N\}$ such that k-MLRC and k-OCDFC hold when $\lambda^{V_a} > 0$; k-MLRC and k-CDFC hold when $\lambda^{V_a} < 0$. For Risk Adverse Principal Model, assume that MLRC and CDFC hold when $\lambda^{V_a} < 0$; CDFC holds when $\lambda^{V_a} = 0$. Then there exists a unique multiplier (b_1^*, b_2^*) to the agent's problem

$$\max_{a \in A} V(w^*, a)$$

at a^* such that $z^* = (w^*, a^*, b_1^*, b_2^*)$ is a local (global) optimal solution of PA-MPEC.

Proof. (I) Because $z^* = (w^*, a^*, b_1^*, b_2^*)$ is the local (global) optimal solution of PA-MPEC and PA-MPEC LICQ or PA-MPEC GMFCQ holds at z^*, the W-stationary condition holds at z^* from Theorem 5.3. If k-MLRC is satisfied when $\lambda^{V_a} \neq 0$, then we can use Proposition 5.5 to get the result that the w_i^* is nondecreasing in i for $\forall i \in \{1, \ldots, k\}$ and nonincreasing in i for $\forall i \in \{k, \ldots, N\}$ when $\lambda^{V_a} < 0$, the w_i^* is nonincreasing in i for $\forall i \in \{1, \ldots, k\}$ and nondecreasing in i for $\forall i \in \{k, \ldots, N\}$ when $\lambda^{V_a} > 0$ with (A.5) for Risk Neutral Principal Model. We can also use Proposition 5.6 to get the result that the w_i^* is nondecreasing in i for $\forall i \in \{1, \ldots, N\}$ when $\lambda^{V_a} \leq 0$ with (A.4) for Risk Adverse Principal Model when MLRC is satisfied with $\lambda^{V_a} < 0$. By Lemma 5.7, $V(w^*, \cdot)$ is concave.

We now prove that (w^*, a^*) is an optimal solution of PA. Since (w^*, a^*, b_1^*, b_2^*) is a local optimal solution of PA-MPEC, the KKT conditions for maximizing $V(w^*, a)$ subject to $\underline{a} \leq a \leq \overline{a}$ hold at $a = a^*$ with multipliers b_1^*, b_2^*. Since $V(w^*, \cdot)$ is concave and the constraint functions are linear, from the standard linear programming theory (see [2,10]) a^* is a maximizer of the agent's problem when the contract is w^*. It is easy to see that any feasible solution of PA is a feasible solution of PA-MPEC. Hence by the optimality of (w^*, a^*, b_1^*, b_2^*), we have $U(w^*, a^*) \geq U(w', a')$. Thus (w^*, a^*) is a local (global) optimal solution of PA.

(II) We suppose that the W-stationary condition holds at any feasible solution (w', a', b_1', b_2') to PA-MPEC, k-MLRC is satisfied when $\lambda^{V_a} \neq 0$. We can use Proposition 5.5 to get the result that the w_i' is nondecreasing in i for $\forall i \in \{1, \ldots, k\}$ and w_i' is nonincreasing in i for $\forall i \in \{k, \ldots, N\}$ when $\lambda^{V_a} < 0$, the w_i' is nonincreasing in i for $\forall i \in \{1, \ldots, k\}$ and nondecreasing in i for $\forall i \in \{k, \ldots, N\}$ when $\lambda^{V_a} > 0$ with (A.5) for Risk Neutral Principal Model. We can also use Proposition 5.6 to get the result that the w_i' is nondecreasing in i for $\forall i \in \{1, \ldots, N\}$ when $\lambda^{V_a} \leq 0$ with (A.4) for Risk Adverse Principal Model when MLRC is satisfied with $\lambda^{V_a} < 0$. By Lemma

5.7, $V(w', \cdot)$ is concave in a.

We now prove that (w^*, a^*, b_1^*, b_2^*) is an optimal solution of PA-MPEC. Let (w', a', b_1', b_2') be a feasible solution to PA-MPEC. Since $V(w', \cdot)$ is a concave function in a, $a' \in \arg\max_{a \in A} V(w', a)$ which means (w', a') is the feasible solution to PA. By the optimality of (w^*, a^*) and the feasibility of (w', a'), we have

$$U(w^*, a^*) \geq U(w', a').$$

Since (w^*, a^*) is a local (global) optimal solution to PA and there exists a unique pair of multipliers (b_1^*, b_2^*) such that the KKT conditions to hold at a^*, (w^*, a^*, b_1^*, b_2^*) is feasible to PA-MPEC. Hence, (w^*, a^*, b_1^*, b_2^*) is a local (global) optimal solution of PA-MPEC.

□

Let $z^* = (w^*, a^*, b_1^*, b_2^*)$ be a local (global) optimal solution of PA-MPEC. By Theorem 5.3, the W-stationary holds at z^* when PA-MPEC GMFCQ holds at z^* and the S-stationary holds at z^* when PA-MPEC LICQ holds at z^*. By the definitions for PA-MPEC LICQ in Definition 5.2 and PA-MPEC GMFCQ in Definition 5.3, we have the following corollaries for each case.

For Case 1:

Corollary 5.9 *Suppose that (A.1), (A.2), (A.6)-(A.8) hold, (A.5) holds for Risk Neutral Principal Model and (A.4) holds for Risk Adverse Principal Model. Let* $z^* = (w^*, a^*, b_1^*, b_2^*)$ *be a local (global) optimal solution of PA-MPEC with* $\underline{a} < a^* < \overline{a}$ *and* $b_1^* = b_2^* = 0$. *We assume that one of the following conditions holds.*

(a) $\nabla V(w^*, a^*), \nabla V_a(w^*, a^*)$ *are linearly independent.*

(b) $\nabla V_a(w^*, a^*) \neq \Theta_{N+1}$ *and* $\exists\, d \in R^{N+1}$ *such that*

$$\nabla V(w^*, a^*)^\top d > 0,$$

$$\nabla V_a(w^*, a^*)^\top d = 0.$$

For Risk Neutral Principal Model, assume that there exists $k \in \{1, \ldots, N\}$ *such that k-MLRC and k-OCDFC hold when* $\lambda^{V_a} > 0$*; k-MLRC and k-CDFC hold when* $\lambda^{V_a} < 0$*. For Risk Adverse Principal Model, assume that MLRC and CDFC hold when* $\lambda^{V_a} < 0$*; CDFC holds when* $\lambda^{V_a} = 0$*. Then* (w^*, a^*) *is a local (global) optimal solution to PA.*

For Case 2:

Corollary 5.10 *Suppose that (A.1), (A.2), (A.6)-(A.8) hold, (A.5) holds for Risk Neutral Principal Model and (A.4) holds for Risk Adverse Principal Model. Let* $z^* = (w^*, a^*, b_1^*, b_2^*)$ *be a local (global) optimal solution of PA-MPEC with* $a^* = \underline{a}$ *and* $b_1^* = b_2^* = 0$*.*

(a) Assume that $\nabla_w V(w^*, a^*)$*,* $\nabla_w V_a(w^*, a^*)$ *are linearly independent. For Risk Neutral Principal Model, assume that there exists* $k \in \{1, \ldots, N\}$ *such that k-MLRC and k-OCDFC hold when* $\lambda^{V_a} > 0$*. For Risk Adverse Principal Model, assume that CDFC holds when* $\lambda^{V_a} = 0$*. Then* (w^*, a^*) *is a local (global) optimal solution to PA.*

(b) We assume that one of the following conditions holds.

(I) $\exists\, d \in R^N, d_{N+1}, d_{N+2} \in R$ *such that*

$$\nabla_w V(w^*, a^*)^\top d + V_a(w^*, a^*) d_{N+1} \geq 0,$$

$$\nabla_w V_a(w^*, a^*)^\top d + V_{aa}(w^*, a^*) d_{N+1} + d_{N+2} = 0$$

and either $d_{N+2} > 0, d_{N+1} \geq 0$ *or* $d_{N+1} > 0, d_{N+2} \geq 0$.

(II) $\nabla_w V(w^*, a^*) \neq \Theta_N$.

(III) $\nabla V_a(w^*, a^*) \neq \Theta_{N+1}$ *and* $\exists\, d \in R^{N+1}$ *such that*

$$\nabla V(w^*, a^*)^\top d > 0,$$

$$\nabla V_a(w^*, a^*)^\top d = 0.$$

For Risk Neutral Principal Model, assume that there exists $k \in \{1, \ldots, N\}$ *such that k-MLRC and k-OCDFC hold when* $\lambda^{V_a} > 0$*; k-MLRC and k-CDFC hold when* $\lambda^{V_a} < 0$. *For Risk Adverse Principal Model, assume that MLRC and CDFC hold when* $\lambda^{V_a} < 0$*; CDFC holds when* $\lambda^{V_a} = 0$. *Then* (w^*, a^*) *is a local (global) optimal solution to PA.*

For Case 3:

Corollary 5.11 *Suppose that (A.1), (A.2), (A.6)-(A.8) hold, (A.5) holds for Risk Neutral Principal Model and (A.4) holds for Risk Adverse Principal Model. Let* $z^* = (w^*, a^*, b_1^*, b_2^*)$ *be a local (global) optimal solution of PA-MPEC with* $a^* = \underline{a}$ *and* $b_1^* > 0, b_2^* = 0$. *Assume that* $\nabla_w V(w^*, a^*) \neq \Theta_N$*, no condition is required to hold for Risk Neutral Principal Model and CDFC holds for Risk Adverse Principal Model, then* (w^*, a^*) *is a local (global) optimal solution to PA.*

For Case 4:

Corollary 5.12 *Suppose that (A.1), (A.2), (A.6)-(A.8) hold, (A.5) holds for Risk Neutral Principal Model and (A.4) holds for Risk Adverse Principal Model. Let* $z^* = (w^*, a^*, b_1^*, b_2^*)$ *be a local (global) optimal solution of PA-MPEC with* $a^* = \overline{a}$ *and* $b_1^* = b_2^* = 0$.

(a) Assume that $\nabla_w V(w^*, a^*)$, $\nabla_w V_a(w^*, a^*)$ *are linearly independent. For Risk Neutral Principal Model, assume that there exists* $k \in \{1, \ldots, N\}$ *such that k-MLRC and k-CDFC hold when* $\lambda^{V_a} < 0$. *For Risk Adverse Principal Model, assume that MLRC and CDFC hold when* $\lambda^{V_a} < 0$; *CDFC holds when* $\lambda^{V_a} = 0$. *Then* (w^*, a^*) *is a local (global) optimal solution to PA.*

(b) We assume that one of the following conditions holds.

(I) $\exists\, d \in R^N, d_{N+1}, d_{N+3} \in R$ *such that*

$$\nabla_w V(w^*, a^*)^\top d + V_a(w^*, a^*) d_{N+1} \geq 0,$$

$$\nabla_w V_a(w^*, a^*)^\top d + V_{aa}(w^*, a^*) d_{N+1} - d_{N+3} = 0$$

and either $d_{N+3} > 0, d_{N+1} \leq 0$ *or* $d_{N+1} < 0, d_{N+3} \geq 0$.

(II) $\nabla_w V(w^*, a^*) \neq \Theta_N$.

(III) $\nabla V_a(w^*, a^*)^T \neq \Theta_{N+1}$ *and* $\exists\, d \in R^{N+1}$ *such that*

$$\nabla V(w^*, a^*)^\top d > 0,$$

$$\nabla V_a(w^*, a^*)^\top d = 0.$$

For Risk Neutral Principal Model, assume that there exists $k \in \{1, \ldots, N\}$ such that k-MLRC and k-OCDFC hold when $\lambda^{V_a} > 0$; k-MLRC and k-CDFC hold when $\lambda^{V_a} < 0$. For Risk Adverse Principal Model, assume that MLRC and CDFC hold when $\lambda^{V_a} < 0$; CDFC holds when $\lambda^{V_a} = 0$. Then (w^*, a^*) is a local (global) optimal solution to PA.

For Case 5:

Corollary 5.13 *Suppose that (A.1), (A.2), (A.6)-(A.8) hold, (A.5) holds for Risk Neutral Principal Model and (A.4) holds for Risk Adverse Principal Model. Let $z^* = (w^*, a^*, b_1^*, b_2^*)$ be a local (global) optimal solution of PA-MPEC with $a^* = \overline{a}$ and $b_1^* = 0, b_2^* > 0$. Assume that $\nabla_w V(w^*, a^*) \neq \Theta_N$, no condition is required to hold for Risk Neutral Principal Model and CDFC holds for Risk Adverse Principal Model, then (w^*, a^*) is a local (global) optimal solution to PA.*

Corollaries 5.9, 5.10, 5.11, 5.12 and 5.13 give the sufficient conditions for the PA-MPEC approach to be valid for each case.

Case 1: The optimal solution is in the interior, so the PA-MPEC approach in this case is actually the first order approach. In the proof of Rogerson [20], the reason that DRPA is used to prove the sufficient condition is to avoid the positive sign of λ^{V_a}. We also can not solve the problem when $\lambda^{V_a} > 0$ in Risk Averse Principal Model. However, there is no restriction on the signs of λ^{V_a} in Risk Neutral Principal Model and we do not have to assume (A.9) and (A.10).

Case 2: The optimal solution (w^*, a^*) is at the lower boundary when $V_a(w^*, a^*) = 0$. In (a) of Corollary 5.10, PA-MPEC LICQ is required, such that S-stationary holds at z^* which is equivalent to the standard KKT conditions from Theorem 5.2. In (b) of Corollary 5.10, PA-MPEC GMFCQ is required, such that M-stationary holds at z^* which is weaker than the standard KKT conditions because M-stationary is weaker than S-stationary by Definition 5.1 and S-stationary is equivalent to the standard KKT conditions by Proposition 5.2.

Case 3: The optimal solution (w^*, a^*) is at the lower boundary when $V_a(w^*, a^*) < 0$. With $V_a(w^*, a^*) < 0$, $V(w, a)$ becomes smaller when a is bigger. That is why $a^* = \underline{a}$. The constraint qualifications PA-MPEC LICQ and PA-MPEC GMFCQ are the same here, so S-stationary (the standard KKT conditions) holds at z^*. Because $\lambda^{V_a} = 0$ in the S-stationary condition, there is no need to assume MLRC or k-MLRC. In Risk Averse Principal Model, k-CDFC is not required when $\lambda^{V_a} = 0$.

Case 4: The optimal solution (w^*, a^*) is at the upper boundary when $V_a(w^*, a^*) = 0$. In (a) of Corollary 5.12, PA-MPEC LICQ is required such that S-stationary holds at z^*. In (b) of Corollary 5.12, PA-MPEC GMFCQ is required, such that M-stationary holds at z^* which is weaker than the standard KKT conditions because M-stationary is weaker than S-stationary by Definition 5.1 and S-stationary is equivalent to the standard KKT conditions by Proposition 5.2.

Case 5: The optimal solution (w^*, a^*) is at the upper boundary when $V_a(w^*, a^*) > 0$. With $V_a(w^*, a^*) > 0$, $V(w, a)$ becomes bigger when a is bigger. That is why $a^* = \overline{a}$. The others are the same as Case 3.

Now we use the PA-MPEC approach to solve Example 3.1 in Chapter 3. We can also take the probability functions in Example 3.1 to show that all the constraint qualifications in Case 4 are satisfied.

Solution. Using the PA-MPEC approach, we can solve Example 3.1 when $a \in A = [0.1, 2]$. By Optimization Toolbox in Matlab, we can find the optimal solution is $(w_1^*, w_2^*, a^*, b_1^*, b_2^*) = (7.1391, 71.2791, 2, 0, 0)$ with $V(w^*, a^*) = V^* = 5$ and $U(w^*, a^*) = 207.2559$. $\nabla V = (0.0468, 0.0444, 0)$ and $\nabla V_a = (-0.0324, 0.0103, -0.6931)$ can satisfy all the constraint qualifications as follows:

(a) Obviously.

(b)

(I) $d = (d_1, d_2)$, $\exists d_1 = 0, d_2 > 0, d_{N+1} > 0, d_{N+3} = 0$ such that

$$-324d_1 + 103d_2 - 6931d_{N+1} = 0,$$

$$468d_1 + 444d_2 - 0 \geq 0.$$

(II) Obviously.

(III) $d = (d_1, d_2, d_3)$, $\exists d_1 > 0, d_2 > 0, d_3 = 0$ and $\frac{d_1}{d_2} = \frac{103}{324}$, such that

$$-324d_1 + 103d_2 = 0,$$

$$468d_1 + 444d_2 > 0.$$

From Example 4.4, Example 3.1 satisfy MLRC and CDFC.

In all, Example 3.1 can satisfy all the conditions in Corollary 5.12 when $k = N = 2$ which is required by the S-stationary condition.

Chapter 6

Conclusion

In this thesis we study the PA as follows:

$$
\begin{aligned}
\text{(PA)} \quad \max_{w,a} \quad & U(w,a) \\
\text{s.t.} \quad & V(w,a) \geq V^*, \\
& a \in \arg\max_{a' \in A} V(w,a').
\end{aligned}
$$

Because $a \in \arg\max_{a' \in A} V(w,a')$ is difficult to treat, the economists use the first-order approach which replaces the PA problem by the RPA problem:

$$
\begin{aligned}
\text{(RPA)} \quad \max_{w,a} \quad & U(w,a) \\
\text{s.t.} \quad & V(w,a) \geq V^*, \\
& V_a(w,a) = 0, \\
& a \in A.
\end{aligned}
$$

The MLRC and CDFC are the most important conditions to prove that the first-order approach is valid. In the proof, we need the DRPA problem:

$$
\begin{aligned}
\text{(DRPA)} \quad \max_{w,a} \quad & U(w,a) \\
\text{s.t.} \quad & V(w,a) \geq V^*, \\
& V_a(w,a) \geq 0, \\
& a \in A.
\end{aligned}
$$

It is involved to prove that the agent's payoff function $V(w,a)$ is concave in a at w^* when (w^*, a^*) solves DRPA under MLRC and CDFC. Thus the stationary point a^* is the global maximum of $V(w^*, a)$ on A. However, with the Assumptions (A.9), (A10), the global maximum a^* can not appear at the boundary of A. In order to make the proof easier to read, we change the structure of the proof and add more details. We also provide the proof that tells how to get wage nondecreasing in discrete case, which is not available in the references.

In this thesis, we use a new approach which replaces the PA problem by the following further relaxed program PA-MPEC:

$$
\begin{aligned}
\text{(PA-MPEC)} \quad \min_{w,a,b_1,b_2} \quad & -U(w,a) \\
\text{s.t.} \quad & V(w,a) \geq V^*, \\
& V_a(w,a) + b_1 - b_2 = 0, \\
& (b_1, b_2) \geq 0, \\
& (a - \underline{a}, \overline{a} - a) \geq 0, \\
& (b_1(a - \underline{a}), b_2(\overline{a} - a)) = 0
\end{aligned}
$$

to study the PA problem. If $V(w^*, \cdot)$ is concave in a, then the optimal solution of

PA-MPEC problem is the optimal solution of the PA problem. To get the concavity of $V(w, \cdot)$, we figure out a bunch of new conditions:

- The W-stationary condition for PA-MPEC in Definition 5.2 which is weaker than KKT conditions
- The PA-MPEC or PA-LICQ constraint qualifications for W-stationary condition to hold in Definitions 5.3, 5.4
- The k-MLRC (Definition 5.5), k-CDFC (Definition 5.6) and k-OCDFC (Definition 5.7) conditions which are more general than MLRC and CDFC.

Now we get the new conditions such that PA-MPEC is equivalent to PA explained in Theorem 5.9. From this theory, we can obtain some corollaries to get the optimal solutions of PA both in the interior and at the boundary of A which is more convenient than the first-order approach.

The PA-MPEC approach is good method to deal with the boundary case of the domain of the action a. My future work would focus on using the PA-MPEC approach to discuss the multitask principal-agent problem. Since for such problems, the boundary cases can not be avoided in general.

Bibliography

[1] K.J. Arrow and A.C. Enthoven, Quasi-concave programming. *Econometrica*. 29(1961), pp. 779-800.

[2] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming Theory and Algorithms*, John Wiley&Sons, New York, Second Edition, 1993.

[3] M. Brown, A new class of sufficient conditions for the first-order approach to the principal-agent problem, *Economics Letter*, 21(1986), pp. 7-11.

[4] E. Alvi, First-order approach to principal-agent problems: a generalization, *Geneva Papers on Risk and Insurance Theory*, 22(1997), pp. 59-65.

[5] A. Araujo and H. Moreira, A general lagrangian approach for nonconcave moral hazard problems, *Journal of Mathematical Economics*, 35(2001), pp. 17-39.

[6] A.V. Fiacco and J. Kyparisis, Convexity and concavity properties of the optimal value function in parametric nonlinear programming, *Journal of optimization theory and applications*, 48(1986), pp. 95-126.

[7] I. Jewitt, Justifying the first-order approach to principal-agent problems, *Econometrica*, 56(1988), pp. 1177- 1190.

[8] H. W. Kuhn and A. W. Tucker, Nonlinear programming, in Proceedings of the Second Berkeley Symposium on *Mathematical Statistics and Probability*, J. Neyman (ed.), University of California Press, Berkeley, California, 1951, pp. 481-492.

[9] Y. Lucet and J.J. Ye, Sensitivity analysis of the value function for optimization problems with variational inequality constraints. *SIAM Journal on Control and Optimization*, 41(2002), pp. 1315-1319.

[10] Z.Q. Luo, J.S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[11] O.L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York, 1969; SIAM, Philadelphia, 1994.

[12] P.R. Milgrom, Good news and bad news: representation theorems and applications, *Bell Journal of Economics*, 12(1981), pp. 380-391.

[13] J. Mirrlees, The optimal structure of authority and incentives within an organization, *Bell Journal of Economics*, 7(1976), pp. 105-131.

[14] J. Mirrlees, The implications of moral hazard for optimal insurance, seminar given at Conference held in honor of Karl Borch, Bergen, Norway, mimeo, 1979.

[15] J. Mirrlees, The theory of moral hazard and unobservable behavior: Part I, *Review of Economic Studies*, 66(1999), pp. 3-21.

[16] B.S. Mordukhovich, Generalized differential calculus for nonsmooth and set- valued mappings, *Journal of Mathematical Analysis and Applications*, 183(1994), pp. 250-288.

[17] M. M. Mäkelä, *Nonsmooth Optimization*, World Scientific Publishing Co., Singapore, 1992.

[18] J.V. Outrata, Optimality conditions for a class of mathematical programs with equilibrium constraints, *Mathematics of Operations Research*, 24(1999), pp. 627-644.

[19] J.V. Outrata, M. Kočvara and J. Zowe, *Nonsmooth Approach to Optimization Problem with Equilibrium Constraints: Theory, Application and Numerical Results,* Kluwer, Dordrecht, The Netherlands, 1998.

[20] W.P. Rogerson, The first-order approach to principal-agent problems, *Econometrica*, 53(1985), pp. 1357-1368.

[21] B. Salanié, *The Economics of Contracts*, The MIT Press, Cambridge, 1999.

[22] H.Scheel and S. Scholtes, Mathematical programs with complementarity constraints: stationary, optimality and sensitivity, *Mathematics of Operations Research*, 25(2000), pp. 1-22.

[23] W. Whitt, Uniform Conditional stochastic order, *Journal of Applied Probability*, 17(1980), pp. 112-123.

[24] Z. Xu, Supply chain contracts under asymmetric information, PhD Thesis (2007), Fudan University, China.

[25] J.J. Ye, Optimality conditions for optimization problems with complementarity constraints, *SIAM Journal on Optimization*, 9(1999), pp. 374-387.

[26] J.J. Ye, Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints, *SIAM Journal on Optimization*, 10(2000), pp. 943-962.

[27] J.J. Ye, Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints, *Journal of Mathematical Analysis and Applications*, 307(2005), pp. 350-369.

[28] J.J. Ye and X.Y. Ye, Necessary optimality conditions for optimization problems with variational inequality constraints, *Mathematics of Operations Research*, 22(1997), pp. 977-977.

[29] J.J.Ye, D.L. Zhu and Q.J. Zhu, Exact penalization and necessary optimality conditions for generalized bilevel programming problems, *SIAM Journal on Optimization*, 2(1997), pp. 481-507.