

The Mathematics of Principal-Agent Problem with Adverse Selection

by

Mojdeh Shadnam

B.Sc., University of Alzahra, 2007

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

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

© Mojdeh Shadnam, 2011

University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopying or other means, without the permission of the author.

The Mathematics of Principal-Agent Problem with Adverse Selection

by

Mojdeh Shadnam

B.Sc., University of Alzahra, 2007

Supervisory Committee

Dr. Martial Agueh, Co-Supervisor
(Department of Mathematics and Statistics)

Dr. Jane Ye, Co-Supervisor
(Department of Mathematics and Statistics)

Supervisory Committee

Dr. Martial Agueh, Co-Supervisor
(Department of Mathematics and Statistics)

Dr. Jane Ye, Co-Supervisor
(Department of Mathematics and Statistics)

ABSTRACT

This thesis studies existence and characterization of optimal solutions to the principal-agent problem with adverse selection for both discrete and continuous problems. The existence results are derived by the abstract concepts of differentiability and convexity. Under the Spence Mirrlees condition, we show that the discrete problem reduces to a problem that always satisfies the linear independence constraint qualification, while the continuum of type problem becomes an optimal control problem. We then use the Ellipsoid algorithm to solve the problem in the discrete and convex case. For the problem without the Spence Mirrlees condition, we consider different classes of constraint qualifications. Then we introduce some easy-to-check conditions to verify these constraint qualifications. Finally we give economic interpretations for several numerical examples.

Contents

Supervisory Committee	ii
Abstract	iii
Table of Contents	iv
Acknowledgements	vi
1 Introduction	1
2 Preliminaries	5
2.1 Some properties of measure spaces	5
2.2 Kuhn-Tucker optimality condition and constraint qualification	7
2.3 Convex programming problem	11
2.3.1 Subgradients	11
2.3.2 Ellipsoid method	11
2.4 Optimal control problem	12
3 Existence of Solutions	14
3.1 The principal's problem as an optimal control problem	15
3.2 Compactness	23
3.3 Existence result for a linear cost function	27
3.4 Existence of solutions for the general cost functions	32
3.5 Discrete problem	33
4 Solutions under the Spence Mirrlees Condition	35
4.1 One dimensional continuum of type problem	36
4.2 Discrete problem	47
5 Solutions without the Spence Mirrlees Condition	56

5.1	Discrete problem	56
5.2	Linear independence constraint qualification (LICQ)	57
5.3	Mangasarian-Fromovitz constraint qualification (MFCQ)	59
5.4	Other constraint qualifications and examples	65
6	Examples and Economic Interpretation	69
6.1	Example 1	69
6.2	Example 2	71
6.3	Example 3	72
7	Conclusion	74
	Bibliography	75

ACKNOWLEDGEMENTS

This research could not be conducted without the financial support of the Department of Mathematics and Statistics at the University of Victoria. The author is indebted to Drs. Martial Agueh and Jane Ye for their encouragement, supervision, patience and support that enabled her to develop a good understanding of the subject.

The author wishes to express her love and gratitude to her family for their understanding and endless love through the duration of her studies. Helpful discussions and assistance of Dr. Masoud Shadnam (Rouen Business School, France) is acknowledged.

Chapter 1

Introduction

The principal-agent problem is a problem which frequently occurs in economics (contract theory) and also political science [6, 18, 28]. It arises when a principal (e.g., firm, organization, employer, seller) assigns a task to an agent (e.g., worker, employee, buyer) through a contract. The goal of the principal is to assign the contract in a way that maximizes his profit while compensating the agent for performing the task required from him.

This problem has been discussed extensively in the literature of mathematics and economics [17, 30]. The theory has the potential to manage a wide range of problems under the same framework without imposing several technical assumptions. There are two types of principal-agent problems based on information asymmetry, that is, when one party of the contract has more or better information than the other: moral-hazard or hidden action (i.e., the case where the agent can take an action unobservable to the principal) and the adverse selection or hidden knowledge (i.e., the case where some relevant information of the agent is unobservable to the principal).

In this thesis, we focus on the principal-agent problem with adverse selection. As explained above, it is based on an economic contract that relates a principal and an agent such that some relevant characteristics (defined here by θ) of the agent is unobservable for the principal [17]. For example, the principal may not know how efficient or trustworthy the agent is. Given some unknown characteristics, the principal seeks to define his contract with the agent in a way that optimizes his profit. These kinds of problems have many applications in management [18].

In this thesis, without loss of generality, we assume that the principal is the owner of a restaurant and his customers are the agents. The asymmetric information in this case is the customer's taste which is not known to the owner of the restaurant. In this

case, the only available information to the owner of the restaurant are the proportions of customers with specific taste-type, for example the probability of higher-taste customers, that of lower-taste customers, etc. Thus, the owner of the restaurant seeks to define a contract with his customers in a way that maximizes his profit.

In our model, θ represents the taste of customers which belongs to some bounded set $\Theta \subset R^n$. The customer with taste θ goes to the restaurant and orders a food with quality $q \in R_+^m$ and pays a monetary transfer $t \in R_+$ for the price of the food. Let $h(\theta, q)$ denote the satisfaction of the customer of type θ buying the food with quality q . Then the welfare or utility of this agent (defined here by “a”) is

$$U_a(\theta) = h(\theta, q(\theta)) - t(\theta).$$

Roughly speaking, $U_a(\theta)$ quantifies how much a customer with taste θ enjoys the food with quality q , knowing that he spends the amount t for it. If $C(q)$ represents the cost of producing the food with quality q , then the utility of the principal (defined here by “p”) is

$$U_p(\theta) = t(\theta) - C(q(\theta)).$$

Here $U_p(\theta)$ can be viewed as the profit that the restaurant’s owner makes in selling the food with quality q to the customer with taste θ .

We observe that even if the owner of the restaurant knows that the higher-taste customer is willing to pay more for a higher quality food, if he chooses to offer just expensive foods to earn more money, he may loose the lower-taste customer, whereas if he only offers cheap food, his business may not produce much profit. Therefore he should offer a mixture of high quality, expensive foods for higher-taste customers and a range of cheap foods for lower-taste customers. The difficulty here is that, if the owner of the restaurant simply offers a mixture of high quality and cheap foods with no other strategy, the higher-taste customer could hide his type and acts as a lower-taste, then choosing the food that is intended for the lower-taste customer and pay less money to earn more utility. Of course, this is not desirable to the owner of the restaurant since his primary goal is to earn more money by making the higher-taste customer choose a high quality food that is targeted for him. So, the challenge for the owner of the restaurant is to anticipate the customer’s choice so that each customer reveals its taste by choosing his favorite food that is targeted for him. Therefore, the principal’s utility $U_p(\theta)$ is subject to some constraints, called incentive compatible constraints, meaning that customer is given incentive to reveal his real

taste. Mathematically, the incentive compatible constraints can be represented as:

$$h(\theta, q(\theta)) - t(\theta) \geq h(\theta, q(\theta')) - t(\theta') \quad \forall \theta, \theta' \in \Theta.$$

So, the principal-agent problem can be formulated as follows:

$$\begin{aligned} (PA) \quad & \max_{q(\theta), t(\theta)} \int_{\Theta} (t(\theta) - C(q(\theta))) f(\theta) d\theta \\ & \text{s.t.} \quad h(\theta, q(\theta)) - t(\theta) \geq 0, \quad \forall \theta \in \Theta \quad (\text{IR}) \\ & \quad \quad h(\theta, q(\theta)) - t(\theta) \geq h(\theta, q(\theta')) - t(\theta'), \quad \forall \theta, \theta' \in \Theta \quad (\text{IC}) \end{aligned}$$

where $f(\theta) \in L^\infty(\Theta)$ is the probability density function representing the distribution of the customer tastes. The first set of constraints represents individual rationality constraint (IR for short) meaning that customers go to the restaurant only if they receive at least zero level of utility, otherwise the customers will choose to go to another restaurant. The second set of constraints is the incentive compatibility constraints (IC for short).

In the discrete case, when we have n customers, using the argument discussed earlier the problem can be formulated as:

$$\begin{aligned} (PA)_d \quad & \max_{q_i, t_i} \sum_{i=1}^n (t_i - C(q_i)) f_i \\ & \text{s.t.} \quad h(\theta_i, q_i) - t_i \geq 0, \quad \forall i = 1, \dots, n \quad (\text{IR}) \\ & \quad \quad h(\theta_i, q_i) - t_i \geq h(\theta_i, q_j) - t_j \quad \forall i, j = 1, \dots, n \quad (\text{IC}) \end{aligned}$$

Existence of solutions to problems (PA) and $(PA)_d$ as well as characterization of solutions have been one of the main issues in the past thirty years for many economists and mathematicians [9, 30]. The concept of the adverse selection in contract theory was first introduced and analyzed by Mussa and Rosen [22] in a model of nonlinear monopoly pricing. Maskin and Riley [20] addressed this problem from a different perspective, by using a graphical method. Rochet and Chone [25] extended the idea of Mussa and Rosen for the existence and characterization of solutions to multidimensional screening when h is linear in θ . They approached the problem in two different ways: the direct approach and the dual approach. After showing the difficulty of the direct approach, they rewrote the problem using an indirect utility function. Then, they used bunching, ironing and sweeping processes to characterize solutions. Carlier

[9] studied the general existence of the solution to the principal-agent problem when the principal is an employer and the agent is his employee. For the characterization of the solution to this problem, it is customary in the literature to require that the utility of the agent satisfies some technical condition, namely the Spence Mirrlees condition [6, 27]. This condition means that the marginal rate of substitution between quality and money is either increasing or decreasing with respect to the customer's taste. In fact the Spence Mirrlees condition allows to simplify the problem as much as possible, leading to a reduced problem which can be solved explicitly.

Although there are many utility functions that do not satisfy the Spence Mirrlees condition, very little is known about the solution to the principal-agent problem in these cases. One notable exception is the paper by Araujo and Moreira [3] that deals with the problem under a condition called U-shaped condition.

One of the goals here is to study the principal-agent problem in discrete case without the Spence Mirrlees condition. Precisely, this thesis has three principal objectives. The primal one is to adapt to problem (PA) and $(PA)_d$ the conditions imposed by Carlier [9] to obtain existence results to the model of employers and employees. The second objective is to characterize the solution to the principal-agent problem with adverse selection in both discrete and continuous cases when the Spence Mirrlees condition holds. The third and the last objective is to study the $(PA)_d$ problem without Spence Mirrlees condition.

The thesis is organized as follows. In Chapter 2, we recall some preliminary results on optimization, measure theory, convex programming and optimal control theory. In Chapter 3, we introduce some sufficient conditions that guarantee the existence of the solution to the principal-agent problem in both the discrete and continuum of type problems. In Chapter 4, we find the solution of the discrete and continuum of type problems when the utility of the agent satisfies the Spence Mirrlees condition. This is followed by the case without the Spence Mirrlees condition that we treat in Chapter 5. In Chapter 6, we present some examples, and use some existing softwares to numerically compute the solutions of the problems. Finally, Chapter 7 is reserved for the conclusion.

Chapter 2

Preliminaries

2.1 Some properties of measure spaces

In this section, we recall some well-known results in Lebesgue and Sobolev spaces.

Definition 2.1.1. (*L^p Space*)

Let X be an arbitrary measure space with a positive measure μ . If $1 \leq p \leq \infty$, we say that a complex measurable function f on X belongs to $L^p(X)$ if

$$\|f\|_p = \left(\int_X |f|^p d\mu \right)^{1/p}$$

is finite. We define also $\|f\|_\infty$ to be the essential supremum of $|f|$ and we let $L^\infty(\mu)$ consist of all f for which $\|f\|_\infty < \infty$. We call $\|f\|_p$ the L^p -norm of f .

Definition 2.1.2. (*Sobolev Space*)

Let $\Omega \subset \mathbb{R}^n$ be an open set and $1 \leq p \leq \infty$. The Sobolev space $W^{1,p}(\Omega)$ is defined as the space of real-valued functions $u \in L^p(\Omega)$ whose weak partial derivatives belong to $L^p(\Omega)$ for every $i = 1, \dots, n$. This space is equipped the following norm,

$$\|u\|_{W^{1,p}} = (\|u\|_{L^p}^p + \|\nabla u\|_{L^p}^p)^{1/p}, \quad \text{if } 1 \leq p < \infty$$

and

$$\|u\|_{W^{1,\infty}} = \max\{\|u\|_{L^\infty}, \|\nabla u\|_{L^\infty}\}, \quad \text{if } p = \infty.$$

Below are the definitions of continuous and compact embeddings which are used in the next theorem.

Definition 2.1.3. (*Continuous Embedding Vector Space*)

A normed vector space is continuously embedded in another normed vector space if the inclusion function between them is continuous.

Definition 2.1.4. (*Compact Embedding Topological Space*)

Let (X, d) be a topological space, and let A and B be subsets of X . We say that A is compactly embedded in B , if $A \subseteq \bar{A} \subseteq \text{int}(B)$, where \bar{A} denotes the closure of A , and $\text{int}(B)$ denotes the interior of B , and \bar{A} is compact.

The following theorem gives the embedding of Sobolev spaces into Lebesgue spaces. For a proof, we refer to [8].

Theorem 2.1.5. (*Rellich-Kondrachov Theorem*)

Let $\Omega \subseteq \mathbb{R}^n$ be an open and bounded Lipschitz domain. Set $p^* := \frac{np}{n-p}$ where $1 \leq p < n$. Then the Sobolev space $W^{1,p}(\Omega)$ is continuously embedded in the Lebesgue space $L^{p^*}(\Omega, \mathbb{R})$ and is compactly embedded in $L^q(\Omega)$ for every $1 \leq q < p^*$.

The next theorem will be used in Chapter 3 to obtain the existence result for the principal-agent problem. Recall that a multifunction $f : S \subset \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is a mapping from S to the collection of all subsets of \mathbb{R}^n . We say that f is closed or non-empty on S if for each x in S the set $f(x)$ is closed or nonempty. It is also said to be measurable if for every open subset C of \mathbb{R}^n the set

$$\{x \in S \mid f(x) \cap C \neq \emptyset\}$$

is Lebesgue measurable.

Theorem 2.1.6. (*Measurable Selection Theorem*)(see e.g., [12])

Let $f : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a multifunction and S be a subset of \mathbb{R}^m . If f is measurable, closed and nonempty on S , then there exists a measurable function $g : S \rightarrow \mathbb{R}^n$ such that $g(x)$ belongs to $f(x)$ for all x in S .

Definition 2.1.7. (*Lower Semi-continuity*)

Let f be a real-valued function on a topological space. If the level set $\{x : f(x) > \alpha\}$ is open for every real α , then f is said to be lower semi-continuous.

We end this section with Fatou's Lemma; see [26].

Lemma 2.1.8. (*Fatou's Lemma*)

If $f_n : X \rightarrow [0, \infty]$ is measurable, for each positive integer n , then

$$\int_X \left(\liminf_{n \rightarrow \infty} f_n \right) d\mu \leq \liminf_{n \rightarrow \infty} \int_X f_n d\mu.$$

2.2 Kuhn-Tucker optimality condition and constraint qualification

Lagrange multipliers play a crucial role in the study of constrained optimization problem with equality constraints. In particular, if x^* is a local minimum of an optimization problem with equality constraints, then x^* should be a feasible point. Moreover, the gradient of the objective function plus a linear combination of the gradient of the constraints should equal zero. Fritz John [14] developed this concept as the necessary optimality condition for the constrained optimization problem with the inequality constraint.

Theorem 2.2.1. (*Fritz John Necessary Optimality Condition*)

Suppose that f and g from R^n to R are differentiable functions and x^* is the local minimum of the following optimization problem,

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0 \quad i = 1, 2, \dots, n. \end{aligned}$$

Then, there exist scalars $\lambda_i, i = 0, 1, 2, \dots, n$ not all zero such that,

$$\lambda_0 \nabla f(x^*) + \sum_{i=1}^n \lambda_i \nabla g_i(x^*) = 0,$$

$$\lambda_i g_i(x^*) = 0, \quad i = 1, 2, \dots, n,$$

$$\lambda_i \geq 0, \quad i = 0, 1, \dots, n.$$

If $\lambda_0 = 0$, then there is no information about the objective function and thus the optimal solution. This is a result of the lack of technical properties named constraint qualifications. The constraint qualification ensures the positiveness of the Fritz John multiplier associated to the objective function (λ_0). Kuhn and Tucker [16] introduced

a condition that could guarantee a non-zero λ_0 in Fritz John's condition; this condition was then named, Kuhn Tucker necessary optimality condition (KKT condition).

Theorem 2.2.2. (*KKT Necessary Optimality Condition*)

Suppose that x^* is a local minimizer of the following minimization problem:

$$(P_0) \quad \begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & g_i(x) = 0, \quad i = 1, 2, \dots, p \\ & g_i(x) \leq 0, \quad i = p + 1, \dots, n. \end{aligned}$$

If x^* satisfies one of the constraint qualifications to be defined below, then there exist Lagrange multipliers λ_i for $i = 1, 2, \dots, n$ such that

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^n \lambda_i \nabla g_i(x^*) &= 0, \\ \lambda_i g_i(x^*) &= 0, \quad \lambda_i \geq 0 \quad \forall p + 1 \leq i \leq n. \end{aligned}$$

Let x^* be a local optimal solution of the optimization problem (P_0) . Define the index set of all active inequality constraints at x^* :

$$I(x^*) = \{i = p + 1, \dots, n : g_i(x^*) = 0\}.$$

Definition 2.2.3. (*Constraint Qualifications*)(see e.g., [19])

(i) *Linear Independence Constraint Qualification (LICQ):*

$$\sum_{i=1}^p \lambda_i \nabla g_i(x^*) + \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = 0 \text{ implies } \lambda_i = 0, \quad i \in I(x^*) \cup \{1, 2, \dots, p\}.$$

(ii) *Mangasarian–Fromowitz Constraint Qualification (MFCQ):* There exists a vector $d \in R^n$ such that

$$\begin{cases} \nabla g_i(x^*)^T d < 0, & \forall i \in I(x^*) \\ \nabla g_i(x^*)^T d = 0, & \forall i = 1, 2, \dots, p. \end{cases}$$

It is known that MFCQ is equivalent to the positively linearly independent constraint

qualification. That is,

$$\sum_{i=1}^p \lambda_i \nabla g_i(x^*) + \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = 0, \quad \lambda_i \geq 0, \quad i \in I(x^*)$$

implies that $\lambda_i = 0, \quad \forall i \in \{1, 2, \dots, p\} \cup I(x^*)$.

(iii) *Linear Constraints Qualification: All the constraint functions $g_i(x)$ are affine.*

There are many other constraint qualifications available, some of which are more useful than the others depending on the cases in consideration. Below is another set of constraint qualifications:

Definition 2.2.4. (*Quasinormality Constraint Qualification*)(see e.g., [5])

Consider the problem (P_0) defined in Theorem 2.2.2. A feasible point x^* is quasinormal if there are no scalars $\lambda_1, \dots, \lambda_n$ and no sequence $\{x^k\}$ such that

$$\begin{aligned} \sum_{i=1}^n \lambda_i \nabla g_i(x^*) &= 0, \\ \lambda_1, \dots, \lambda_n &\text{ are not all zero,} \\ \lambda_i &\geq 0, \quad i = p + 1, \dots, n, \\ \{x^k\} &\text{ converges to } x^*, \text{ for all } k \text{ and for all } i \text{ with } \lambda_i \neq 0, \lambda_i g_i(x^k) > 0. \end{aligned}$$

The next constraint qualification is more abstract and is related to the pseudoconvexity property of the constraint functions. First we give the meaning of pseudoconvex function.

Definition 2.2.5. (*Pseudoconvex Function*)

Let $D \subset R^n$ be an open set and $\hat{x} \in D$. A differentiable function $f : D \rightarrow R$ is pseudoconvex at \hat{x} if for any $x \in D$,

$$\nabla f(\hat{x}^T)(x - \hat{x}) \geq 0 \Rightarrow f(x) \geq f(\hat{x}).$$

f is pseudoconcave if and only if $-f$ is pseudoconvex.

Now, we define the Arrow-Hurwicz-Uzawa constraint qualification.

Definition 2.2.6. (*Arrow-Hurwicz-Uzawa Constraint Qualification*)[4]

A feasible point x^* of problem (P_0) satisfies the Arrow-Hurwicz-Uzawa Constraint

Qualification if the equality constraints are pseudoaffine (pseudoconvex and pseudoconcave) and there exists a vector $d \in R^n$ such that,

$$\begin{aligned} \nabla g_i(x^*)^T d &< 0, \text{ where } g_i \text{ is pseudoconcave at } x^* \text{ for } i \in I(x^*), \\ \nabla g_i(x^*)^T d &\leq 0, \text{ where } g_i \text{ is not pseudoconcave at } x^* \text{ for } i \in I(x^*), \\ \nabla g_i(x^*)^T d &= 0, \forall i = 1, 2, \dots, p. \end{aligned}$$

We end the subsection with one of the most abstract constraint qualification, the Abadie constraint qualification.

Definition 2.2.7. (*Abadie Constraint Qualification*)[1]

We say that the Abadie constraint qualification holds at x^* if the linearized cone is equal to the tangent cone at x^* . In other words Abadie constraint qualification holds at x^* if,

$$T(x^*) = L(x^*),$$

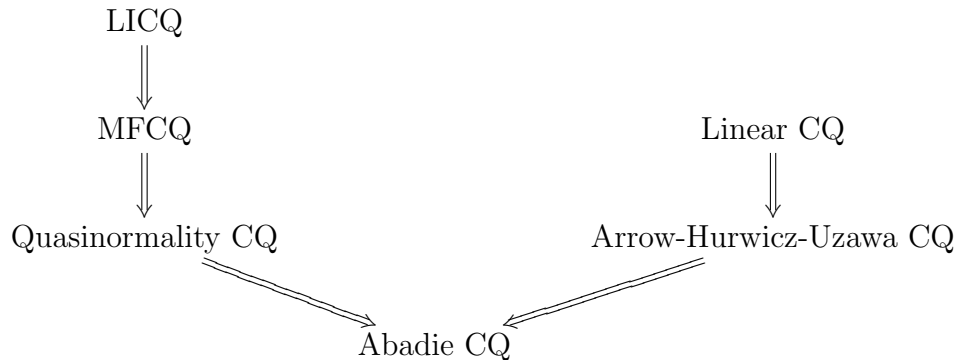
where the tangent cone is

$$T(x^*) := \{d \in R^n \mid \exists x^k \subseteq X, t_k \rightarrow 0 : x^k \rightarrow x^* \text{ and } \frac{x^k - x^*}{t_k} \rightarrow d\},$$

and X denotes the feasible set of problem (P_0) , and, the linearized cone is

$$L(x^*) = \{d \in R^n \mid \nabla g_i(x^*)d \leq 0 \text{ for } i \in I(x^*), \nabla g_i(x^*)d = 0 \text{ for } i = 1, \dots, p\}.$$

The relationships between the above constraint qualifications are given in the following charts



2.3 Convex programming problem

The concept of convexity plays an important role in optimization. One of the most important properties of the convex programming problem is that a local minimizer is also a global minimizer of the problem. As a result, the KKT condition is a necessary condition for convex programming problems.

There are different numerical methods to solve the convex programming problems in a finite dimensional space. Among them, cutting plane methods are very common, of which Ellipsoid [2] is one of the most popular ones.

2.3.1 Subgradients

Subgradients are the extension of usual gradients for the nondifferentiable functions.

Definition 2.3.1. *Let $f : D \rightarrow R$ be a convex function on an open convex set $D \subset R^n$. Then, a vector $d \in R^n$ is a subgradient of f at \hat{x} if,*

$$f(x) \geq f(\hat{x}) + d^T(x - \hat{x}), \quad \forall x \in D.$$

If f is not differentiable, then the subgradient may not be unique. We denote the set of all subgradients by

$$\partial f(\hat{x}) := \{d \in R^n : f(x) \geq f(\hat{x}) + d^T(x - \hat{x}), \quad \forall x \in D\}.$$

In the case when f is convex and differentiable at \hat{x}

$$\partial f(\hat{x}) = \{\nabla f(\hat{x})\}.$$

Every convex function has a subgradient at each of the points in its domain of definition. However, the subgradient may not be unique when the function is not differentiable. For example, $|x|$ is a convex function but non differentiable at 0; its subgradients at 0 is $[-1, 1]$. The concept of subgradients plays an important role in the following cutting plane method.

2.3.2 Ellipsoid method

Ellipsoid method is developed in 1970 for the first time by Shor, Nemirovski and Yudin [2]. This method is one of the practical numerical methods for solving nonlinear

convex programming problems. The basic idea of the Ellipsoid method is to generate an initial ellipse that contains the minimum. Although we do not know where the optimum point is, we can choose an ellipse large enough to ensure that the optimum point is contained inside it. After that, we cut the ellipse by a line passing through the center and use the gradient to determine which part of the ellipse contains the optimum point. Then we make a new ellipse with the smallest volume such that it contains one half of the initial ellipse that we have produced from the previous step. This procedure is then continuously repeated by cutting the previous ellipse and choosing the half which contains the optimum point. In each step the volume of the ellipse is decreased. The procedure will stop when the volume of the ellipse tends to zero.

2.4 Optimal control problem

An optimal control problem is the generalization of a calculus of variations problem. It can be used on a problem for which the classical calculus of variations is not applicable. It is also an important tool in solving continuous optimization problem of the form:

$$(OP) \quad \begin{aligned} & \max_{x(t), u(t)} \int_0^T F(x(t), u(t), t) dt \\ & \text{s.t.} \quad x'_i(t) = g_i(x(t), u(t), t), \quad i = 1, \dots, n \\ & \quad \quad u(t) \in \Omega, \end{aligned}$$

where $F : R^n \times R^m \times R \rightarrow R$, $g_i : R^n \times R^m \times R \rightarrow R$ and $x'_i(t)$ denotes the derivative of function $x_i(t)$. In such a problem the variables are divided into two classes, state and control variables where the state variables is governed by the first order differential equation. In the above problem $x(t) = (x_1(t), \dots, x_n(t)) \in R^n$ and $u(t) = (u_1(t), \dots, u_m(t)) \in R^m$ represent the state and control variables respectively, and $t \in [0, \infty)$ denotes the time and Ω is a given set in R^m . Assume that F , g_i , $\frac{\partial f}{\partial x_j}$ and $\frac{\partial g_i}{\partial x_j}$ are continuous with respect to all their arguments for all $i = 1, \dots, n$ and $j = 1, \dots, n$. The following theorem gives the necessary optimality conditions for the problem (OP).

Theorem 2.4.1. *(Pontryagin's Maximum Principle)(see e.g., [13, 15, 24])*
 Let $(x^*(t), u^*(t))$ be optimal for the problem (OP), then it is necessary that Pon-

Pontryagin's Maximum Principle holds. This means that there exist absolute continuous functions $\lambda(t) = (\lambda_1(t), \dots, \lambda_n(t))$, $0 \leq t \leq T$, such that

- for all $u \in \Omega$

$$H(x^*(t), u, \lambda(t), t) \leq H(x^*(t), u^*(t), \lambda(t), t)$$

where the Hamiltonian function H is defined as

$$H(x, u, \lambda, t) = F(x, u, t) + \sum_{i=1}^n \lambda_i g_i(x, u, t).$$

- except at the points of discontinuity of u^* ,

$$\frac{\partial H}{\partial x_i}(t, x^*(t), u^*(t), \lambda(t)) = -\lambda'_i(t), \quad i = 1, \dots, n.$$

- transversality conditions are satisfied, i.e.,

$$\lambda(T) = \lambda(0) = 0.$$

Moreover if H is a concave function in x and u , then the above Pontryagin's Maximum Principle is also sufficient for optimality [15].

Chapter 3

Existence of Solutions

In this chapter, we study the general existence of solutions to the principal-agent problem (*PA*). We will adapt the proof of Carlier [9] for our model of the monopolist and his customers.

In this model, θ represents the taste of a customer which is not detectable to the owner of the restaurant. We will assume that θ belongs to some open bounded convex subset Θ of R^p with C^1 boundary. We denote by $\bar{\Theta}$ the closure of set Θ . The utility of the agent (i.e., the customer) is given by,

$$U_a(\theta, q, t) = h(\theta, q) - t$$

where $h(\theta, q) : \bar{\Theta} \times R_+^m \rightarrow R$ denotes the satisfaction of the customer of type θ ordering a food with quality q and paying a price $t \in R_+$ for it. The utility of the principal (i.e., the owner of the restaurant) is

$$U_p(t, q) = t - C(q)$$

where $C(q)$ is the cost of producing food with quality q .

Our model can then be formulated as the following maximization problem:

$$\begin{aligned}
 (PA) \quad & \max_{q(\theta), t(\theta)} \int_{\Theta} (t(\theta) - C(q(\theta))) f(\theta) d\theta \\
 & \text{s.t.} \quad h(\theta, q(\theta)) - t(\theta) \geq 0, \quad \forall \theta \in \Theta \quad (\text{IR}) \\
 & \quad \quad h(\theta, q(\theta)) - t(\theta) \geq h(\theta, q(\theta')) - t(\theta'), \quad \forall \theta, \theta' \in \Theta \quad (\text{IC}).
 \end{aligned}$$

where $f(\theta) \in L^\infty(\Theta)$ is the probability of having a customer with taste θ and the essential infimum of f is positive.

Carlier in [9] studied the general existence of the principal-agent problem in the context of employers and employees. In this chapter, we adapt the hypotheses in Carlier's proof to our model which deals with a monopolist and his consumers (or, equivalently, the owner of a restaurant and his customers). Because of the similarities between these two models, most of the propositions in [9] are also applicable here. Before stating the main theorem, we first introduce some definitions and lemmas which will be useful in the proof of the existence theorem. First, we write the principal problem as an optimal control problem.

3.1 The principal's problem as an optimal control problem

In this section, we state some results that will help us to rewrite the principal problem as an optimal control problem. The notions of h-convexity and h-differentiability will be useful.

Definition 3.1.1. (*Implementability*)

- A contract is a pair of functions (q, t) from Θ to $R_+^m \times R_+$.
- A function $q : \Theta \rightarrow R_+^m$ is called implementable if there exists a $t : \Theta \rightarrow R_+$ such that the pair (q, t) is an incentive compatible contract (IC), i.e.,

$$h(\theta, q(\theta)) - t(\theta) \geq h(\theta, q(\theta')) - t(\theta'), \quad \forall (\theta, \theta') \in \Theta^2.$$

Definition 3.1.2. Suppose that (q, t) is a contract. The potential associated with (q, t) is the function denoted $U_{q,t} : \Theta \rightarrow R$ defined by,

$$U_{q,t}(\theta) = h(\theta, q(\theta)) - t(\theta).$$

Definition 3.1.3. (*h-convexity*)

A function $V : \Theta \rightarrow R \cup \{+\infty\}$ is h-convex if there exists a nonempty subset A of

$R_+^m \times R_+$ such that

$$V(\theta) = \sup_{(q,t) \in A} \{h(\theta, q) - t\}.$$

Definition 3.1.4. (*h-differentiability*)

Let $V : \Theta \rightarrow R \cup \{+\infty\}$. A vector $q \in R_+^m$ is called a *h-subgradient* of V at θ if

$$V(\theta') \geq V(\theta) + h(\theta', q) - h(\theta, q), \quad \forall \theta' \in \Theta.$$

The set of all *h-subgradients* of V at θ is called the *h-subdifferential* of V at θ and is denoted by $\partial^h V(\theta)$.

Moreover, V is said to be *h-subdifferentiable* at $\theta \in \Theta$ if $\partial^h V(\theta) \neq \emptyset$.

The following proposition states the relation between implementability and the notions of h-convexity and h-subdifferentiability.

Proposition 3.1.5. A function $q : \Theta \rightarrow R_+^m$ is implementable if and only if there exists some h-convex and h-subdifferentiable mapping $V : \Theta \rightarrow R$ such that $q(\theta) \in \partial^h V(\theta)$ for all $\theta \in \Theta$.

Proof. Let us first assume that q is implementable. By Definition 3.1.1, there exists t such that (q, t) is an incentive compatible contract. This means,

$$h(\theta, q(\theta)) - t(\theta) \geq h(\theta, q(\theta')) - t(\theta'), \quad \forall \theta, \theta' \in \Theta. \quad (\text{IC})$$

Let V be the potential associated with (q, t) , i.e., $V(\theta) = h(\theta, q(\theta)) - t(\theta)$ by Definition 3.1.2. From (IC) we have

$$V(\theta) = \sup_{\theta' \in \Theta} \{h(\theta, q(\theta')) - t(\theta')\}, \quad \forall \theta \in \Theta.$$

So, V is h-convex by Definition 3.1.3. Since $t(\theta') = h(\theta', q(\theta')) - V(\theta')$, we have,

$$\begin{aligned} V(\theta) &\geq h(\theta, q(\theta')) - t(\theta') \quad \forall \theta' \in \Theta, \\ &\geq V(\theta') + h(\theta, q(\theta')) - h(\theta', q(\theta')). \end{aligned}$$

This means that $q(\theta') \in \partial^h V(\theta'), \forall \theta' \in \Theta$.

Now, we will show the reverse implication. Assume that there exists some h-convex and h-subdifferentiable mapping V such that $q(\theta) \in \partial^h V(\theta), \forall \theta \in \Theta$. Define $t(\theta) =$

$h(\theta, q(\theta)) - V(\theta)$. Then, since we know that $q(\theta') \in \partial^h V(\theta')$, then for all $(\theta, \theta') \in \Theta^2$ we have

$$V(\theta) \geq V(\theta') + h(\theta, q(\theta')) - h(\theta', q(\theta'))$$

Substituting $V(\theta') = h(\theta', q(\theta')) - t(\theta')$ and $V(\theta) = h(\theta, q(\theta)) - t(\theta)$ yields

$$h(\theta, q(\theta)) - t(\theta) \geq h(\theta, q(\theta')) - t(\theta').$$

Thus, (q, t) is an incentive compatible contract. \square

In the sequel, by $\theta_2 > \theta_1$ we mean that $(\theta_2)_i \geq (\theta_1)_i$ for all $1 \leq i \leq n$, that is, the customer with taste θ_2 is a higher taste-customer than the customer with taste θ_1 .

From now on, we assume the following hypotheses.

H1. $h \in C^0(\bar{\Theta} \times R_+^m, R)$ and for every $q \in R_+^m$, $h(\cdot, q)$ is nondecreasing in θ .

H2. For every $(\theta, q) \in \Theta \times R_+^m$ the partial derivative $\frac{\partial h}{\partial \theta}(\theta, q)$ exists, and the map $\frac{\partial h}{\partial \theta}(\cdot, \cdot)$ is continuous with respect to both arguments. Moreover, for every compact subset K of $\Theta \times R_+^m$, there exists k such that for all $((\theta, q), (\theta', q)) \in K^2$

$$\left\| \frac{\partial h}{\partial \theta}(\theta, q) - \frac{\partial h}{\partial \theta}(\theta', q) \right\| \leq k \|\theta - \theta'\|.$$

H3. For every $M > 0$ there exists $r > 0$ such that for all $(\theta, q) \in \Theta \times R_+^m$

$$\|q\| \geq r \Rightarrow \sum_{i=1}^n \frac{\partial h}{\partial \theta_i}(\theta, q) \geq M.$$

Remark 3.1.6. *By Hypothesis (H1), we can conclude that*

$$V(\theta) := \sup_{(q,t) \in A} \{h(\theta, q) - t\}$$

is also nondecreasing in θ .

The following proposition gives the relation between h-convex and h-differentiable functions.

Proposition 3.1.7. *Let V from Θ to $(-\infty, +\infty]$ be h-convex. If K is a compact subset of Θ , $\delta > 0$ and $R > 0$ satisfy*

(i) $K + \delta \bar{B}(0, 1) \subset \Theta$

(ii) $|V(\theta)| \leq R, \quad \forall \theta \in K + \delta \bar{B}(0, 1)$, (where $\bar{B}(0, 1)$ is the closed unit ball of R^p)

then

1. V is h -subdifferentiable at every point of K .
2. There exists some positive constant $M(R, K, \delta)$ such that

$$\forall \theta \in K, \forall q \in \partial^h V(\theta), \quad \|q\| \leq M(R, K, \delta).$$

Proof. Step 1. Let us fix $\theta_0 \in K$. Since V is h -convex, (i) and (ii) imply that there exists a sequence (q_n, t_n) of elements of $R_+^m \times R_+$ such that

$$\forall \theta \in \Theta, \quad \forall n, \quad V(\theta) \geq h(\theta, q_n) - t_n, \quad (3.1)$$

$$\lim_{n \rightarrow +\infty} h(\theta_0, q_n) - t_n = V(\theta_0) \quad \text{and} \quad |h(\theta, q_n) - t_n| \leq R, \quad \forall n, \forall \theta \in K + \delta \bar{B}(0, 1). \quad (3.2)$$

For all $u \in \bar{B}(0, 1)$ taking $\theta = \theta_0 + \delta u$, (ii) implies that $|h(\theta_0 + \delta u, q_n) - t_n| \leq R$. This equation and (3.2) yield

$$2R \geq h(\theta_0 + \delta u, q_n) - h(\theta_0, q_n), \quad \forall n, \forall u \in \bar{B}(0, 1).$$

Let us show that (q_n) is bounded. If it is not, there exists a (non-relabelled) subsequence such that $\lim_{n \rightarrow +\infty} \|q_n\| = +\infty$. Hence,

$$\frac{2R}{\delta} \geq \frac{h(\theta_0 + \delta u, q_n) - h(\theta_0, q_n)}{\delta},$$

which yields

$$\frac{2R}{\delta} \geq \int_0^1 \frac{\partial h}{\partial \theta}(\theta_0 + t\delta u, q_n) \cdot u dt.$$

Taking u to be the unit vector of R^p with each component equal to $p^{-1/2}$, we obtain

$$\frac{2R}{\delta} \geq p^{-1/2} \sum_{i=1}^n \int_0^1 \frac{\partial h}{\partial \theta_i}(\theta_0 + t\delta u, q_n) dt.$$

Pick $M > n$, then by (H3) there exists some $r > 0$ such that for all $(\theta, q) \in \Theta \times R_+^m$

$$\|q\| > r \Rightarrow \sum_{i=1}^n \frac{\partial h}{\partial \theta_i}(\theta, q) \geq M.$$

But since $\lim_{n \rightarrow \infty} \|q_n\| = +\infty$, then for some large value of n we have $\|q_n\| > r$. So,

$\sum_{i=1}^n \frac{\partial h}{\partial \theta_i}(\theta_0 + t\delta u, q_n) \geq M$. This implies

$$\frac{2R}{\delta} \geq p^{-\frac{1}{2}} M \geq p^{-\frac{1}{2}} n.$$

Then, let $n \rightarrow +\infty$ gives a contradiction. Therefore, (q_n) is bounded. Up to a subsequence, we may assume that (q_n) converges to $\bar{q} \in R_+^m$, so that t_n is also convergent with limit $\bar{t} = h(\theta_0, \bar{q}) - V(\theta_0)$. Passing to the limit in equation (3.1) yields

$$\forall \theta \in \Theta, \quad V(\theta) \geq h(\theta, \bar{q}) - \bar{t} = V(\theta_0) + h(\theta, \bar{q}) - h(\theta_0, \bar{q}).$$

This means $\bar{q} \in \partial^h V(\theta_0)$, which completes the proof of (1).

Step 2. Let $\theta_0 \in K$ and $q \in \partial^h V(\theta_0)$, then for every $u \in \bar{B}(0, 1)$,

$$V(\theta_0 + \delta u) \geq V(\theta_0) + h(\theta_0 + \delta u, q) - h(\theta_0, q).$$

With (ii) we obtain,

$$\begin{aligned} 2R &\geq V(\theta_0 + \delta u) - V(\theta_0) \\ &\geq h(\theta_0 + \delta u, q) - h(\theta_0, q). \end{aligned}$$

Using the same argument as in the previous step, we have

$$\frac{2R}{\delta} \geq \frac{h(\theta_0 + \delta u, q) - h(\theta_0, q)}{\delta},$$

and then

$$\frac{2R}{\delta} \geq \int_0^1 \frac{\partial h}{\partial \theta}(\theta_0 + t\delta u, q) \cdot u dt,$$

which yields

$$\frac{2R}{\delta} \geq p^{-1/2} \sum_{i=1}^n \int_0^1 \frac{\partial h}{\partial \theta}(\theta_0 + t\delta u, q) dt,$$

where u is the unit vector defined in step (1). Using the similar argument as step 1 and (H3), there exists $M = M(R, K, \delta)$ such that

$$\forall \theta \in K, \forall q \in \partial^h V(\theta), \quad \|q\| \leq M.$$

□

Remark 3.1.8. As a result of the proposition 3.1.7, if V is h -convex and locally bounded, then the set valued map $\partial^h V(\cdot)$ takes non-empty compact values.

Definition 3.1.9. V is called locally semi-convex if and only if for all convex compact subsets K of Θ there exists $\lambda > 0$ such that $V_\lambda(\cdot) := V(\cdot) + \lambda \|\cdot\|^2$ is convex in K . Any λ which has this property is called a semi-convexity modulus of V in K .

The next proposition is one of the main propositions that will be used in the proof of the existence theorem. It states that h -convex potentials are locally semi-convex.

Proposition 3.1.10. Let V from Θ to $(-\infty, +\infty]$ be h -convex. If K is a convex compact subset of Θ , $\delta > 0$ and $R > 0$ satisfy

(i) $K + \delta \bar{B}(0, 1) \subset \Theta$

(ii) $|V(\theta)| \leq R, \quad \forall \theta \in K + \delta \bar{B}(0, 1)$, (where $\bar{B}(0, 1)$ is the closed unit ball of R^p) then V is locally semi-convex in K .

In particular, any locally bounded h -convex mapping V is locally semi-convex in Θ .

Proof. Step 1: Let F be bounded in R_+^m . Define,

$$h_\lambda(\theta, q) = h(\theta, q) + \lambda \|\theta\|^2, \quad \forall (\theta, q) \in K \times F$$

with $\lambda \geq (1/2)Lip(K \times F, \frac{\partial h}{\partial \theta})$ and

$$Lip(K \times F, \frac{\partial h}{\partial \theta}) := \sup_{\{(\theta, \theta', q) \in K^2 \times F: \theta \neq \theta'\}} \left\| \frac{\partial h}{\partial \theta}(\theta, q) - \frac{\partial h}{\partial \theta}(\theta', q) \right\| \|\theta - \theta'\|^{-1}.$$

Thus, we have

$$\begin{aligned} H(h_\lambda) &= \left\langle \frac{\partial h_\lambda}{\partial \theta}(\theta, q) - \frac{\partial h_\lambda}{\partial \theta}(\theta', q), \theta - \theta' \right\rangle \\ &\geq 2\lambda \|\theta - \theta'\|^2 - \sup \left\| \frac{\partial h}{\partial \theta}(\theta, q) - \frac{\partial h}{\partial \theta}(\theta', q) \right\| \|\theta - \theta'\|. \end{aligned}$$

Consequently,

$$\left\langle \frac{\partial h_\lambda}{\partial \theta}(\theta, q) - \frac{\partial h_\lambda}{\partial \theta}(\theta', q), \theta - \theta' \right\rangle \geq \left[2\lambda - Lip(K \times F, \frac{\partial h}{\partial \theta}) \right] \|\theta - \theta'\|^2.$$

Using the definition of λ we can conclude that

$$\left\langle \frac{\partial h_\lambda}{\partial \theta}(\theta, q) - \frac{\partial h_\lambda}{\partial \theta}(\theta', q), \theta - \theta' \right\rangle \geq 0.$$

Since the $H(h_\lambda)$ is non-negative, $h_\lambda(\cdot, q)$ is convex in K , for all $q \in F$.

Step 2: Let V be h-convex, satisfying (i) and (ii). Then

$$V(\theta) = \sup_{(q,t) \in A} \{h(\theta, q) - t\}, \quad \forall \theta \in \Theta.$$

Let us first assume that the set

$$\Pi_1(A) := \{q \in R_+^m : \exists t \in R_+ : (q, t) \in A\}$$

is bounded, $\Pi_1(A) \subset \bar{B}(0, M)$ for some $M > 0$. Then define

$$V_\lambda(\cdot) := V(\cdot) + \lambda \|\cdot\|^2,$$

and note that

$$V_\lambda(\theta) = \sup_{(q,t) \in A} \{h_\lambda(\theta, q) - t\}.$$

From step 1, we have that if

$$\lambda \geq (1/2) \text{Lip} \left(K \times (\bar{B}(0, M) \cap R_+^m), \frac{\partial h}{\partial \theta} \right)$$

then V_λ is convex in K as the upper envelope of convex functions. Note that λ can be chosen only depending on K and M .

Step 3: General case.

Since V is h-convex, then

$$V(\theta) = \sup_{(q,t) \in A} \{h(\theta, q) - t\}, \quad \forall \theta \in \Theta.$$

We know from Proposition 3.1.7 that V is h-subdifferentiable on K and that there exists some positive constant $M(R, K, \delta)$ such that

$$\forall \theta \in K, \forall q \in \partial^h V(\theta), \quad \|q\| \leq M(R, K, \delta).$$

Finally define, for all $\theta \in \Theta$

$$\tilde{V}(\theta) := \sup_{\theta' \in K, q \in \partial^h V(\theta')} \{V(\theta') + h(\theta, q) - h(\theta', q)\}.$$

Since V is h-differentiable, then

$$\tilde{V}(\theta) = V(\theta), \quad \forall \theta \in K.$$

It follows then from Step 2 that any λ with

$$\lambda \geq (1/2)Lip \left(K \times \bar{B}(0, M(R, K, \delta)) \cap R_+^m, \frac{\partial h}{\partial \theta} \right)$$

is a semi-convexity modulus of V in K . Such a λ can be chosen independently of V since $M(R, k, \delta)$ is independent of V . \square

The following proposition relates h-subdifferentiability to the classical notion of gradient.

Proposition 3.1.11. *Let V be a function from Θ to R . Assume $q \in \partial^h V(\theta)$, where $\theta \in \Theta$ and V is differentiable at θ . Then, $\nabla V(\theta) = \frac{\partial h}{\partial \theta}(\theta, q)$.*

Proof. Let $\epsilon > 0$ be such that $\theta + B(0, \epsilon) \subset \Theta$ and let k be such that $\|k\| < \epsilon$. We then have

$$\begin{aligned} V(\theta + k) &= V(\theta) + \langle \nabla V(\theta), k \rangle + o(k) \\ &\geq V(\theta) + h(\theta + k, q) - h(\theta, q) \\ &= V(\theta) + \left\langle \frac{\partial h}{\partial \theta}(\theta, q), k \right\rangle + o(k), \end{aligned}$$

where the first inequality comes from the fact that $q \in \partial^h V(\theta)$.

This yields,

$$\left\langle \nabla V(\theta) - \frac{\partial h}{\partial \theta}(\theta, q), k \right\rangle \geq o(k).$$

Changing k to $-k$, we obtain

$$\left\langle \nabla V(\theta) - \frac{\partial h}{\partial \theta}(\theta, q), k \right\rangle = o(k),$$

for all k such that $\|k\| < \epsilon$. Letting $\|k\| \rightarrow 0$ we conclude that

$$\nabla V(\theta) = \frac{\partial h}{\partial \theta}(\theta, q).$$

\square

Remark 3.1.12. *Combining Propositions 3.1.7, 3.1.10 and 3.1.11 and Rademacher's Theorem [7], we obtain that every locally bounded h-convex potential V is almost everywhere differentiable, everywhere h-subdifferentiable so that*

$$\nabla V(\theta) = \frac{\partial h}{\partial \theta}(\theta, q), \text{ for almost every } \theta \in \Theta, \text{ for every } q \in \partial^h V(\theta).$$

Now, we can rewrite the following principal's problem

$$(PA) \begin{cases} \inf \Pi(q, t) := \int_{\Theta} [C(q) - t(\theta)] f(\theta) d\theta \\ \text{s.t.} \\ (q, t) \text{ incentive-compatible} \\ h(\theta, q(\theta)) - t(\theta) \geq 0, \quad \forall \theta \in \Theta \end{cases}$$

as a variational problem with a h-convexity constraint.

Proposition 3.1.13. *The principal's problem (PA) is equivalent to*

$$(PA') \begin{cases} \inf J(q, V) := \int_{\Theta} [C(q) - h(\theta, q(\theta)) + V(\theta)] f(\theta) d\theta \\ \text{s.t.} \\ V \text{ is h-convex,} \\ q(\theta) \in \partial^h V(\theta), \quad \forall \theta \in \Theta \\ V(\theta) \geq 0, \quad \forall \theta \in \Theta \end{cases}$$

Proof. The proof follows directly from Proposition 3.1.5. □

3.2 Compactness

In this section, by $V : \Theta \rightarrow R$ h-convex, we mean that there exists a h-convex function $\tilde{V} : \Theta \rightarrow R$ such that $\tilde{V} = V$ almost everywhere in Θ .

Notation $\omega \subset\subset \Theta$ means that the closure of ω is included in Θ .

The proof of the next proposition can be found in Carlier [10, 11].

Proposition 3.2.1. *Let (u_n) be a sequence of convex functions in Θ such that for every open convex set ω with $\omega \subset\subset \Theta$, the following holds:*

$$\sup_n \|u_n\|_{W^{1,1}(\omega)} < +\infty.$$

Then, there exists a function \bar{u} that is convex in Θ , a measurable subset A of Θ and a subsequence again labeled (u_n) , such that:

1. (u_n) converges to \bar{u} uniformly on compact subsets of Θ ,
2. (∇u_n) converges to $\nabla \bar{u}$ pointwise on A and $\dim_H(\Theta \setminus A) \leq n-1$, where $\dim_H(\Theta \setminus A)$ is the Hausdorff dimension of $\Theta \setminus A$. In particular, (∇u_n) converges to $\nabla \bar{u}$ almost everywhere in Θ .

This proposition extends to h-convex functions as follows:

Proposition 3.2.2. *Let (V_n) be a sequence of h-convex functions in Θ such that the following holds:*

$$\sup_n \|V_n\|_{W^{1,1}(\Theta)} < +\infty.$$

Then, there exists a function $\bar{V} \in W^{1,1}(\Theta)$, that is h-convex in Θ , a measurable subset A of Θ and a subsequence again labeled (V_n) , such that:

1. (V_n) converges to \bar{V} uniformly on compact subsets of Θ ,
2. (∇V_n) converges to $\nabla \bar{V}$ pointwise in A and $\dim_H(\Theta \setminus A) \leq n-1$. In particular, (∇V_n) converges to $\nabla \bar{V}$ almost everywhere in Θ .

Proof. Step 1: Let us prove that for all $\omega \subset\subset \Theta$ we have

$$\sup_n \|V_n\|_{L^\infty(\bar{\omega})} < +\infty.$$

Otherwise, there exists a sequence (θ_n) of elements of ω such that

$$\limsup_n |V_n(\theta_n)| = +\infty.$$

Extracting subsequences, if necessary, we may also assume $\theta_n \rightarrow \bar{\theta} \in \bar{\omega}$ and without loss of generality $V_n(\theta_n) \rightarrow +\infty$. Let $\theta \in \Theta$ be such that $\theta_i > \bar{\theta}_i, i = 1, \dots, p$. Then, there exists n_0 such that $n \geq n_0 \Rightarrow \theta \geq \theta_n$ and, since every V_n is nondecreasing (by H1) we get, for all $n \geq n_0$

$$V_n(\theta) \geq V_n(\theta_n).$$

Hence

$$\begin{aligned} \int_{\{\alpha \in \Theta, \alpha \geq \theta\}} V_n(\alpha) d\alpha &\geq \int_{\{\alpha \in \Theta, \alpha \geq \theta\}} V_n(\theta_n) d\alpha \\ &= V_n(\theta_n) \int_{\{\alpha \in \Theta, \alpha \geq \theta\}} d\alpha. \end{aligned}$$

Consequently,

$$\int_{\{\alpha \in \Theta, \alpha \geq \theta\}} V_n(\alpha) d\alpha \geq \text{meas}\{\alpha \in \Theta, \alpha \geq \theta\} V_n(\theta_n) \rightarrow +\infty,$$

which is in contradiction to our original assumption.

Step 2: Let us define for every k

$$\omega_k := \{\theta \in \Theta : \text{dist}(\theta, \partial\Theta) > \frac{1}{2^{k+1}}\}$$

where $\partial\Theta$ represents the boundary of Θ .

From Step 1, we may define the nondecreasing sequence

$$M_k := \sup_n \|V_n\|_{L^\infty(\bar{\omega}_{k+1})} < +\infty, \quad \forall k.$$

Since V_n is h-convex and bounded in $\bar{\omega}_{k+1}$, using Proposition 3.1.10, there exists a sequence $\lambda_k > 0$ such that for all n , the function

$$V_{n,k}(\cdot) := V_n(\cdot) + \lambda_k \|\cdot\|^2$$

is convex in $\bar{\omega}_{k+1}$.

Using Proposition 3.2.1, there exists a subsequence of $(V_{n,1})$ denoted $(V_{p_1(n),1})$ a convex function $V + \lambda_1 \|\cdot\|^2$ and a measurable subset $A_1 \subset \Theta$ such that

$$\begin{cases} (V_{p_1(n),1}) \text{ converges uniformly to } \bar{V} + \lambda_1 \|\cdot\|^2 \text{ on } \bar{\omega}_1, \\ (\nabla V_{p_1(n),1}) \text{ converges to } \nabla(\bar{V} + \lambda_1 \|\cdot\|^2) \text{ pointwise on } A_1, \\ A_1 \subset \omega_1, \dim_H(\omega_1 \setminus A_1) \leq p-1, \end{cases}$$

which also implies

$$\begin{cases} (V_{p_1(n)}) \text{ converges uniformly to } \bar{V} \text{ on } \bar{\omega}_1, \\ (\nabla V_{p_1(n)}) \text{ converges to } \nabla \bar{V} \text{ pointwise on } A_1. \end{cases}$$

By induction for given $k \geq 1$, there exists a subsequence $\{V_{p_k(n)}\}$ of $\{V_{p_{k-1}(n)}\}$ and a measurable subset $A_k \subset \omega_k$ such that

$$\begin{cases} (V_{p_k(n)}) \text{ converges uniformly to } \bar{V} \text{ on } \bar{\omega}_k, \\ (\nabla V_{p_k(n)}) \text{ converges to } \nabla \bar{V} \text{ pointwise on } A_k, \\ A_{k-1} \subset A_k. \end{cases}$$

Defining $\Psi(n) := p_n(n)$ for all n , we have

$$\begin{cases} (V_{\Psi(n)}) \text{ converges uniformly to } \bar{V} \text{ on compact subsets of } \Theta, \\ (\nabla V_{\Psi(n)}) \text{ converges to } \nabla \bar{V} \text{ pointwise on } A, \end{cases}$$

with

$$A := \bigcap_n A_n,$$

so that $\dim_H(\Theta \setminus A) \leq p - 1$. Fatou's Lemma yields

$$\int_{\Theta} \bar{V}(\theta) + |\nabla \bar{V}(\theta)| d\theta \leq \liminf \int_{\Theta} V_{\Psi(n)}(\theta) + |\nabla V_{\Psi(n)}|(\theta).$$

Since $\sup \|V_n\|_{W^{1,1}(\Theta)} < +\infty$, this implies $\bar{V} \in W^{1,1}(\Theta)$. Note also that for all k , λ_k is a modulus of semi-convexity of \bar{V} in $\bar{\omega}_{k+1}$.

Step 3: \bar{V} is h-convex.

First, let us relabel $V_{\Psi(n)}$ to V_n . Define, for all $\theta \in \Theta$

$$\tilde{V}(\theta) := \sup_{\{\theta' \in \Theta, q \in F(\theta')\}} \{\bar{V}(\theta') + h(\theta, q) - h(\theta', q)\}$$

where

$$F(\theta') = \bigcap_{N \geq 1} \bigcup_{n \geq N} \overline{\partial^h V_n(\theta')}.$$

Note first that $F(\theta') \neq \emptyset$ by Proposition 3.1.7. \tilde{V} is then well-defined and h-convex. Next we show that $\tilde{V} = \bar{V}$. It is clear that $\tilde{V} \geq \bar{V}$ (choose $\theta' = \theta$). To show the converse inequality, let $(\theta, \theta') \in \Theta^2$ and $q \in F(\theta')$. There exists $n_k \rightarrow +\infty$ as $k \rightarrow +\infty$, $q_{n_k} \in \partial^h V_{n_k}(\theta')$ for all k with $q = \lim_k q_{n_k}$. Then for all k ,

$$V_{n_k}(\theta) \geq V_{n_k}(\theta') + h(\theta, q_{n_k}) - h(\theta', q_{n_k})$$

which, at the limit, yields $\bar{V}(\theta) \geq \bar{V}(\theta') + h(\theta, q) - h(\theta', q)$ for all $(\theta, \theta') \in \Theta^2$, thus showing $\bar{V} \geq \tilde{V}$. Therefore, $\bar{V} = \tilde{V}$ is h-convex. \square

3.3 Existence result for a linear cost function

For convenience, we will first show the existence result for linear cost functions. Then, we will extend the argument to more general cost functions. In this section, we assume that $C(q) = \langle p, q \rangle$, where $p \in R_+^m$ denotes the range of food prices. Then, the principal-agent problem becomes:

$$(PA) \left\{ \begin{array}{l} \inf \Pi(q, t) := \int_{\Theta} [\langle p, q \rangle - t(\theta)] f(\theta) d\theta \\ s.t. \\ (q, t) \text{ is incentive-compatible} \\ h(\theta, q(\theta)) - t(\theta) \geq 0, \forall \theta \in \Theta, \end{array} \right.$$

or equivalently

$$(PA') \left\{ \begin{array}{l} \inf J(q, V) := \int_{\Theta} [\langle p, q \rangle - h(\theta, q(\theta)) + V(\theta)] f(\theta) d\theta \\ s.t. \\ V \text{ is h-convex,} \\ q(\theta) \in \partial^h V(\theta), \forall \theta \in \Theta \\ V(\theta) \geq 0, \forall \theta \in \Theta. \end{array} \right.$$

In addition to the previously mentioned hypotheses, we assume the following technical hypotheses:

H4. There exist $\alpha \leq 1$, $a > 0$ and $b \in R$, such that for all $(\theta, q) \in \Theta \times R_+^m$, we have

$$h(\theta, q) \leq a \|q\|^\alpha - b.$$

and if $\alpha = 1$, $a < \min_{1 \leq i \leq m} p_i$.

H5. There exist $\beta \in (0, \alpha)$, $c > 0$, $d \in R$ such that for all $(\theta, q) \in \Theta \times R_+^m$

$$\left\| \frac{\partial h}{\partial \theta}(\theta, q) \right\| \leq c \|q\|^\beta + d.$$

Under the above assumptions, the following existence result holds.

Theorem 3.3.1. *(PA') admits at least one solution.*

Proof. Consider an arbitrary V that is both h-convex and locally bounded. Then, Proposition 3.1.7 and the measurable selection Theorem (see chapter 2, Theorem

2.1.6) imply that both set-valued maps $\partial^h V(\cdot)$ and

$$\Phi_V : \theta \rightarrow \underset{\partial^h V(\theta)}{\operatorname{argmin}}\{-h(\theta, \cdot) + \langle p, \cdot \rangle\}$$

are non-empty and compact-valued and admit measurable selections.

Let (V_n, q_n) be a minimizing sequence of (PA') . Without loss of generality, we may assume that for all n , q_n is measurable and $q_n(\theta) \in \Phi_{V_n}(\theta) \forall \theta \in \Theta$. Then,

$$\left| \int_{\Theta} (V_n(\theta) - h(\theta, q_n(\theta)) + \langle p, q_n(\theta) \rangle) f(\theta) d\theta \right| \leq C, \quad (3.3)$$

where C is a positive constant.

Since $V_n(\theta)$ and $f(\theta)$ are both positive, we have

$$\int_{\Theta} (-h(\theta, q_n(\theta)) + \langle p, q_n(\theta) \rangle) f(\theta) d\theta \leq C.$$

By adding $\langle p, q_n(\theta) \rangle$ to both sides of the inequality in Hypothesis (H4), we have

$$\int_{\Theta} (-a \|q_n(\theta)\|^\alpha + \langle p, q_n(\theta) \rangle) f(\theta) d\theta \leq \int_{\Theta} (-h(\theta, q_n(\theta)) - b + \langle p, q_n(\theta) \rangle) f(\theta) d\theta \leq C',$$

where $C' = C - b$, $C' > 0$ by choosing C large enough.

Equivalently,

$$-a \int_{\Theta} \|q_n(\theta)\|^\alpha f(\theta) d\theta + \int_{\Theta} \langle p, q_n(\theta) \rangle f(\theta) d\theta \leq C'. \quad (3.4)$$

Let $M = \min_{1 \leq i \leq m} p_i > 0$, then

$$\begin{aligned} \int_{\Theta} \langle p, q_n(\theta) \rangle f(\theta) d\theta &= \int_{\Theta} \sum_{i=1}^m p_i q_n^i f(\theta) d\theta \\ &\geq M \sum_{i=1}^m \int_{\Theta} q_n^i f(\theta) d\theta \\ &= M \int_{\Theta} \|q_n(\theta)\| f(\theta) d\theta. \end{aligned} \quad (3.5)$$

Inserting (3.5) into (3.4) yields

$$-a \int_{\Theta} \|q_n(\theta)\|^\alpha f(\theta) d\theta + M \int_{\Theta} \|q_n(\theta)\| f(\theta) d\theta \leq C' \quad (3.6)$$

At this step we consider two different cases: where $\alpha < 1$ and where $\alpha = 1$. First assume $\alpha < 1$, Then by Young's inequality we have

$$\begin{aligned}\|q_n(\theta)\|^\alpha &= \delta \|q_n(\theta)\|^\alpha \frac{1}{\delta} \\ &\leq \frac{(\delta \|q_n(\theta)\|^\alpha)^\eta}{\eta} + \frac{|\frac{1}{\delta}|^{\eta'}}{\eta'}\end{aligned}$$

where η and η' are positive real numbers such that $\frac{1}{\eta} + \frac{1}{\eta'} = 1$. Now, by choose $\eta = \frac{1}{\alpha}$ we have

$$\|q_n(\theta)\|^\alpha \leq \alpha \delta^{\frac{1}{\alpha}} \|q_n(\theta)\| + \frac{1 - \alpha}{\delta^{\frac{1}{1-\alpha}}}. \quad (3.7)$$

Inserting (3.7) into (3.6) yields

$$(M - a\alpha\delta^{\frac{1}{\alpha}}) \int_{\Theta} \|q_n(\theta)\| f(\theta) d\theta - \frac{a(1-\alpha)}{\delta^{\frac{1}{1-\alpha}}} \int_{\Theta} f(\theta) d\theta \leq C'.$$

Consequently,

$$(M - a\alpha\delta^{\frac{1}{\alpha}}) \int_{\Theta} \|q_n(\theta)\| f(\theta) \leq C'',$$

where $C'' = C' + \frac{a(1-\alpha)}{\delta^{\frac{1}{1-\alpha}}}$.

Choosing δ small enough so that $M - a\alpha\delta^{\frac{1}{\alpha}} > 0$ and the fact that $f \in L^\infty(\Theta)$ has a positive essential infimum ensure that (q_n) is bounded in $L^1(\Theta, R_+^m)$. Now we consider the case that $\alpha = 1$. Rewriting equation (3.6) for $\alpha = 1$ gives us

$$(M - a) \int_{\Theta} \|q_n(\theta)\| f(\theta) d\theta \leq C'.$$

We know from (H4) that $M - a > 0$. Dividing both side of the above equation by $(M - a)$ and the fact that $f \in L^\infty(\Theta)$ has a positive essential infimum implies that (q_n) is bounded is $L^1(\Theta, R_+^m)$. Moreover, by (H4) we know that

$$h(\theta, q_n(\theta)) \leq a \|q_n(\theta)\|^\alpha - b. \quad (3.8)$$

At the same time using Young's inequality we have

$$\|q_n(\theta)\|^\alpha \leq \frac{(\|q_n(\theta)\|^\alpha)^\eta}{\eta} + \frac{1}{\eta'}$$

where η and η' are positive real numbers such that $\frac{1}{\eta} + \frac{1}{\eta'} = 1$. Inserting the above

equation into (3.8) implies

$$h(\theta, q_n(\theta)) \leq a \left(\frac{\|q_n(\theta)\|^{\alpha\eta}}{\eta} + \frac{1}{\eta'} \right) - b.$$

Choosing $\eta = \frac{1}{\alpha}$ and using the fact that (q_n) is bounded in L^1 , then $h(\theta, q_n)$ is bounded in L^1 . Using this fact and that facts that (q_n) is bounded in $L^1(\Theta)$ and $f \in L^\infty(\Theta)$ has a positive essential infimum, equation (3.3) ensure that (V_n) is also bounded in $L^1(\Theta, R_+)$. For all n , V_n is locally bounded. From Propositions 3.1.10 and 3.1.11, we deduce that for all n and for almost every $\theta \in \Theta$,

$$\nabla V_n(\theta) = \frac{\partial h}{\partial \theta}(\theta, q_n(\theta)).$$

Using Hypothesis (H4) we get,

$$\begin{aligned} \|\nabla V_n\| &= \left\| \frac{\partial h}{\partial \theta}(\theta, q_n(\theta)) \right\| \\ &\leq c\|q\|^\beta + d, \quad \beta \in (0, \alpha) \\ &\leq c(1 + \|q_n(\theta)\|) + d \quad \text{a.e. in } \Theta. \end{aligned}$$

Thus, ∇V_n is bounded in $L^1(\Theta)$ and V_n satisfies the assumptions of Proposition 3.2.1. Consequently, we may now assume that

$$\begin{cases} (V_n) \text{ converges in } L^1(\Theta) \text{ and uniformly on compact subsets of } \Theta, \\ (\nabla V_n) \text{ converges a.e. to } \nabla \bar{V}, \end{cases}$$

where $\bar{V} \in W^{1,1}(\Theta, R_+)$ is h-convex.

Finally, define $\bar{q}(\cdot)$ as a measurable selection of $\Phi_{\bar{V}}(\cdot)$.

First, since (V_n) converges to \bar{V} in $L^1(\Theta)$ and $f \in L^\infty(\Theta)$ we have

$$\lim_n \int_{\Theta} V_n(\theta) f(\theta) d\theta = \int_{\Theta} \bar{V}(\theta) f(\theta) d\theta. \quad (3.9)$$

Fatou's Lemma yields,

$$\liminf_n \int_{\Theta} [-h(\theta, q_n(\theta)) + C(q_n)] f(\theta) d\theta \geq \int_{\Theta} \liminf_n [-h(\theta, q_n(\theta)) + C(q_n(\theta))] f(\theta) d\theta. \quad (3.10)$$

Let us define for all fixed θ ,

$$\alpha(\theta) := \liminf_n \{-h(\theta, q_n(\theta)) + C(q_n(\theta))\}.$$

Since $(q_n(\theta))$ is bounded in L^1 by Hypothesis (H4), up to a subsequence, we may assume that

$$\begin{cases} \alpha(\theta) = \liminf_n -h(\theta, q_n(\theta)) + C(q_n(\theta)), \\ q_n(\theta) \rightarrow y(\theta) \quad a.e. \end{cases}$$

We know that for all $\theta' \in \Theta$ and all n ,

$$V_n(\theta') \geq V_n(\theta) + h(\theta', q_n(\theta)) - h(\theta, q_n(\theta)).$$

In the limit, we obtain

$$\bar{V}(\theta') \geq \bar{V}(\theta) + h(\theta', y(\theta)) - h(\theta, y(\theta)).$$

The above equation means that $y(\theta) \in \partial^h \bar{V}(\theta)$.

Then, we get

$$\alpha(\theta) = -h(\theta, y(\theta)) + C(y(\theta)) \geq -h(\theta, \bar{q}(\theta)) + C(\bar{q}(\theta)). \quad (3.11)$$

Therefore, we have

$$\begin{aligned} \inf (PA') &= \liminf_n \int_{\Theta} [V_n(\theta) - h(\theta, q_n(\theta)) + C(q_n(\theta))] f(\theta) d\theta \\ &\geq \int_{\Theta} [\bar{V}(\theta) - h(\theta, \bar{q}(\theta)) + C(\bar{q}(\theta))] f(\theta) d\theta \\ &= J(\bar{V}, \bar{q}) \end{aligned}$$

where in the above inequality we used equations (3.9), (3.10) and (3.11).

This shows that (\bar{V}, \bar{q}) is a solution of (PA') . □

3.4 Existence of solutions for the general cost functions

In this section, we extend Carlier's proof [9] to our model for the general cost functions. Suppose h satisfies (H1)-(H3) and (H5). We recall the minimization problem (PA') :

$$(PA') \left\{ \begin{array}{l} \min_{q(\theta), V(\theta)} \int_{\Theta} \phi(\theta, V(\theta), q(\theta)) d\theta \\ \text{s.t.} \\ V \text{ is h-convex,} \\ q(\theta) \in \partial^h V(\theta), \\ V(\theta) \geq 0 \end{array} \right.$$

where $\phi(\theta, V(\theta), q(\theta)) = [V(\theta) - h(\theta, q(\theta)) + C(q(\theta))]f(\theta)$. To prove the extension of the previous result regarding existence of solutions, we generalize (H4) to a larger class of cost functions and we further add (H6):

H4'. There exist $\alpha \leq 1$, $a > 0$ and $b \in \mathbb{R}$, such that for all $(\theta, q) \in \Theta \times \mathbb{R}_+^m$

$$h(\theta, q) \leq a\|q\|^\alpha - b.$$

H6. $\phi(\cdot, \cdot, \cdot)$ is a normal integrand, which means that for almost every $\theta \in \Theta$, $\phi(\theta, \cdot, \cdot)$ is lower semi-continuous and that there exists a borelian map $\bar{\phi}$ such that $\phi(\theta, \cdot, \cdot) = \bar{\phi}(\theta, \cdot, \cdot)$ for almost every $\theta \in \Theta$. There exist $A > 0$ and $\Psi \in L^1(\Theta)$ such that for almost every $\theta \in \Theta$ and every $(V, q) \in \mathbb{R} \times \mathbb{R}_+^m$

$$\begin{aligned} \phi(\theta, V, q) &= V - h(\theta, q) + C(q) \\ &\geq A(|V| + \|q\|^\gamma) + \Psi(\theta), \quad \gamma \geq 1. \end{aligned}$$

Theorem 3.4.1. *Problem (PA') admits at least one solution.*

Proof. The proof is similar to that of Theorem 3.3.1. Indeed, according to (H6), we know that (V_n) is bounded in $W^{1,1}(\Theta)$. Using Proposition 3.2.2 we have:

$$\left\{ \begin{array}{l} V_n \text{ converges to } \bar{V} \text{ in } L^1(\Theta) \text{ and uniformly on compact subsets of } \Theta, \\ \nabla V_n \text{ converges almost everywhere in } \Theta \text{ to } \nabla \bar{V}, \end{array} \right.$$

where \bar{V} is h-convex and belongs to $W^{1,1}(\Theta)$.

Following the proof of Theorem 3.3.1, we can choose q_n to be measurable and such

and slightly modify (H6):

H6'. $\phi(\cdot, \cdot)$ is a normal integrand and there exist $A > 0$ and $\Psi \in L^1(\Theta)$ such that for almost every $\theta \in \Theta$

$$\begin{aligned}\phi(\theta, q) &= -h(\theta, q) + C(q) \\ &\geq A\|q\|^\gamma + \Psi(\theta), \quad \gamma \geq 1.\end{aligned}$$

Under assumptions we imposed above, the following existence result holds.

Theorem 3.5.1. *The discrete problem (PD) has at least one solution.*

Proof. From the (IR) constraint, we have

$$t_{i_1, \dots, i_p} \leq h(\theta_{i_1, \dots, i_p}, q_{i_1, \dots, i_p}). \quad (3.12)$$

and then

$$t_{i_1, \dots, i_p} - C(q_{i_1, \dots, i_p}) \leq h(\theta_{i_1, \dots, i_p}, q_{i_1, \dots, i_p}) - C(q_{i_1, \dots, i_p}). \quad (3.13)$$

Recall by (H6') that, we have for fixed value of $\theta \in \Theta$, if $q \rightarrow \infty$, then $-h(\theta, q) + C(q) \rightarrow +\infty$. As a result, $h(\theta, q) - C(q) \rightarrow -\infty$. As a result, by (3.13), $t_{i_1, \dots, i_p} - C(q_{i_1, \dots, i_p})$ will go to negative infinity which does not affect our maximization problem. This implies q_{i_1, \dots, i_p} is bounded. So is $h(\theta_{i_1, \dots, i_p}, q_{i_1, \dots, i_p})$ using (H4'). This means that t_{i_1, \dots, i_p} is bounded by (3.12). Therefore, we are searching for a maximum of the upper semi-continuous function on a compact set which definitely exist (Weiestrass-Lebesgue Lemma). \square

Chapter 4

Solutions under the Spence Mirrlees Condition

In the previous chapter we found sufficient conditions under which both discrete and continuous problems have at least one solution. Although the conditions guarantee the existence of a solution, there might be some other cases that violate those conditions but a solution still exists. Hence, in the rest of the thesis we will deal with the original problem without imposing any conditions given in Chapter 3. The only condition that we impose here is that the utility of the agent, $h(\theta, q)$, satisfies the Spence Mirrlees Condition (SMC) to be defined below. The purpose of this chapter is to introduce conditions under which we can characterize the solution of the problem under SMC.

The SMC was first introduced by Spence [29] in 1973. Mirrlees [21] also used this condition to study a taxation problem. Principal-agent problem with adverse selection under this condition was also studied in a few books [6, 17, 27]. For a quasi-linear utility function $U_a(\theta, q, t) = h(\theta, q) - t$, that we consider in our model, the SMC is equivalent to the following condition:

Definition 4.0.2. (*Spence Mirrlees condition*)

A function $h(\theta, q)$ satisfies the Spence Mirrlees condition (SMC) if

$$\forall \theta, q \quad \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q) > 0.$$

Remark 4.0.3. In fact, the Spence Mirrlees condition states that $\frac{\partial^2 h}{\partial \theta \partial q}(\theta, q)$ does not change signs for all (θ, q) . For simplicity we assume that the sign is positive. The

result under SMC when the sign is negative is similar.

The Spence Mirrlees condition is important in the study of the principal-agent problem, since if the customer's utility function satisfies the Spence Mirrlees condition then we could reduce the number of incentive compatible constraints.

In the following, we will first examine the problem of continuum of types and then study the discrete problem, using the reduction of incentive compatible constraints. I borrow the idea from the case where the customer's satisfaction is linear in θ as discussed by Patrick Bolton and Mathias Dewatripont [6].

4.1 One dimensional continuum of type problem

In this section, we consider a special case of (PA) where Θ is a closed interval $[\theta_1, \theta_2]$. The problem can be formulated as:

$$\begin{aligned}
 (PA)_c \quad & \max_{q(\theta), t(\theta)} \int_{\theta_1}^{\theta_2} (t(\theta) - C(q(\theta))) f(\theta) d\theta \\
 \text{s.t.} \quad & h(\theta, q(\theta)) - t(\theta) \geq 0, \quad \forall \theta \in [\theta_1, \theta_2] \quad (\text{IR}) \\
 & h(\theta, q(\theta)) - t(\theta) \geq h(\theta, q(\theta')) - t(\theta'), \quad \forall \theta, \theta' \in [\theta_1, \theta_2] \quad (\text{IC})
 \end{aligned}$$

where $f(\theta)$ denotes the density function of having a customer of taste θ with the cumulative distribution function $F(\theta) = \int_{\theta_1}^{\theta} f(\theta') d\theta'$.

In this section, we assume that the following hypotheses hold.

H1. Assume that $h(\theta, q)$ satisfies the Spence Mirrlees condition and $h_q(\cdot) > 0$, $h_\theta(\cdot) > 0$.

H2. Suppose $C(\cdot)$ is a twice differentiable, increasing and strictly convex function such that $C'(0) = 0$.

Remark 4.1.1. Under the condition that $h(\theta, q)$ is increasing in θ , the (IR) constraints can be replaced by $h(\theta_1, q(\theta_1)) - t(\theta_1) \geq 0$. Indeed, for all $\theta \in (\theta_1, \theta_2]$,

$$\begin{aligned}
 h(\theta, q(\theta)) - t(\theta) & \geq h(\theta, q(\theta_1)) - t(\theta_1) && \text{by (IC),} \\
 & > h(\theta_1, q(\theta_1)) - t(\theta_1) && \text{since } h(\theta, q) \text{ is increasing in } \theta, \\
 & \geq 0 && \text{by (IR).}
 \end{aligned}$$

The key for solving the problem $(PA)_c$ is to reduce the constraints to the simplest possible form.

Theorem 4.1.2. *Let $(q(\theta), t(\theta))$ be an optimal solution of $(PA)_c$. Then the constraint (IR) is active at $\theta = \theta_1$, that is, $h(\theta_1, q(\theta_1)) - t(\theta_1) = 0$.*

Proof. We prove by contradiction. Suppose that $h(\theta_1, q(\theta_1)) - t(\theta_1) > 0$. Then,

$$\begin{aligned} h(\theta, q(\theta)) - t(\theta) &\geq h(\theta, q(\theta_1)) - t(\theta_1) && \text{by (IC).} \\ &> h(\theta_1, q(\theta_1)) - t(\theta_1) && \text{since } h(\theta, q) \text{ is increasing in } \theta. \\ &> 0. \end{aligned}$$

Consequently a sufficiently small ϵ can be found such that,

$$h(\theta, q(\theta)) - (t(\theta) + \epsilon) > 0,$$

$$h(\theta, q(\theta)) - (t(\theta) + \epsilon) \geq h(\theta, q(\theta')) - (t(\theta') + \epsilon)$$

which imply that $(q(\theta), t(\theta) + \epsilon)$ is still feasible. However the objective function value in $(PA)_c$ at $(q(\theta), t(\theta) + \epsilon)$ is larger than the one at $(q(\theta), t(\theta))$. This violates the fact that $(q(\theta), t(\theta))$ is an optimal solution. \square

The fact that the utility of the customer satisfies the SMC enables us to reduce the set of (IC) constraints.

Theorem 4.1.3. *Under SMC, a pair of differentiable functions $(q(\theta), t(\theta))$ satisfies the (IC) constraint if and only if it satisfies monotonicity of q in θ , and the first order condition, that is,*

$$\frac{dq}{d\theta}(\theta) \geq 0, \quad \forall \theta \in (\theta_1, \theta_2), \quad (M)$$

$$\frac{\partial h}{\partial q}(\theta, q)q'(\theta) - t'(\theta) = 0, \quad \forall \theta \in (\theta_1, \theta_2) \quad (FOC).$$

Proof. Suppose that $q(\theta)$ and $t(\theta)$ satisfy the (IC) constraint. Then it holds that

$$h(\theta, q(\theta)) - t(\theta) = \max_{\theta' \in [\theta_1, \theta_2]} \{h(\theta, q(\theta')) - t(\theta')\}.$$

Since $\theta' = \theta$ is the maximizer of the above equation, if $\theta \in (\theta_1, \theta_2)$, the first and

second order conditions

$$\frac{\partial h}{\partial q}(\theta, q(\theta))q'(\theta) - t'(\theta) = 0, \quad \forall \theta \in (\theta_1, \theta_2) \quad (\text{FOC})$$

$$\frac{\partial^2 h}{\partial q^2}(\theta, q(\theta))(q'(\theta))^2 + \frac{\partial h}{\partial q}(\theta, q(\theta))q''(\theta) - t''(\theta) \leq 0, \quad \forall \theta \in (\theta_1, \theta_2) \quad (\text{SOC})$$

hold. Taking the derivatives of the first order condition (FOC) with respect to θ gives

$$\frac{\partial^2 h}{\partial \theta \partial q}(\theta, q(\theta))q'(\theta) + \frac{\partial^2 h}{\partial q^2}(\theta, q(\theta))(q'(\theta))^2 + \frac{\partial h}{\partial q}(\theta, q(\theta))q''(\theta) - t''(\theta) = 0.$$

Comparing this equation with the second order condition (SOC) it yields

$$\frac{\partial^2 h}{\partial \theta \partial q}(\theta, q(\theta))q'(\theta) \geq 0,$$

which implies by the SMC ($\frac{\partial^2 h}{\partial \theta \partial q}(\cdot, \cdot) > 0$) that

$$q'(\theta) \geq 0.$$

Hence, we have shown that if $(q(\theta), t(\theta))$ satisfies the (IC) constraint then it should also satisfy the monotonicity and the first order condition.

We prove the reverse by contradiction. Suppose that $(q(\theta), t(\theta))$ satisfies (M), (FOC) but not (IC). That is, there are $\theta' \neq \theta \in [\theta_1, \theta_2]$ such that

$$h(\theta, q(\theta)) - t(\theta) < h(\theta, q(\theta')) - t(\theta').$$

Equivalently

$$\int_{\theta}^{\theta'} \left[\frac{\partial h}{\partial q}(\theta, q(x))q'(x) - t'(x) \right] dx > 0. \quad (4.1)$$

Without loss of generality, assume that $\theta < \theta'$. Then, since $\frac{\partial h}{\partial q}(\theta, q(x)) < \frac{\partial h}{\partial q}(x, q(x))$ and $q'(x) \geq 0$, we have

$$\int_{\theta}^{\theta'} \left(\frac{\partial h}{\partial q}(\theta, q(x))q'(x) - t'(x) \right) dx \leq \int_{\theta}^{\theta'} \left(\frac{\partial h}{\partial q}(x, q(x))q'(x) - t'(x) \right) dx = 0,$$

where the last equality is the result of the first order condition (FOC). But it contradicts (4.1). This completes the proof. \square

Remark 4.1.4. Under Spence Mirrlees condition the set of (IC) constraints is equivalent to the set of monotonicity conditions of q and the set of first order conditions. Since the first order conditions ensure that θ is a local maximizer of $\theta' \mapsto h(\theta, q(\theta')) - t(\theta')$, we call these conditions the local (IC) constraints.

Using Theorem 4.1.3 and Remark 4.1.1 under SMC, we can reformulate the problem $(PA)_c$ as follows:

$$\begin{aligned} \max_{q(\theta), t(\theta)} \quad & \int_{\theta_1}^{\theta_2} (t(\theta) - C(q(\theta)))f(\theta)d\theta, \\ \text{s.t.} \quad & h(\theta_1, q(\theta_1)) - t(\theta_1) \geq 0, \\ & q'(\theta) \geq 0, \quad \forall \theta \in (\theta_1, \theta_2), \quad (\text{M}) \\ & \frac{\partial h}{\partial q}(\theta, q(\theta))q'(\theta) - t'(\theta) = 0, \quad \forall \theta \in (\theta_1, \theta_2) \quad (\text{FOC}). \end{aligned}$$

We know by Theorem 4.1.2 that the constraint $h(\theta_1, q(\theta_1)) - t(\theta_1) \geq 0$ is active at the optimal solution. Hence under SMC an optimal solution $(q^*(\theta), t^*(\theta))$ of $(PA)_c$ must be an optimal solution of the following problem

$$\begin{aligned} (PA)_c' \quad \max_{q(\theta), t(\theta)} \quad & \int_{\theta_1}^{\theta_2} (t(\theta) - C(q(\theta)))f(\theta)d\theta, \\ \text{s.t.} \quad & h(\theta_1, q(\theta_1)) - t(\theta_1) = 0, \quad (\text{IR}_{\theta_1}) \\ & q'(\theta) \geq 0, \quad \forall \theta \in (\theta_1, \theta_2), \quad (\text{M}) \\ & \frac{\partial h}{\partial q}(\theta, q(\theta))q'(\theta) - t'(\theta) = 0, \quad \forall \theta \in (\theta_1, \theta_2) \quad (\text{FOC}). \end{aligned}$$

Consider the following relaxed problem of $(PA)_c'$ where the monotonicity constraint, $q'(\theta) \geq 0$ is omitted.

$$\begin{aligned} (PA)_c'' \quad \max_{q(\theta), t(\theta)} \quad & \int_{\theta_1}^{\theta_2} (t(\theta) - C(q(\theta)))f(\theta)d\theta, \\ \text{s.t.} \quad & h(\theta_1, q(\theta_1)) - t(\theta_1) = 0, \quad (\text{IR}_{\theta_1}) \\ & \frac{\partial h}{\partial q}(\theta, q(\theta))q'(\theta) - t'(\theta) = 0, \quad \forall \theta \in (\theta_1, \theta_2) \quad (\text{FOC}). \end{aligned}$$

Theorem 4.1.5. Let $(q^*(\theta), t^*(\theta))$ be a feasible solution of $(PA)_c''$. $(q^*(\theta), t^*(\theta))$ is an optimal solution of $(PA)_c''$ if and only if for all $\theta \in (\theta_1, \theta_2)$, the following equation

holds:

$$C'(q^*(\theta)) = \frac{\partial h}{\partial q}(\theta, q^*(\theta)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q^*(\theta)) \frac{1 - F(\theta)}{f(\theta)} \quad (C).$$

Proof. Step 1. Let $(q(\theta), t(\theta))$ be a feasible solution of $(PA)_c''$. We show that $t(\theta)$ can be replaced by a function of $q(\theta)$. Recall, for all $\theta \in [\theta_1, \theta_2]$ the utility of the customer was defined as

$$U_a(\theta) = h(\theta, q(\theta)) - t(\theta). \quad (4.2)$$

and hence,

$$t(\theta) = h(\theta, q(\theta)) - U_a(\theta).$$

By (FOC) we have

$$\frac{dU_a}{d\theta}(\theta) = \frac{\partial h}{\partial \theta}(\theta, q(\theta)), \quad \forall \theta \in (\theta_1, \theta_2).$$

Integrating the above equation gives

$$\int_{\theta_1}^{\theta} \frac{dU_a}{dx}(x) dx = \int_{\theta_1}^{\theta} \frac{\partial h}{\partial x}(x, q(x)) dx, \quad \forall \theta \in (\theta_1, \theta_2).$$

Consequently, we have

$$U_a(\theta) - U_a(\theta_1) = \int_{\theta_1}^{\theta} \frac{\partial h}{\partial x}(x, q(x)) dx.$$

Since $U_a(\theta_1) = h(\theta_1, q(\theta_1)) - t(\theta_1) = 0$, we have

$$U_a(\theta) = \int_{\theta_1}^{\theta} \frac{\partial h}{\partial x}(x, q(x)) dx.$$

Hence, we can rewrite the objective function of $(PA)_c''$ as:

$$\begin{aligned} \int_{\theta_1}^{\theta_2} (t(\theta) - C(q(\theta))) f(\theta) d\theta &= \int_{\theta_1}^{\theta_2} \left(h(\theta, q(\theta)) - \int_{\theta_1}^{\theta} \frac{\partial h}{\partial x}(x, q(x)) dx - C(q(\theta)) \right) f(\theta) d\theta \\ &= \int_{\theta_1}^{\theta_2} [h(\theta, q(\theta)) - C(q(\theta))] f(\theta) d\theta - \int_{\theta_1}^{\theta_2} \int_{\theta_1}^{\theta} \left(\frac{\partial h}{\partial x}(x, q(x)) \right) f(\theta) d\theta. \end{aligned} \quad (4.3)$$

Using integration by parts, the last terms of the above equation is equal to

$$\begin{aligned}
\int_{\theta_1}^{\theta_2} \int_{\theta_1}^{\theta} \left(\frac{\partial h}{\partial x}(x, q(x)) dx \right) f(\theta) d\theta &= \left[F(\theta) \int_{\theta_1}^{\theta} \frac{\partial h}{\partial x}(x, q(x)) dx \right]_{\theta_1}^{\theta_2} - \int_{\theta_1}^{\theta_2} \frac{\partial h}{\partial \theta}(\theta, q(\theta)) F(\theta) d\theta \\
&= [F(\theta)W(\theta)]_{\theta_1}^{\theta_2} - \int_{\theta_1}^{\theta_2} \frac{\partial h}{\partial \theta}(\theta, q(\theta)) F(\theta) d\theta \\
&= W(\theta_2)F(\theta_2) - \int_{\theta_1}^{\theta_2} \frac{\partial h}{\partial \theta}(\theta, q(\theta)) F(\theta) d\theta \\
&= W(\theta_2) - \int_{\theta_1}^{\theta_2} \frac{\partial h}{\partial \theta}(\theta, q(\theta)) F(\theta) d\theta \\
&= \int_{\theta_1}^{\theta_2} \frac{\partial h}{\partial \theta}(\theta, q(\theta)) d\theta - \int_{\theta_1}^{\theta_2} \frac{\partial h}{\partial \theta}(\theta, q(\theta)) F(\theta) d\theta \\
&= \int_{\theta_1}^{\theta_2} \frac{\partial h}{\partial \theta}(\theta, q(\theta)) (1 - F(\theta)) d\theta, \tag{4.4}
\end{aligned}$$

where $W(\theta) = \int_{\theta_1}^{\theta} \frac{\partial h}{\partial x}(x, q(x)) dx$ and $F(\theta) = \int_{\theta_1}^{\theta} f(\theta') d\theta'$ is a cumulative distribution function of f .

Inserting (4.4) into (4.3) gives

$$\int_{\theta_1}^{\theta_2} [t(\theta) - C(q(\theta))] f(\theta) d\theta = \int_{\theta_1}^{\theta_2} \left[(h(\theta, q(\theta)) - C(q)) f(\theta) - \frac{\partial h}{\partial \theta}(\theta, q(\theta)) (1 - F(\theta)) \right] d\theta.$$

Step 2. Since the objective function is continuous with respect to q , by Theorem 3.1 in [23] $q^*(\theta)$ is an optimal solution of

$$\max_{q(\theta)} \int_{\theta_1}^{\theta_2} \left[(h(\theta, q(\theta)) - C(q)) f(\theta) - \frac{\partial h}{\partial \theta}(\theta, q(\theta)) (1 - F(\theta)) \right] d\theta$$

if and only if it is a pointwise maximizer of the integrand for almost all θ . Since the integrand is concave in q by (H1) and (H2), then q^* is an optimal solution if and only if:

$$f(\theta) \left(\frac{\partial h}{\partial q}(\theta, q^*(\theta)) - C'(q^*(\theta)) \right) = \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q^*(\theta)) (1 - F(\theta)).$$

Equivalently q^* is an optimal solution if and only if,

$$C'(q^*(\theta)) = \frac{\frac{\partial h}{\partial q}(\theta, q^*(\theta)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q^*(\theta)) (1 - F(\theta))}{f(\theta)}. \tag{4.5}$$

□

This shows that $C'(q^*(\theta)) = \frac{\partial h}{\partial q}(\theta, q^*(\theta))$ for $\theta = \theta_2$ and $C'(q^*(\theta)) < \frac{\partial h}{\partial q}(\theta, q^*(\theta))$ for all the customers with taste $\theta < \theta_2$.

Theorem 4.1.6. *If $(q^*(\theta), t^*(\theta))$ is an optimal solution of $(PA)_c''$ and $\frac{dq^*}{d\theta}(\theta) \geq 0$ for all $\theta \in (\theta_1, \theta_2)$, then $(q^*(\theta), t^*(\theta))$ is an optimal solution of*

$$\begin{aligned}
(PA)_c' \quad & \max_{q(\theta), t(\theta)} \int_{\theta_1}^{\theta_2} (t(\theta) - C(q(\theta)))f(\theta)d\theta, \\
\text{s.t.} \quad & h(\theta_1, q(\theta_1)) - t(\theta_1) = 0, \quad (IR_{\theta_1}) \\
& \frac{dq}{d\theta}(\theta) \geq 0, \quad \forall \theta \in (\theta_1, \theta_2), \quad (M) \\
& \frac{\partial h}{\partial q}(\theta, q(\theta))q'(\theta) - t'(\theta) = 0, \quad \forall \theta \in (\theta_1, \theta_2) \quad (FOC).
\end{aligned}$$

Proof. Let $(q(\theta), t(\theta))$ be a feasible solution of $(PA)_c'$. Then it is also a feasible solution of $(PA)_c''$. Since $(q^*(\theta), t^*(\theta))$ is an optimal solution of $(PA)_c''$, we have

$$\int_{\theta_1}^{\theta_2} (t(\theta) - C(q(\theta)))f(\theta)d\theta \leq \int_{\theta_1}^{\theta_2} (t^*(\theta) - C(q^*(\theta)))f(\theta)d\theta.$$

Since $q^*(\theta)$ satisfies the monotonicity condition $\frac{dq^*}{d\theta}(\theta) \geq 0$ for all θ in (θ_1, θ_2) , then $(q^*(\theta), t^*(\theta))$ is also a feasible solution for problem $(PA)_c'$. Hence it implies that $(q^*(\theta), t^*(\theta))$ is an optimal solution of $(PA)_c'$ with $h(\theta_1, q(\theta_1)) - t(\theta_1) = 0$. \square

Now, we check the satisfaction of the monotonicity condition which we ignored so far for the solution that we found, $q^*(\theta)$.

Theorem 4.1.7. *Let $q^*(\theta)$ be a solution of (4.5). Suppose that $h_{qq}(\cdot) < 0$, $h_{\theta qq}(\cdot) > 0$, $h_{\theta\theta q}(\cdot) < 0$ and SMC holds, if $(\frac{1-F(\theta)}{f(\theta)})' < 1$, then $\frac{dq^*}{d\theta}(\theta) \geq 0$.*

Proof. Taking the derivative of the equation (4.5) with respect to θ gives us

$$\begin{aligned}
C''(q^*(\theta))\frac{dq^*}{d\theta}(\theta) &= \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q^*(\theta)) + \frac{\partial^2 h}{\partial q^2}(\theta, q^*(\theta))\frac{dq^*}{d\theta}(\theta) \\
&\quad - \left(\frac{\partial^3 h}{\partial \theta^2 \partial q}(\theta, q^*(\theta)) + \frac{\partial^3 h}{\partial \theta \partial q^2}(\theta, q^*(\theta))\frac{dq^*}{d\theta}(\theta) \right) \frac{1-F(\theta)}{f(\theta)} \\
&\quad - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q^*(\theta)) \left(\frac{1-F(\theta)}{f(\theta)} \right)'.
\end{aligned}$$

This equation can be rewritten as

$$\begin{aligned} \frac{dq^*}{d\theta}(\theta) \left(\frac{\partial^2 h}{\partial q^2}(\theta, q^*(\theta)) - \frac{\partial^3 h}{\partial \theta \partial q^2}(\theta, q^*(\theta)) \left(\frac{1-F(\theta)}{f(\theta)} \right) - C''(q^*(\theta)) \right) = \\ \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q^*(\theta)) \left(\left(\frac{1-F(\theta)}{f(\theta)} \right)' - 1 \right) + \frac{\partial^3 h}{\partial \theta^2 \partial q}(\theta, q^*(\theta)) \left(\frac{1-F(\theta)}{f(\theta)} \right). \end{aligned}$$

By the assumptions (H1) and (H2), the coefficient of $\frac{dq^*}{d\theta}(\theta)$ has a negative sign. Hence, a sufficient condition for monotonicity of $q^*(\theta)$ is that the right hand side also has a negative sign. Equivalently, a sufficient condition for $\frac{dq^*}{d\theta}(\theta) \geq 0$ is that

$$\left(\frac{1-F(\theta)}{f(\theta)} \right)' < 1.$$

□

Remark 4.1.8. We know that the hazard rate is $\tilde{h}(\theta) = \frac{f(\theta)}{1-F(\theta)}$. Therefore, a sufficient condition for the monotonicity constraint to be satisfied is $(\frac{1}{\tilde{h}(\theta)})' \leq 0$. This means, $\frac{1}{\tilde{h}(\theta)}$ is non-increasing in θ or equivalently, $\tilde{h}(\theta)$ is non-decreasing in θ .

The hazard rate of the most frequently used distributions, e.g. normal, uniform and exponential distributions, is non-decreasing in θ .

Corollary 4.1.9. Suppose that SMC holds and the $f(\theta)$ is the density function whose hazard rate is non-decreasing in θ . If $(q^*(\theta), t^*(\theta))$ is an optimal solution of $(PA)_c''$ which satisfies

$$\begin{cases} h(\theta_1, q^*(\theta_1)) - t^*(\theta_1) = 0, & (IR_{\theta_1}) \\ \frac{\partial h}{\partial q}(\theta, q^*(\theta)) \frac{dq^*}{d\theta}(\theta) - \frac{dt^*}{d\theta}(\theta) = 0, & (FOC) \\ C'(q^*(\theta)) = \frac{\partial h}{\partial q}(\theta, q^*(\theta)) - \frac{\frac{\partial^2 h}{\partial \theta \partial q}(\theta, q^*(\theta))(1-F(\theta))}{f(\theta)}, & (C) \end{cases}$$

then it is a solution of $(PA)_c'$.

The hazard rate decreases in θ if the density $f(\theta)$ decreases too rapidly in θ . In other words, the hazard rate decreases if the higher-taste customer becomes much less likely to come to the restaurant. In this case, the monotonicity condition might be violated for the solution of the relaxed problem. We now consider the original problem $(PA)_c'$ without removing the monotonicity constraint. Using Theorem 4.1.5

(equations (4.3) and (4.4)) the problem $(PA)_c'$ is equivalent to

$$\begin{aligned} \max_{q(\theta)} \quad & \int_{\theta_1}^{\theta_2} \left(h(\theta, q(\theta)) - C(q(\theta)) - \frac{\partial h}{\partial \theta}(\theta, q(\theta)) \frac{1}{\tilde{h}(\theta)} \right) f(\theta) d\theta, \\ \text{s.t.} \quad & q'(\theta) \geq 0. \end{aligned}$$

Suppose that $q'(\theta)$ changes sign in finitely many times. We start the process of ironing. The above problem can be reformulated as the following optimal control problem

$$\begin{aligned} (PA)_{oc} \quad & \max_{q(\theta), \mu(\theta)} \int_{\theta_1}^{\theta_2} \left(h(\theta, q(\theta)) - C(q(\theta)) - \frac{\partial h}{\partial \theta}(\theta, q(\theta)) \frac{1}{\tilde{h}(\theta)} \right) f(\theta) d\theta, \\ \text{s.t.} \quad & q'(\theta) = \mu(\theta), \quad a.e. \\ & \mu(\theta) \geq 0 \end{aligned}$$

where $\mu(\theta)$ is the control variable and $q(\theta)$ is the state variable.

Theorem 4.1.10. *Suppose that q' has finitely many discontinuities in $[\theta_1, \theta_2]$. Then $(q(\theta), t(\theta))$ is an optimal solution of $(PA)_{oc}$ if and only if in any of the subintervals (without loss of generality $[\theta_3, \theta_4]$) in which q' is continuous, either*

(1) $q'(\theta) > 0, \quad \forall \theta \in [\theta_3, \theta_4]$, in which case (C) holds, i.e.,

$$C'(q(\theta)) = \frac{\partial h}{\partial q}(\theta, q(\theta)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q(\theta)) \frac{1}{\tilde{h}(\theta)}, \quad \forall \theta \in [\theta_3, \theta_4]$$

or,

(2) $q'(\theta) = 0$, for all $[\theta_3, \theta_4]$ in which case $q(\theta) = \bar{q}$ for all $\theta \in [\theta_3, \theta_4]$. \bar{q} , θ_3 and θ_4 can be determined by the system of three equations with three unknowns

$$\begin{aligned} C'(\bar{q}) &= \frac{\int_{\theta_3}^{\theta_4} \left(\frac{\partial h}{\partial q}(\theta, \bar{q}) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, \bar{q}) \frac{1}{\tilde{h}(\theta)} \right) f(\theta) d\theta}{\int_{\theta_3}^{\theta_4} f(\theta) d\theta}, \\ C'(\bar{q}) &= C'(q(\theta_3)) = \frac{\partial h}{\partial q}(\theta_3, q(\theta_3)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta_3, q(\theta_3)) \frac{1}{\tilde{h}(\theta_3)}, \\ C'(\bar{q}) &= C'(q(\theta_4)) = \frac{\partial h}{\partial q}(\theta_4, q(\theta_4)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta_4, q(\theta_4)) \frac{1}{\tilde{h}(\theta_4)}. \end{aligned}$$

Proof. The Hamiltonian associated with the problem $(PA)_{oc}$ is

$$H(\theta, q, \mu, \lambda) = \left(h(\theta, q) - C(q) - \frac{\partial h}{\partial \theta}(\theta, q) \frac{1}{\tilde{h}(\theta)} \right) f(\theta) + \lambda \mu$$

which is concave in q and μ by (H1) and (H2).

Using the Pontryagin Maximum Principle (Theorem 2.4.1), $(q(\theta), \mu(\theta))$ is optimal if and only if there exist absolutely continuous functions $\lambda(\theta) = (\lambda_1(\theta), \dots, \lambda_n(\theta))$ for all $\theta_1 \leq \theta \leq \theta_2$ such that

1. $\mu(\cdot)$ maximizes the Hamiltonian function, that is,

$$H(\theta, q(\theta), \mu(\theta), \lambda(\theta)) = \max_{\mu \geq 0} H(\theta, q(\theta), \mu, \lambda(\theta)).$$

- 2.

$$\begin{aligned} \frac{d\lambda(\theta)}{d\theta} &= -\frac{\partial H}{\partial q}(\cdot) \\ &= -\left(\frac{\partial h}{\partial q}(\theta, q(\theta)) - C'(q(\theta)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q(\theta)) \frac{1}{\tilde{h}(\theta)} \right) f(\theta). \end{aligned} \quad (4.6)$$

3. The transversality conditions, $\lambda(\theta_1) = \lambda(\theta_2) = 0$.

The first condition implies that $\lambda(\theta) \leq 0$. Moreover, $\mu(\theta) = 0$ whenever $\lambda < 0$.

Integrating the equation (4.6) yields

$$\lambda(\theta) = -\int_{\theta_1}^{\theta} \left(\frac{\partial h}{\partial q}(\theta, q(\theta)) - C'(q(\theta)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q(\theta)) \frac{1}{\tilde{h}(\theta)} \right) f(\theta) d\theta, \quad \forall \theta \in (\theta_1, \theta_2). \quad (4.7)$$

Suppose that $q(\theta)$ is strictly increasing over $[\theta_3, \theta_4]$, which means that, $\mu(\theta) = q'(\theta)$ is larger than zero. Then, $\lambda(\theta)$ is equal to zero on this interval. So, $\frac{d\lambda}{d\theta}(\theta)$ is equal to zero. Using the second condition of Pontryagin Maximum Principle, it yields

$$C'(q(\theta)) = \frac{\partial h}{\partial q}(\theta, q(\theta)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q(\theta)) \frac{1}{\tilde{h}(\theta)}, \quad (4.8)$$

which is exactly the solution (C) on interval $[\theta_3, \theta_4]$

Now, we consider the case where $q(\theta) = \bar{q}$ for $\theta \in [\theta_3, \theta_4]$. Since $\mu(\theta)$ is not equal to zero at the left hand side of θ_3 and right hand side of θ_4 , then

$$\lambda(\theta_3) = \lambda(\theta_4) = 0.$$

Therefore, using equation (4.7), we have

$$\begin{aligned}\lambda(\theta_3) &= - \int_{\theta_1}^{\theta_3} \left(\frac{\partial h}{\partial q}(\theta, q(\theta)) - C'(q(\theta)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q(\theta)) \frac{1}{\tilde{h}(\theta)} \right) f(\theta) d\theta \\ &= 0,\end{aligned}$$

and

$$\begin{aligned}\lambda(\theta_4) &= - \int_{\theta_1}^{\theta_4} \left(\frac{\partial h}{\partial q}(\theta, q(\theta)) - C'(q(\theta)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q(\theta)) \frac{1}{\tilde{h}(\theta)} \right) f(\theta) d\theta \\ &= 0.\end{aligned}$$

Hence,

$$\int_{\theta_3}^{\theta_4} \left(\frac{\partial h}{\partial q}(\theta, q(\theta)) - C'(q(\theta)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, q(\theta)) \frac{1}{\tilde{h}(\theta)} \right) f(\theta) d\theta = 0.$$

Thus,

$$C'(\bar{q}) = \frac{\int_{\theta_3}^{\theta_4} \left(\frac{\partial h}{\partial q}(\theta, \bar{q}) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta, \bar{q}) \frac{1}{\tilde{h}(\theta)} \right) f(\theta) d\theta}{\int_{\theta_3}^{\theta_4} f(\theta) d\theta}. \quad (4.9)$$

On the left hand side of θ_3 and right hand side of θ_4 where $q(\theta)$ is strictly increasing, we proved that \bar{q} should coincide with (C). Hence, using equation (C) at θ_3 and θ_4 and the continuity of \bar{q} implies that

$$C'(\bar{q}) = C'(q(\theta_3)) = \frac{\partial h}{\partial q}(\theta_3, q(\theta_3)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta_3, q(\theta_3)) \frac{1}{\tilde{h}(\theta_3)},$$

$$C'(\bar{q}) = C'(q(\theta_4)) = \frac{\partial h}{\partial q}(\theta_4, q(\theta_4)) - \frac{\partial^2 h}{\partial \theta \partial q}(\theta_4, q(\theta_4)) \frac{1}{\tilde{h}(\theta_4)}.$$

Now, considering the above equation we can find the value of θ_3 , θ_4 and \bar{q} from the above equations and equation (4.9). \square

The interval $[\theta_3, \theta_4]$ is called the bunching interval. To have a better view of bunching process consider that the density function is in the lowest position on interval $[\theta_3, \theta_4]$. In the bunching process the owner keeps the same quality for this interval. In this case, the customer will pay a lower transfer but there is another advantage for the owner. Since the incentive constraints are relaxed for all $\theta \geq \theta_4$, he can receive higher transfer from all those types. The profit of the owner will be really huge if the

likelihood of the buyer's type falls sufficiently low.

4.2 Discrete problem

In this section, we assume that n is the number of the customers with different tastes. Suppose that θ_i denotes the i -th customer's taste and $\theta_1 < \theta_2 < \dots < \theta_n$. We consider a discretized version of problem $(PA)_c$,

$$(PA)_d \quad \max_{q_i, t_i} \quad \sum_{i=1}^n (t_i - C(q_i)) f_i$$

$$\text{s.t.} \quad h(\theta_i, q_i) - t_i \geq 0, \quad i = 1, \dots, n, \quad (\text{IR})$$

$$h(\theta_i, q_i) - t_i \geq h(\theta_i, q_j) - t_j, \quad i \neq j = 1, \dots, n \quad (\text{IC})$$

Definition 4.2.1. (*Discrete Spence Mirrlees condition*)

A function $h(\theta, q)$ satisfies the discrete Spence Mirrlees condition if

$$\forall \theta_i > \theta_j, \forall q_m > q_n \quad h(\theta_i, q_m) - h(\theta_j, q_m) > h(\theta_i, q_n) - h(\theta_j, q_n).$$

Remark 4.2.2. In fact, the Spence Mirrlees condition states that $\forall \theta' > \theta \quad h(\theta', q) - h(\theta, q)$ increases or decreases in q . For simplicity we assume that it is always increasing in q . The result under SMC when it is decreasing is similar.

In this section, we assume that the following hypotheses hold.

H3. Suppose that $h(\theta, q)$ is an increasing function of θ .

H4. Suppose that $C(\cdot)$ is a twice differentiable, increasing and strictly convex function such that $C'(0) = 0$.

Remark 4.2.3. If the function $h(\theta, q)$ is increasing in θ , then the (IR) constraints can be replaced by $h(\theta_1, q_1) - t_1 \geq 0$.

Proof. Using the facts that for every $i \neq 1$, $\theta_i > \theta_{i-1}$ and $h(\theta, q)$ is increasing in θ we have:

$$\begin{aligned} h(\theta_i, q_i) - t_i &\geq h(\theta_i, q_1) - t_1 && \text{by (IC),} \\ &\geq h(\theta_1, q_1) - t_1 && \text{since } h(\theta, q) \text{ is increasing in } \theta, \\ &\geq 0 && \text{by (IR).} \end{aligned}$$

This gives all the other individual rationality constraints. \square

Same as the problem of continuum of types, the strategy is to rewrite the problem in the simplest possible form.

Theorem 4.2.4. *Let the vector (q, t) with $q = (q_1, \dots, q_n)$, $t = (t_1, \dots, t_n)$ be an optimal solution for problem $(PA)_d$. Then the (IR) constraint is active at θ_1 , that is, $h(\theta_1, q_1) - t_1 = 0$.*

Proof. The proof is by contradiction. Suppose that (q, t) is an optimal solution and $h(\theta_1, q_1) - t_1 > 0$. Then,

$$\begin{aligned} h(\theta_i, q_i) - t_i &\geq h(\theta_i, q_1) - t_1 && \text{by (IC),} \\ &\geq h(\theta_1, q_1) - t_1 && \text{since } h(\theta, q) \text{ is increasing in } \theta, \\ &> 0. \end{aligned}$$

Consequently, for all i ,

$$h(\theta_i, q_i) - (t_i + \epsilon) > 0, \quad \text{for some small enough } \epsilon > 0.$$

One can find $\epsilon > 0$ small enough such that $(q_i, t_i + \epsilon)$ is still feasible but give a bigger objective value in $(PA)_d$. This violates the fact that (q, t) is an optimal solution. \square

Now, we want to reduce the (IC) constraints.

Theorem 4.2.5. *Under the discrete Spence Mirrlees condition a feasible solution (q, t) of $(PA)_d$ satisfies the (IC) constraints if and only if it satisfies*

- (1) *the monotonicity constraint, that is, for all i*

$$q_{i+1} \geq q_i \geq q_{i-1},$$

and

- (2) *the set of local (IC) constraints, that is, for all i*

$$\begin{aligned} h(\theta_i, q_i) - t_i &\geq h(\theta_i, q_{i+1}) - t_{i+1}, \\ h(\theta_i, q_i) - t_i &\geq h(\theta_i, q_{i-1}) - t_{i-1}. \end{aligned}$$

Proof. First, assume that (q_i, t_i) satisfies the (IC) constraints. Then, it is clear that it satisfies (2), since (2) is a subset of constraints we defined in (IC). The proof of (1)

is by contradiction. With no loss of generality, suppose $\theta_i \leq \theta_j$, then we will show that $q_i \leq q_j$. The incentive compatibility constraints are

$$h(\theta_i, q_i) - t_i \geq h(\theta_i, q_j) - t_j,$$

$$h(\theta_j, q_j) - t_j \geq h(\theta_j, q_i) - t_i.$$

Adding these two constraints, implies

$$h(\theta_j, q_j) - h(\theta_j, q_i) \geq h(\theta_i, q_j) - h(\theta_i, q_i). \quad (4.10)$$

Now, if $q_j - q_i < 0$, then Spence Mirrlees condition implies

$$h(\theta_j, q_i) - h(\theta_i, q_i) > h(\theta_j, q_j) - h(\theta_i, q_j).$$

Equivalently,

$$h(\theta_i, q_j) - h(\theta_i, q_i) > h(\theta_j, q_j) - h(\theta_j, q_i)$$

which contradicts equation (4.10).

Now, we want to prove the reverse. Suppose (q, t) satisfies (1) and (2). Consider the three tastes $\theta_{i-1} < \theta_i < \theta_{i+1}$ and the following local downward incentive constraints:

$$h(\theta_{i+1}, q_{i+1}) - t_{i+1} \geq h(\theta_{i+1}, q_i) - t_i,$$

$$h(\theta_i, q_i) - t_i \geq h(\theta_i, q_{i-1}) - t_{i-1}.$$

Adding these constraints yields

$$h(\theta_{i+1}, q_{i+1}) - h(\theta_{i+1}, q_i) + h(\theta_i, q_i) - h(\theta_i, q_{i-1}) \geq t_{i+1} - t_{i-1}.$$

Since we know that $h(\theta, q)$ satisfies the Spence Mirrlees condition or equivalently, $h(\theta_{i+1}, q_i) - h(\theta_i, q_i) \geq h(\theta_{i+1}, q_{i-1}) - h(\theta_i, q_{i-1})$, we have

$$h(\theta_{i+1}, q_{i+1}) - h(\theta_{i+1}, q_{i-1}) \geq t_{i+1} - t_{i-1}.$$

Rearranging the above equation will give us one of the global downward incentive compatible constraints,

$$h(\theta_{i+1}, q_{i+1}) - t_{i+1} \geq h(\theta_{i+1}, q_{i-1}) - t_{i-1}.$$

Hence, we could reduce the set of downward incentive constraints to a set of local downward incentive constraints plus the monotonicity condition $q_i \geq q_{i-1}$. The same conclusion is true for the set of upward constraints. Consider $\theta_{i-1} < \theta_i < \theta_{i+1}$ and only local upwards constraints

$$\begin{aligned} h(\theta_{i-1}, q_{i-1}) - t_{i-1} &\geq h(\theta_{i-1}, q_i) - t_i \\ h(\theta_i, q_i) - t_i &\geq h(\theta_i, q_{i+1}) - t_{i+1}. \end{aligned}$$

Adding the above constraints, using Spence Mirrlees condition and the fact that $q_{i+1} \geq q_i$, gives the non-local incentive compatible constraints: $h(\theta_{i-1}, q_{i-1}) - t_{i-1} \geq h(\theta_{i-1}, q_{i+1}) - t_{i+1}$. \square

Using Theorem 4.2.5 and Remark 4.2.3 under the discrete SMC we can reformulate the problem $(PA)_d$ as follows:

$$\begin{aligned} \max_{q_i, t_i} \quad & \sum_{i=1}^n (t_i - C(q_i)) f_i \\ \text{s.t.} \quad & h(\theta_1, q_1) - t_1 \geq 0, & (\text{IR}_{\theta_1}) \\ & h(\theta_i, q_i) - t_i \geq h(\theta_i, q_{i+1}) - t_{i+1}, & \forall i = 1, \dots, n, & (\text{LU}) \\ & h(\theta_i, q_i) - t_i \geq h(\theta_i, q_{i-1}) - t_{i-1}, & \forall i = 1, \dots, n, & (\text{LD}) \\ & q_i \geq q_j, & \forall i > j. & (\text{M}) \end{aligned}$$

We know by Theorem 4.2.4 that the constraint $h(\theta_1, q_1) - t_1 \geq 0$ is active at an optimal solution. Hence an optimal solution (q, t) of $(PA)_d$ must be an optimal solution of the following problem

$$\begin{aligned} (PA)_d' \quad \max_{q_i, t_i} \quad & \sum_{i=1}^n (t_i - C(q_i)) f_i \\ \text{s.t.} \quad & h(\theta_1, q_1) - t_1 = 0, & (\text{IR}_{\theta_1}) \\ & h(\theta_i, q_i) - t_i \geq h(\theta_i, q_{i+1}) - t_{i+1}, & \forall i = 1, \dots, n, & (\text{LU}) \\ & h(\theta_i, q_i) - t_i \geq h(\theta_i, q_{i-1}) - t_{i-1}, & \forall i = 1, \dots, n, & (\text{LD}) \\ & q_i \geq q_j, & \forall i > j. & (\text{M}) \end{aligned}$$

where (LU) represents the set of local upward constraints and (LD) denotes the set of local downward constraints.

Theorem 4.2.6. *Let the vector (q, t) be an optimal solution of $(PA)_d'$. Then under Spence Mirrlees condition*

- (1) *all the local downward incentive constraints are active, and consequently*
- (2) *the set of local downward incentive constraints together with the monotonicity condition (M) implies the set of local upward incentive constraints.*

Proof. The proof of (1) is by contradiction. Suppose that one of the constraints is not active, i.e., $h(\theta_i, q_i) - t_i > h(\theta_i, q_{i-1}) - t_{i-1}$. Then,

$$\begin{aligned}
 h(\theta_i, q_i) - t_i &> h(\theta_i, q_{i-1}) - t_{i-1} && \text{by local downward constraints,} \\
 &\geq h(\theta_{i-1}, q_{i-1}) - t_{i-1} && \text{since } h(\theta, q) \text{ is increasing in } \theta, \\
 &\geq h(\theta_{i-1}, q_{i-2}) - t_{i-2} && \text{by local downward constraints,} \\
 &\vdots \\
 &\geq h(\theta_1, q_1) - t_1 && \text{since } h(\theta, q) \text{ is increasing in } \theta, \\
 &= 0 && \text{by Theorem 4.2.4.}
 \end{aligned}$$

Then, for all $j \geq i$, $h(\theta_j, q_j) - (t_j + \epsilon) > 0$ for some small enough $\epsilon > 0$. Such a $(q_j, t_j + \epsilon)$ is still feasible but gives a bigger objective value in $(PA)_d'$. This violates the fact that (q_i, t_i) is an optimal solution. In simple words, by increasing all t_j where $j \geq i$ with the same rate the owner of the restaurant can earn more money while all the constraints are satisfied. This is in contradiction with being at the optimum point.

Now, we want to prove (2). We know from step (1) that the set of local downward constraints is active, i.e.,

$$h(\theta_i, q_i) - t_i = h(\theta_i, q_{i-1}) - t_{i-1}.$$

Equivalently,

$$h(\theta_i, q_i) - h(\theta_i, q_{i-1}) = t_i - t_{i-1}.$$

Since h satisfies the discrete SMC and $q_i \geq q_{i-1}$, we have

$$h(\theta_i, q_i) - h(\theta_i, q_{i-1}) = t_i - t_{i-1} \geq h(\theta_{i-1}, q_i) - h(\theta_{i-1}, q_{i-1}).$$

The above equation can be rewritten as

$$h(\theta_{i-1}, q_{i-1}) - t_{i-1} \geq h(\theta_{i-1}, q_i) - t_i$$

which is the local upward constraint. \square

Since we deal with an optimization problem over (q, t) , we are going to use the Kuhn-Tucker or KKT condition to solve it. Although we already know that the KKT condition is necessary and sufficient for concave programming problems, problem $(PA)_d$ does not belong to the class of concave programming problems in general. Since even if $h(\theta, q)$ is a concave function, there is no guarantee that their difference $(h(\theta, q) - h(\theta, q'))$ is concave as well. So, we are going to consider two different cases: where we are dealing with a concave programming and where we are dealing with a non-concave programming problem. In the first case, we can find the solution explicitly if the solution exists. However, in the second case, KKT condition just gives us a set of stationary points to which the global maximum belongs.

Theorem 4.2.7. *Under Spence Mirrlees condition, the problem $(PA)_d$ always satisfies LICQ.*

Proof. LICQ holds if the gradient of the active inequality constraints and the gradient of the equality constraints are linearly independent at the optimum. By Theorems 4.2.4 and 4.2.6 we know that only (LD) and (IR_{θ_1}) are active at the optimal point. With no loss of generality, we are assuming that the monotonicity constraints are active at the optimum and prove that the problem always satisfies the LICQ. On that case even if at least a subset of the monotonicity constraints are inactive, the problem still satisfies the LICQ. Since any subset of linearly independent vectors is itself linearly independent, we are considering the worst case where all of our inequality constraints are active at the optimum. In order to have an idea of the proof, let's first assume $n = 2$. In this case the active constraints at the optimum are:

$$\begin{aligned} h(\theta_1, q_1) - t_1 &= 0, \\ h(\theta_2, q_2) - t_2 &= h(\theta_2, q_1) - t_1 \\ q_2 &= q_1. \end{aligned}$$

Hence, the equations are

$$\lambda_1 \begin{pmatrix} -\frac{\partial h}{\partial q}(\theta_1, q_1) \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_{21} \begin{pmatrix} \frac{\partial h}{\partial q}(\theta_2, q_1) \\ -\frac{\partial h}{\partial q}(\theta_2, q_2) \\ -1 \\ 1 \end{pmatrix} + \mu_{21} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \mathbf{0}.$$

It implies $\lambda_1 = \lambda_{21} = \mu_{21} = 0$. Thus, these vectors are linearly independent.

In general case of having n taste-type customers we are dealing with the following gradient vectors

$$\lambda_1 \begin{pmatrix} -\frac{\partial h}{\partial q}(\theta_1, q_1) \\ \mathbf{0} \\ \text{---} \\ 1 \\ \mathbf{0} \end{pmatrix} + \sum_{i=2}^n \lambda_{i(i-1)} \begin{pmatrix} \mathbf{0} \\ \frac{\partial h}{\partial q}(\theta_i, q_{i-1}) \\ -\frac{\partial h}{\partial q}(\theta_i, q_i) \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \\ -1 \\ 1 \\ \mathbf{0} \end{pmatrix} \begin{matrix} (i-1)\text{-th row} \\ i\text{-th row} \\ \\ \\ \\ (i-1)\text{-th row} \\ i\text{-th row} \end{matrix} \\ + \sum_{i=2}^n \mu_{i(i-1)} \begin{pmatrix} \mathbf{0} \\ -1 \\ 1 \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \end{pmatrix} \begin{matrix} (i-1)\text{-th row} \\ i\text{-th row} \\ \\ \\ \end{matrix} = \mathbf{0}.$$

where the horizontal line (---) divides the matrix into two parts. The first part is the gradient with respect to q and the second part is the gradient with respect to t . The $\mathbf{0}$ denotes the zero vector.

Since the i -th row of the lower part of the matrices for $i \geq 2$ is zero except in the second matrix, it forces all $\lambda_{i(i-1)}$'s to be zero. Right after that consider the first row of the lower part of the matrices which are all zero except for the first matrix forces λ_1 to become zero. With all the other multipliers being zero $\mu_{i(i-1)}$'s will also be forced to be zero. Thus, LICQ holds. \square

Same as the continuous case we first solve the relaxed problem of $(PA)_d'$ with ignoring the monotonicity condition (M). The relaxed problem can be written as

$$\begin{aligned}
(PA)_d'' \quad & \max_{q,t} \quad \sum_{i=1}^n (t_i - C(q_i)) f_i \\
\text{s.t.} \quad & h(\theta_1, q_1) - t_1 = 0, & \text{(IR)} \\
& h(\theta_i, q_i) - t_i = h(\theta_i, q_{i-1}) - t_{i-1}, & \forall i = 2, \dots, n \quad \text{(LD)}
\end{aligned}$$

Theorem 4.2.8. *Suppose that $(PA)_d''$ is a concave programming problem and (q^*, t^*) is a feasible solution of $(PA)_d''$. Then (q^*, t^*) is an optimal solution of $(PA)_d''$ if and only if the following set of conditions holds:*

$$\begin{aligned}
C'(q_n^*) &= \frac{\partial h}{\partial q}(\theta_n, q_n^*) \\
C'(q_i^*) &= \frac{\partial h}{\partial q}(\theta_i, q_i^*) - \frac{\lambda_{i+1} \left(\frac{\partial h}{\partial q}(\theta_{i+1}, q_{i+1}^*) - \frac{\partial h}{\partial q}(\theta_i, q_i^*) \right)}{f_i} \quad \text{where } 1 < i < n. & \text{(C}_d\text{)}
\end{aligned}$$

Proof. The Lagrangian formulation associated with the problem is

$$L = \sum_{i=1}^n [f_i(C(q_i) - t_i)] - \sum_{i=2}^n [\lambda_i(h(\theta_i, q_i) - t_i - h(\theta_i, q_{i-1}) + t_{i-1})] - \mu(h(\theta_1, q_1) - t_1).$$

The first-order conditions for $1 < i < n$ and $i = n$ can be characterized as

$$\begin{aligned}
\partial L / \partial q_i = 0 & \quad \Rightarrow \quad f_i C'(q_i) = -\lambda_{i+1} \frac{\partial h}{\partial q}(\theta_{i+1}, q_i) + \lambda_i \frac{\partial h}{\partial q}(\theta_i, q_i), \\
\partial L / \partial t_i = 0 & \quad \Rightarrow \quad f_i + \lambda_{i+1} = \lambda_i, \\
\partial L / \partial q_n = 0 & \quad \Rightarrow \quad f_n C'(q_n) = \lambda_n \frac{\partial h}{\partial q}(\theta_n, q_n), \\
\partial L / \partial t_n = 0 & \quad \Rightarrow \quad f_n = \lambda_n.
\end{aligned}$$

From the above equations we can find:

$$C'(q_n^*) = \frac{\partial h}{\partial q}(\theta_n, q_n^*)$$

$$C'(q_i^*) = \frac{\partial h}{\partial q}(\theta_i, q_i^*) - \frac{\lambda_{i+1} \left(\frac{\partial h}{\partial q}(\theta_{i+1}, q_i^*) - \frac{\partial h}{\partial q}(\theta_i, q_i^*) \right)}{f_i} \text{ where } 1 < i < n.$$

□

If the problem is non-concave, we have the following remark:

Remark 4.2.9. *If $(PA)_d''$ is a non-concave programming problem then the condition (C_d) is only a necessary condition for the optimality of a feasible point (q^*, t^*) of $(PA)_d''$.*

Now, we need to check if the solution that we found from Theorem 4.2.8 satisfy the monotonicity condition or not. If it satisfies the monotonicity condition it is also the solution of the problem $(PA)'_c$.

Chapter 5

Solutions without the Spence Mirrlees Condition

In the previous chapter we considered the problem under the Spence Mirrlees Condition. In that case, the Spence Mirrlees Condition gives us the monotonicity in q and as a result we have many reductions on the set of incentive compatible constraints that enable us to achieve the satisfaction of LICQ. However, when the problem does not satisfy the Spence Mirrlees Condition we cannot go through those simplifications any more. Although there are lots of utility functions that do not satisfy the Spence Mirrlees Condition (SMC), I only know Araujo and Moreira [3] who have studied the problem of continuum of types without the SMC. In their paper they characterized the problem in a special case using the decreasing limit curve that separates the region $\Theta \times R_+^m$ into two parts where the SMC is positive (above the limit curve) and where it is negative (below the limit curve). In this chapter, we are going to study the discrete problem when SMC does not hold.

5.1 Discrete problem

Considering the original problem

$$\begin{aligned}
 (PA_d) \quad & \max_{q_i, t_i} \sum_{i=1}^n (t_i - C(q_i)) f_i \\
 & \text{s.t.} \quad h(\theta_i, q_i) - t_i \geq 0, & i = 1, \dots, n, & \quad (\text{IR}) \\
 & \quad \quad h(\theta_i, q_i) - t_i \geq h(\theta_i, q_j) - t_j, & i \neq j = 1, \dots, n & \quad (\text{IC})
 \end{aligned}$$

the aim of this chapter is to find sufficient and easy to check conditions or methods which guarantee the satisfaction of at least one of the constraint qualifications discussed in Chapter 2. In that case, we can derive the KKT condition for the problem and solve it explicitly when the problem is concave and find the set of stationary points when the problem is not concave. We are going to start with the most common constraint qualifications: Linear Independence Constraint Qualification (LICQ) and Mangasarian-Fromovitz Constraint Qualification (MFCQ).

5.2 Linear independence constraint qualification (LICQ)

Although the discrete problem under the Spence Mirrlees Condition always satisfies the LICQ, it is not the case when we do not have the Spence Mirrlees Condition. The difficulty in this case is that we are not able to say which of our constraints are active and which are inactive at the optimum. We have to consider the worst case where all of the inequality constraints are active at the optimum. As explained in the previous chapter, any subset of the linearly independent vectors is itself linearly independent. This means we are after a set of sufficient conditions for the satisfaction of the LICQ.

Theorem 5.2.1. *LICQ holds for $n = 2$ if,*

$$\left\{ \begin{array}{l} \frac{\partial h}{\partial q}(\theta_1, q_1) \neq \frac{\partial h}{\partial q}(\theta_2, q_1); \\ \text{and} \\ \frac{\partial h}{\partial q}(\theta_2, q_2) \neq \frac{\partial h}{\partial q}(\theta_1, q_2). \end{array} \right.$$

Proof. When we have only 2 cutomers, the constraints of the problem are:

$$\begin{aligned} h(\theta_1, q_1) - t_1 &\geq 0, \\ h(\theta_2, q_2) - t_2 &\geq 0, \\ h(\theta_1, q_1) - t_1 &\geq h(\theta_1, q_2) - t_2, \\ h(\theta_2, q_2) - t_2 &\geq h(\theta_2, q_1) - t_1. \end{aligned}$$

Suppose all of our constraints are active at optimum. Then the linear combination of

the gradient vectors can be written as

$$\lambda_1 \begin{pmatrix} -\frac{\partial h}{\partial q}(\theta_1, q_1) \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -\frac{\partial h}{\partial q}(\theta_2, q_2) \\ 0 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} -\frac{\partial h}{\partial q}(\theta_1, q_1) \\ \frac{\partial h}{\partial q}(\theta_1, q_2) \\ 1 \\ -1 \end{pmatrix} + \lambda_4 \begin{pmatrix} \frac{\partial h}{\partial q}(\theta_2, q_1) \\ -\frac{\partial h}{\partial q}(\theta_2, q_2) \\ -1 \\ 1 \end{pmatrix} = 0.$$

This gives four equations with four unknowns, $\lambda_1, \lambda_2, \lambda_3, \lambda_4$,

$$\begin{cases} -\lambda_1 \frac{\partial h}{\partial q}(\theta_1, q_1) - \lambda_3 \frac{\partial h}{\partial q}(\theta_1, q_1) + \lambda_4 \frac{\partial h}{\partial q}(\theta_2, q_1) = 0, \\ -\lambda_2 \frac{\partial h}{\partial q}(\theta_2, q_2) + \lambda_3 \frac{\partial h}{\partial q}(\theta_1, q_2) - \lambda_4 \frac{\partial h}{\partial q}(\theta_2, q_2) = 0, \\ \lambda_1 + \lambda_3 - \lambda_4 = 0, \\ \lambda_2 - \lambda_3 + \lambda_4 = 0. \end{cases}$$

Finding λ_1 and λ_2 from the last two equations and substituting them back into the rest of the equations yields

$$\begin{cases} -\lambda_4 \frac{\partial h}{\partial q}(\theta_1, q_1) + \lambda_4 \frac{\partial h}{\partial q}(\theta_2, q_1) = 0, \\ -\lambda_3 \frac{\partial h}{\partial q}(\theta_2, q_2) + \lambda_3 \frac{\partial h}{\partial q}(\theta_1, q_2) = 0. \end{cases}$$

By the above equations, it is obvious that if

$$\begin{cases} \frac{\partial h}{\partial q}(\theta_1, q_1) \neq \frac{\partial h}{\partial q}(\theta_2, q_1); \\ \text{and} \\ \frac{\partial h}{\partial q}(\theta_2, q_2) \neq \frac{\partial h}{\partial q}(\theta_1, q_2) \end{cases}$$

then all the $\lambda_i, i = 1, 2, 3, 4$ have to be zero. This implies the satisfaction of LICQ. \square

The theorem below shows the weakness of LICQ for dealing with the principal-agent problem without the Spence Mirrlees Condition in general.

Theorem 5.2.2. *LICQ is never satisfied for $n \geq 3$ when all of our constraints are active.*

Proof. Suppose λ_i are the coefficients of the gradient vectors. Then LICQ holds if

$$\sum_{i=1}^{n^2} \lambda_i \begin{pmatrix} \frac{\partial W_i}{\partial q_1} \\ \vdots \\ \frac{\partial W_i}{\partial q_n} \\ \frac{\partial W_i}{\partial t_1} \\ \vdots \\ \frac{\partial W_i}{\partial t_n} \end{pmatrix} = 0,$$

implies all λ_i are zero where W_i represents the i -th constraint. The number of λ_i s is equal to the number of constraints, which is n^2 . On the other hand, the number of equations is equal to the number of components of each gradient matrix, which is $2n$. For $n \geq 3$ the number of λ_i s is larger than the number of equations. Hence, λ_i s are always dependent. \square

In the next section, we will see a weaker constraint qualification called MFCQ.

5.3 Mangasarian-Fromovitz constraint qualification (MFCQ)

First, recall that the Mangasarian-Fromovitz Constraint Qualification (MFCQ) holds if and only if the gradient of the active inequality constraints and the gradient of the active constraints are positively linearly independent. We assume the worst case scenario, which is the case in which all of the inequality constraints are active at the optimum. If only a subset of the inequality constraints were active, then our sufficient condition would still make sure that the vectors corresponding to these quality constraints are positively linearly independent because the set of vectors would be a subset of all the vectors that we prove to be positively linearly independent under our sufficient condition.

Definition 5.3.1. (*Principal-Agent Problem Constraint Qualification*)

Principal-Agent Problem Constraint Qualification (PAPCQ) holds at a certain point if for each fixed value of j we have,

$$\frac{\partial h}{\partial q}(\theta_i, q_j) > \frac{\partial h}{\partial q}(\theta_j, q_j), \quad \forall i \neq j.$$

Then we have the following theorem.

Theorem 5.3.2. *MFCQ holds if PAPCQ holds.*

Proof. Suppose λ_i is the multiplier of the (IR) constraint associated with the customer with type θ_i and λ_{ij} is the multiplier of the (IC) constraint associated with the customer with type θ_i who likes to hide his type as θ_j . In order to have an idea about the proof, let's first assume $n = 2$. In this case the equations are:

$$\lambda_1 \begin{pmatrix} -\frac{\partial h}{\partial q}(\theta_1, q_1) \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -\frac{\partial h}{\partial q}(\theta_2, q_2) \\ 0 \\ 1 \end{pmatrix} + \lambda_{12} \begin{pmatrix} -\frac{\partial h}{\partial q}(\theta_1, q_1) \\ \frac{\partial h}{\partial q}(\theta_1, q_2) \\ 1 \\ -1 \end{pmatrix} + \lambda_{21} \begin{pmatrix} \frac{\partial h}{\partial q}(\theta_2, q_1) \\ -\frac{\partial h}{\partial q}(\theta_2, q_2) \\ -1 \\ 1 \end{pmatrix} = 0$$

which gives us four equations with four unknowns, $\lambda_1, \lambda_2, \lambda_{12}$ and λ_{21} ,

$$\begin{cases} -\lambda_1 \frac{\partial h}{\partial q}(\theta_1, q_1) - \lambda_{12} \frac{\partial h}{\partial q}(\theta_1, q_1) + \lambda_{21} \frac{\partial h}{\partial q}(\theta_2, q_1) = 0, \\ -\lambda_2 \frac{\partial h}{\partial q}(\theta_2, q_2) + \lambda_{12} \frac{\partial h}{\partial q}(\theta_1, q_2) - \lambda_{21} \frac{\partial h}{\partial q}(\theta_2, q_2) = 0, \\ \lambda_1 + \lambda_{12} - \lambda_{21} = 0, \\ \lambda_2 - \lambda_{12} + \lambda_{21} = 0. \end{cases}$$

Substituting $\lambda_1 = -\lambda_{12} + \lambda_{21}$ and $\lambda_2 = \lambda_{12} - \lambda_{21}$ in the rest of the equations gives us

$$\begin{cases} \lambda_{12} \left(\frac{\partial h}{\partial q}(\theta_1, q_2) - \frac{\partial h}{\partial q}(\theta_2, q_2) \right) = 0, \\ \lambda_{21} \left(\frac{\partial h}{\partial q}(\theta_2, q_1) - \frac{\partial h}{\partial q}(\theta_1, q_1) \right) = 0. \end{cases}$$

By the above equations, it is obvious that if

$$\begin{cases} \frac{\partial h}{\partial q}(\theta_1, q_1) \neq \frac{\partial h}{\partial q}(\theta_2, q_1); \\ \text{and} \\ \frac{\partial h}{\partial q}(\theta_2, q_2) \neq \frac{\partial h}{\partial q}(\theta_1, q_2), \end{cases}$$

then λ_{12} and λ_{21} have to be zero which completes the proof. As shown before in this case we have even LICQ which is a stronger condition. In the general case of

having n taste-type cutomers we deal with the following gradient vectors

$$\sum_{i=1}^n \lambda_i \begin{pmatrix} \mathbf{0} \\ -\frac{\partial h}{\partial q}(\theta_i, q_i) \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \\ 1 \\ \mathbf{0} \end{pmatrix} \begin{matrix} i\text{-th row} \\ \\ \\ \\ i\text{-th row} \end{matrix} + \sum_{i=1}^n \sum_{j=1}^n \lambda_{ij} \begin{pmatrix} \mathbf{0} \\ -\frac{\partial h}{\partial q}(\theta_i, q_i) \\ \frac{\partial h}{\partial q}(\theta_i, q_j) \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \\ 1 \\ -1 \\ \mathbf{0} \end{pmatrix} \begin{matrix} i\text{-th row} \\ j\text{-th row} \\ \\ \\ \\ i\text{-th row} \\ j\text{-th row} \end{matrix} = 0.$$

This implies the set of equations:

$$\begin{cases} -\lambda_i \frac{\partial h}{\partial q}(\theta_i, q_i) - \sum_{j=1, j \neq i}^n \lambda_{ij} \frac{\partial h}{\partial q}(\theta_i, q_i) + \sum_{j=1, j \neq i}^n \lambda_{ji} \frac{\partial h}{\partial q}(\theta_j, q_i) = 0, & \forall 1 \leq i \leq n \\ \lambda_i + \sum_{j=1}^n \lambda_{ij} - \sum_{j=1}^n \lambda_{ji} = 0. & \forall 1 \leq i \leq n \end{cases}$$

From the last equation we know that $\lambda_i = -\sum_{j=1}^n \lambda_{ij} + \sum_{j=1}^n \lambda_{ji}$ for all i . Substituting in the first equation yields

$$\sum_{j=1}^k \lambda_{ji} \left(-\frac{\partial h}{\partial q}(\theta_j, q_i) + \frac{\partial h}{\partial q}(\theta_i, q_i) \right) = 0, \quad \forall i.$$

Hence, if for all $\left(-\frac{\partial h}{\partial q}(\theta_j, q_i) + \frac{\partial h}{\partial q}(\theta_i, q_i) \right)$ have the same sign then all the λ_{ji} s will be forced to be zero. Thus, we will have MFCQ. \square

If the problem satisfies the PAPCQ then we are able to write the KKT condition for the problem. If the problem is concave, the KKT condition is necessary and sufficient for optimality. Otherwise it just gives us a set of stationary points that the optimal solution will belong to.

The following theorem characterize the set that the optimal solution belongs to.

Theorem 5.3.3. (*Necessary Optimality Condition*)

Suppose (q^, t^*) be an optimal solution of $(PA)_d$, then it should satisfies the following properties:*

The problem can be formulated as:

$$\begin{aligned}
& \max_{q_1, q_2, q_3, t_1, t_2, t_3} && t_1 + t_2 + t_3 - C(q_1) - C(q_2) - C(q_3), \\
& \text{s.t.} && h(\theta_1, q_1) - t_1 \geq 0 \\
& && h(\theta_2, q_2) - t_2 \geq 0, \\
& && h(\theta_3, q_3) - t_3 \geq 0, \\
& && h(\theta_1, q_1) - t_1 \geq h(\theta_1, q_2) - t_2, \\
& && h(\theta_1, q_1) - t_1 \geq h(\theta_1, q_3) - t_3, \\
& && h(\theta_2, q_2) - t_2 \geq h(\theta_2, q_1) - t_1, \\
& && h(\theta_2, q_2) - t_2 \geq h(\theta_2, q_3) - t_3, \\
& && h(\theta_3, q_3) - t_3 \geq h(\theta_3, q_1) - t_1, \\
& && h(\theta_3, q_3) - t_3 \geq h(\theta_3, q_2) - t_2.
\end{aligned}$$

We claim that the optimal point is $(q_1^*, q_2^*, q_3^*, t_1^*, t_2^*, t_3^*) = (0, \frac{5\pi}{4}, \frac{7\pi}{4}, 1, 1, 1)$, since at this point the negative part of the objective function, $C(q)$, will disappear and the positive term, t , achieves its maximum value; using the fact that $h(\theta, q)$ is an upper bound for t . In the other words, $\cos(\cdot)$ is the upper bound for t . So, the maximum value that t can achieve is 1. This point also satisfies all of our constraints. This means (q^*, t^*) is feasible. Hence, this is a real optimal solution for this problem.

The function $h(\theta, q)$ does not satisfy the Spence Mirrlees Condition in the discrete term. To see the point, without loss of generality suppose $q_1 > q_3$. Under SMC we should have,

$$\begin{aligned}
& h(\theta_3, q_1) - h(\theta_2, q_1) > h(\theta_3, q_3) - h(\theta_2, q_3), \\
& h(\theta_3, q_1) - h(\theta_1, q_1) > h(\theta_3, q_3) - h(\theta_1, q_3), \\
& h(\theta_2, q_1) - h(\theta_1, q_1) > h(\theta_2, q_3) - h(\theta_1, q_3).
\end{aligned}$$

This case is impossible since $h(\theta_3, q_1) - h(\theta_1, q_1) = -1$ and $h(\theta_3, q_3) - h(\theta_1, q_3) = 1 - 0.7 = 0.3$. Now, assume $q_1 < q_3$. Then we should have,

$$\begin{aligned}
& h(\theta_3, q_3) - h(\theta_2, q_3) > h(\theta_3, q_1) - h(\theta_2, q_1), \\
& h(\theta_3, q_3) - h(\theta_1, q_3) > h(\theta_3, q_1) - h(\theta_1, q_1), \\
& h(\theta_2, q_3) - h(\theta_1, q_3) > h(\theta_2, q_1) - h(\theta_1, q_1).
\end{aligned}$$

This case is also impossible since, $h(\theta_3, q_3) - h(\theta_2, q_3) = 0.08$ and $h(\theta_3, q_1) - h(\theta_2, q_1) = 0.5$. It is also possible to make the function $h(\theta, q)$ continuous such that the Spence Mirrlees Condition fails. To do that, we need to define a small enough neighborhood around each point θ_1, θ_2 and θ_3 such that their pairwise intersection is empty. Then, the function $h(\theta, q)$ that we already defined can be used in each small neighborhood. Any other function can be defined between those neighborhoods to make the whole function continuous. In that case, since on a small enough neighborhood for each of θ_1, θ_2 and θ_3 the satisfaction of the customer is just a function of q , SMC will fail at θ_1, θ_2 and θ_3 .

Although this function does not satisfy the Spence Mirrlees Condition, it satisfies the PAPCQ. To see that,

$$\frac{\partial h}{\partial q}(\theta, q) = \begin{cases} -\sin q & \theta = \theta_1 \\ -\frac{15}{4} \sin\left(\frac{15}{4}q - \frac{43\pi}{16}\right) & \theta = \theta_2 \\ -\frac{10}{7} \sin\left(\frac{10}{7}q + \frac{3\pi}{2}\right) & \theta = \theta_3. \end{cases}$$

Then, we have:

$$\begin{aligned} \frac{\partial h}{\partial q}(\theta_2, q_1) &= -\frac{15}{4} \sin\left(-\frac{43\pi}{16}\right) &> \frac{\partial h}{\partial q}(\theta_1, q_1) = 0, \\ \frac{\partial h}{\partial q}(\theta_3, q_1) &= -\frac{10}{7} \sin\left(\frac{3\pi}{2}\right) &> \frac{\partial h}{\partial q}(\theta_1, q_1) = 0, \\ \frac{\partial h}{\partial q}(\theta_1, q_2) &= -\sin\left(\frac{5\pi}{4}\right) &> \frac{\partial h}{\partial q}(\theta_2, q_2) = 0, \\ \frac{\partial h}{\partial q}(\theta_3, q_2) &= -\frac{10}{7} \sin\left(\frac{23\pi}{7}\right) &> \frac{\partial h}{\partial q}(\theta_2, q_2) = 0, \\ \frac{\partial h}{\partial q}(\theta_1, q_3) &= -\sin\left(\frac{7\pi}{4}\right) &> \frac{\partial h}{\partial q}(\theta_3, q_3) = 0, \\ \frac{\partial h}{\partial q}(\theta_2, q_3) &= -\frac{15}{4} \sin\left(\frac{31\pi}{8}\right) &> \frac{\partial h}{\partial q}(\theta_3, q_3) = 0. \end{aligned}$$

This means, for all i and j , we have

$$\frac{\partial h}{\partial q}(\theta_i, q_j) > \frac{\partial h}{\partial q}(\theta_j, q_j)$$

which is the PAPCQ condition. So, the problem satisfies the MFCQ constraint qualification at the optimum point. This example can be generalized to any number of

customers.

5.4 Other constraint qualifications and examples

The aim of this section is to consider some particular examples and provide a set of constraint qualifications that might be more applicable to the principal-agent problem. However, some of these are not easily verifiable.

The first class of functions that we want to consider is the class of functions $h(\theta, q)$ which are linear in q . In that case all of our constraints will be linear in q and t . This problem always satisfies the linear constraint qualification which is defined in Chapter 2. Suppose, for example, $h(\theta, q) = (\theta - \theta^*)^2 q$. It is clear that Spence Mirrlees Condition will fail at $\theta = \theta^*$. However, it always satisfies the constraint qualification. So, if the function $C(q)$ is linear, then the KKT condition is a necessary and sufficient condition for the problem and we can find the solution if it exists.

If MFCQ does not satisfy for the problem, we need to check a weaker constraint qualification. Below is a couple of constraint qualifications which are easy to verify.

Quasi-normality Constraint Qualification

Suppose that (q^*, t^*) is a feasible vector of problem (PA_d) . We say that (q^*, t^*) is quasinormal if

$$\sum_{i=1}^n \lambda_i \begin{pmatrix} \mathbf{0} \\ -\frac{\partial h}{\partial q}(\theta_i, q_i) \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \\ 1 \\ \mathbf{0} \end{pmatrix} \begin{matrix} i\text{-th row} \\ \\ \\ \\ i\text{-th row} \end{matrix} + \sum_{i=1}^n \sum_{j=1, j \neq i}^n \lambda_{ij} \begin{pmatrix} \mathbf{0} \\ \frac{\partial h}{\partial q}(\theta_i, q_j) \\ -\frac{\partial h}{\partial q}(\theta_i, q_i) \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \\ -1 \\ 1 \\ \mathbf{0} \end{pmatrix} \begin{matrix} j\text{-th row} \\ i\text{-th row} \\ \\ \\ \\ j\text{-th row} \\ i\text{-th row} \end{matrix} = 0.$$

implies either all $\lambda_i = 0$ or $\{i : \lambda_i \neq 0\} = \emptyset$ and there is no sequence $q^k \rightarrow q^*, t^k \rightarrow t^*$

such that

$$h(\theta_i, q_i^k) - t_i^k > h(\theta_i, q_j^k) - t_j^k \quad \text{if } \lambda_{ij} \neq 0,$$

$$h(\theta_i, q_i^k) - t_i^k > 0 \quad \text{if } \lambda_i \neq 0.$$

In particular, if at least one of the λ 's associated with inequality inactive constraints becomes nonzero.

In the following constraint qualifications, we define the following index sets:

$$I_g = \{i = 1, \dots, n : h(\theta_i, q_i^*) - t_i^* = 0\},$$

$$I_p = \{i = 1, \dots, n : h(\theta_i, q_i^*) - t_i^* = h(\theta_i, q_j^*) - t_j^*\}$$

Arrow-Hurwicz-Uzawa Constraint Qualification

A feasible point (q^*, t^*) satisfies Arrow-Hurwicz-Uzawa Constraint Qualification if there exists a vector d such that for the pseudoconcave constraints at (q^*, t^*) , we have

$$\left\{ \begin{array}{l} \left(\begin{array}{l} \mathbf{0} \\ -\frac{\partial h}{\partial q}(\theta_i, q_i^*) \\ \frac{\partial h}{\partial q}(\theta_i, q_j^*) \\ \mathbf{0} \\ - - - - - \\ \mathbf{0} \\ 1 \\ -1 \\ \mathbf{0} \end{array} \right) \begin{array}{l} i\text{-th row} \\ j\text{-th row} \\ \\ \\ \\ i\text{-th row} \\ j\text{-th row} \\ \\ \end{array} \\ \left(\begin{array}{l} \mathbf{0} \\ -\frac{\partial h}{\partial q}(\theta_i, q_i^*) \\ \mathbf{0} \\ - - - - - \\ \mathbf{0} \\ 1 \\ \mathbf{0} \end{array} \right) \begin{array}{l} i\text{-th row} \\ \\ \\ \\ i\text{-th row} \\ \\ \end{array} \end{array} \right. \quad \begin{array}{l} d < 0, \quad i \in I_p \\ \\ \\ \\ \\ d < 0, \quad i \in I_g \end{array}$$

and for the non-pseudoconcave constraints at (q^*, t^*) , we have

$$\left\{ \begin{array}{l} \left(\begin{array}{c} \mathbf{0} \\ -\frac{\partial h}{\partial q}(\theta_i, q_i^*) \\ \frac{\partial h}{\partial q}(\theta_i, q_j^*) \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \\ 1 \\ -1 \\ \mathbf{0} \end{array} \right) \begin{array}{l} \mathbf{0} \\ i\text{-th row} \\ j\text{-th row} \\ \\ \\ \\ i\text{-th row} \\ j\text{-th row} \\ \\ \end{array} \\ \\ \left(\begin{array}{c} \mathbf{0} \\ -\frac{\partial h}{\partial q}(\theta_i, q_i^*) \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \\ 1 \\ \mathbf{0} \end{array} \right) \begin{array}{l} i\text{-th row} \\ \\ \\ \\ \\ i\text{-th row} \\ \\ \end{array} \end{array} \right. \quad \begin{array}{l} \\ \\ \\ d \leq 0, \quad i \in I_p \\ \\ \\ \\ \\ d \leq 0, \quad i \in I_g \end{array}$$

Finally, we will see that the Abadie Constraint Qualification is one of the weakest constraint qualifications. However, it is not easily verifiable.

Abadie Constraint Qualification

Let (q^*, t^*) be a feasible point of our optimization problem. We say the Abadie Constraint Qualification holds at (q^*, t^*) if the linearized cone is equal to the tangent cone at (q^*, t^*) . In terms of notation we have,

$$T(q^*, t^*) = L(q^*, t^*)$$

where

$$T(q^*, t^*) := \{d \in R^n \mid \exists (q^k, t^k) \subseteq X, p^k \rightarrow 0 : (q^k, t^k) \rightarrow (q^*, t^*) \text{ and } \frac{(q^k, t^k) - (q^*, t^*)}{p^k} \rightarrow d\}$$

and X denotes the feasible set and,

$$L(q^*, t^*) = \left\{ d \in R^n \mid \left(\begin{array}{c} \mathbf{0} \\ -\frac{\partial h}{\partial q}(\theta_i, q_i^*) \\ \frac{\partial h}{\partial q}(\theta_i, q_j^*) \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \\ 1 \\ -1 \\ \mathbf{0} \\ \left(\begin{array}{c} \mathbf{0} \\ -\frac{\partial h}{\partial q}(\theta_i, q_i^*) \\ \mathbf{0} \\ \text{---} \\ \mathbf{0} \\ 1 \\ \mathbf{0} \end{array} \right) \end{array} \right) \left. \begin{array}{l} i\text{-th row} \\ j\text{-th row} \\ \\ \\ \\ i\text{-th row} \\ j\text{-th row} \\ \\ i\text{-th row} \\ \\ \\ i\text{-th row} \end{array} \right\} \left. \begin{array}{l} \\ \\ \\ d \leq 0, i \in I_p, \\ \\ \\ \\ d \leq 0, i \in I_g \end{array} \right\}.$$

Chapter 6

Examples and Economic Interpretation

In this section, we are going to see some numerical examples and applications of the problem discussed earlier. We also developed the algorithms of the Ellipsoid Method introduced in Chapter 2 to solve the principal-agent problem when we have any finite number of customers. We can also use CVX software to compare our results.

6.1 Example 1

Assume that there is a restaurant on an island. Each dish in this restaurant contains two separate pieces: the main food and desert. The cost of preparing the main food with quality q_i is q_i^4 which contains the cost of using good products and professional cook. If the cost of preparing desert with quality q_i be q_i^3 then the total cost for the owner of the restaurant would be

$$C(q_i) = q_i^4 + q_i^3 + 15$$

where 15 indicates the service cost which is constant for all the customers in this restaurant. It is obvious that $C(q_i)$ is convex, since the second derivative of the function is always positive (considering that each q_i is positive). Suppose the satisfaction function of the customer is

$$h(\theta, q) = \theta q.$$

Note that the satisfaction of the customer will increase if the quality increases. If the owner of the restaurant knows how many customers with different tastes are living on this island and what their probabilities are then, he can maximize his profit by using the Ellipsoid algorithm defined in Chapter 2. In other words, he would be able to optimize the quality of each dish in the menu in order to maximize his profit. To see the problem with more details we assume there are just 5 customers with different tastes living in this island. Now, we are going to try different quality distributions by changing the probability of the customers.

a) First, suppose $f = [0.7, 0.2, 0.05, 0.0499, 0.0001]$ then by using the Ellipsoid algorithm after $k = 481$ iterations and 0.124428 seconds we can find

$$Q = [0.3589, 0.5393, 0.6070, 0.8024, 0.8765]^T$$

and the optimal value = 14.6303 which is the optimal profit of the owner of the restaurant. As we expect for each $i, q_{i+1} \geq q_i$, otherwise the person with taste θ_{i+1} may order a cheaper dish. It is interesting to see that although the probability of the θ_5 (the sophisticated customer) is 0.0001 but the amount that we found for q_5 is larger than the others. This means that although the probability of encountering such a sophisticated customer is too small, but if he or she come to the restaurant, the owner has to give him or her the best quality that he can, because he can earn lots of money out of it. However, for the other types of customers it is just enough to give them sufficient quality to make sure that they choose a dish that was targeted for them. These cases were shown by incentive compatible constraints in our model. We can achieve the same result by using CVX after 10 iterations and taking 0.438250 seconds.

b) If we increase the probability of having customer with taste θ_5 from 0.0001 to 0.8 then we expect the value for q_5 to be reduced. Suppose $f = [0.05, 0.05, 0.05, 0.05, 0.8]$. By using both CVX and the Ellipsoid algorithms the result is

$$Q = [0.0000, 0.0000, 0.0000, 0.0000, 0.8766]^T$$

and the optimal profit is 12.5049. As we can see it suggests the owner to produce the food with quality 0 for the customers with tastes 1 to 4. Now, this question may come to mind that why the quality for all customers with tastes 1 to 4 is zero? How we can overcome the incentive compatible constraints with this result? The answer is that in

this case, the high frequency of high taste customers compensate for the improbable loss of other types of customers. We should not forget that in this example any small increase or decrease in quality can make a huge difference in our cost function.

c) By letting $f = [0.6, 0.05, 0.1, 0.15, 0.1]$, we get

$$Q = [0.0000, 0.0000, 0.3388, 0.7463, 0.8766]^T.$$

This means although the probability of taste one is larger than the probability of taste five, but since the owner will receive more money from the customer with taste five it is more profitable for him to give him an excellent quality. In this case, the optimal value for the owner is 14.4121, which is larger than the previous value that we calculated (in b) since in this case the decreased probability of the coming of sophisticated customers make two other types of customers profitable, and the generated new profit is larger than the loss from the high taste customers.

The methods that we discussed in this section are particularly useful when our cost function is complicated and we cannot solve the problem by using simple methods of optimization. In the following we are going to discuss a different cost function in a strange example.

6.2 Example 2

In this example, we want to replace all we had in the restaurant's example by a group of thieves who want to rob a bank. They know how many cameras are available in the bank but they do not know the place of each camera. So, in this example θ is the number of cameras that are installed in different places of a bank for protection and q is the quality of the thieves' robbing project. Depending on the number of cameras, their project plan for robbing changes, thus their cost function changes. Suppose

$$h(\theta, q) = \theta q$$

denotes the discomfort level of the stealing project. This means when q increases then the discomfort of the robbing project also increases. Their cost function contains the number and skill-level of the people that they need and the kind of equipment for their operation. We assume that by adding each one camera in the bank their cost

function almost doubles. Suppose,

$$C(q_i) = 2^{q_i} - 3q_i + 25$$

where 2^q is the cost of skilled thieves and equipment, $3q$ is the profit that they will gain by stealing the expensive lenses of those cameras (besides robbing the cash, they also steal the lenses of security cameras), and 25 is the fixed cost of their transportation. In this example, the cost function is convex since $C''(q_i) = 2^{q_i}(\ln 2)^2 \geq 0$. Suppose t_i is the money that they earn from robbing the bank. So, the thieves want to maximize their profit which is equal to the money that they earn from the bank minus their cost. Suppose the probability of encountering a bank with one camera is 0.5, with two camera is 0.2 and with three camera is 0.4. This means $\alpha = [0.4, 0.2, 0.4]$ then by going through the algorithm of Ellipsoid we can find the maximum profit of the thieves, the best value for the quality of the thieves' planning is

$$Q = [1.8507, 2.1137, 3.1137]^T.$$

This gives us an optimal value of 20.1789. It seems logical that the quality of their operation should be the best when they are dealing with three cameras.

6.3 Example 3

This example is similar to Example 1 with a major difference which is here we no longer have the Spence Mirrlees Condition for the $h(\theta, q)$. Again there is a restaurant on an island. Each dish in this restaurant contains two separate pieces: the main food and desert. The cost of preparing the main food with quality q_i is q_i^4 which contains the cost of ingredients and cooking. If the cost of preparing desert with quality q_i be $5q_i^2$ then the total cost for the owner of the restaurant would be

$$C(q_i) = q_i^4 + 5q_i^2 + 25$$

where 25 indicates the service cost which is constant for all the customers in this restaurant. It is obvious that $C(q_i)$ is concave, since the second derivative of the function is always positive (considering that each q_i is positive). Assume that we have three types of cutomers with tastes $\theta_1 = 1, \theta_2 = 3$ and $\theta_3 = 6$. The satisfaction

function $h(\theta, q)$ is defined as follows:

$$h(\theta, q) = \left(\theta - \frac{3}{2}\right)^2 q + 200.$$

In this function the satisfaction of the customer will increase if the quality increases. Note that the SMC will fail at the point $\theta = \frac{3}{2}$.

If the owner of the restaurant knows how many customers with different tastes live on this island and what their probabilities are then, he can maximize his profit by using Fmincon algorithm. This is because the objective function is concave and the constraints are linear. In other words, he would be able to optimize the quality of each dish in the menu in order to maximize his profit. Now, we are going to try different quality distributions by changing the probability of the customers.

a) First, suppose $f = [0.1, 0.8, 0.1]$ then by using Fmincon we can find

$$Q = [0.0000, 0.0250, 1.5463]^T,$$

$$T = [1.5962, 1.6024, 32.4085]^T$$

and the optimal value = 22.0872 which is the optimal profit of the owner of the restaurant. Same as the Example 1 for each $i, q_{i+1} \geq q_i$, otherwise the person with taste θ_{i+1} may order a cheaper dish. It is interesting to see that although the probability of the θ_3 (the sophisticated customer) is 0.1 but the amount that we found for q_3 is larger than the others. This means that although the probability of encountering such a sophisticated customer is too small, but if he or she comes to the restaurant, the owner has to give him or her the best quality that he can, because he can earn lots of money out of it. However, for the other types of customers it is just enough to give them sufficient quality to make sure that they choose a dish that was targeted for them. These cases were shown by incentive compatible constraints in our model.

b) By letting $f = [0.45, 0.05, 0.5]$, we get

$$Q = [6.0160 \times 10^{-20}, 3.1510 \times 10^{-20}, 0.4903]$$

$$T = [0.4903, 0.4903, 2.1452].$$

As we expect $q_2 < q_1$. This show our result in Chapter 5 that when the Spence Mirrlees Condition fails, there is no guarantee for the satisfaction of the monotonicity condition.

Chapter 7

Conclusion

In this thesis, we study the principal-agent problem with adverse selection from the mathematical point of view.

For the model we consider, to our knowledge there were no existence results available although the existence results for another model was established by Carlier in [9]. Since Carlier's existence result can not be directly applied to obtain the existence result for our model, we use Carlier's technique to derive an existence of optimal solutions for our model.

For the principal-agent problem with adverse selection it is known that when the Spence Mirrlees condition holds an optimal solution can be explicitly found. We survey these results with slight improvements so that the function $h(\theta, q)$ is not linear in θ .

For principal-agent problem when Spence Mirrlees condition does not hold, we study constraint qualifications and we found a sufficient condition for MFCQ.

Finally we give economic interpretation on some numerical examples.

Bibliography

- [1] J.M. Abadie. *On the Kuhn-Tucker Theorem*, in: *J. Abadie (Ed.), Nonlinear Programming*. Wiley, New York, 1967.
- [2] A. Antoniou and W.S. Lu. *Practical Optimization: Algorithms and Engineering Applications*. Springer Science+Business Media, New York, 2007.
- [3] A. Araujo and H. Moreira. *Adverse selection problems without the Spence-Mirrlees condition*. *Journal of Economic Theory*, 145: 1113-1141, 2010.
- [4] K. J. Arrow, L. Hurwiz, and H. Uzawa (Eds.). *Studies in Linear and Nonlinear Programming*. Stanford University Press, Stanford, CA, 1958.
- [5] D. P. Bertsekas and A. E. Ozoaglar. *Pseudonormality and a Lagrange multiplier theory for constrained optimization*. *Journal of Optimization Theory and Applications*, 114: 287343, 2002.
- [6] P. Bolton and M. Dewatripont. *Contract Theory*. The MIT Press, London, England, 2005.
- [7] J. M. Borwein and A. S. Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples*. Springer Science+Business Media, Inc., New York, 2006.
- [8] H. Brzis. *Analyse Fonctionnelle Theorie et Applications*. Masson, Paris, 1984.
- [9] G. Carlier. *A general existence result for the principal-agent problem with adverse selection*. *Journal of Mathematical Economics*, 35: 129-150, 2000.
- [10] G. Carlier. *Calculus of variations for non-decreasing functions*. mimeo, 2000a.
- [11] G. Carlier. *Calculus of variations with convexity constraint*. mimeo, 2000b.

- [12] F. H. Clarke. *Optimization and Non-smooth Analysis*. Wiley-Interscience, New York, 1983.
- [13] R. Guesnerie and J. J. Laffont. *A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm*. *Journal of Public Economics*, 25: 329-369, 1984.
- [14] F. John and J. Moser. ed. *Fritz John: Collected Papers*. Birkhuser, Boston, 1948.
- [15] M. I. Kamien and N.L. Schwartz. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*. North Holland, New York, 1981.
- [16] H.W. Kuhn and A. W. Tucker. *Nonlinear programming, in: J. Neyman (Ed.)*. *Proceedings of the second Berkeley Symposium Mathematical Statistics and Probability*, University of California Press, Berkeley, CA, 1951.
- [17] J. J. Laffont and D. Martimort. *The Theory of Incentives: the Principal-Agent Model*. Princeton University Press, Princeton, New Jersey, 1947.
- [18] V. Levin. *Reduced cost function and their applications*. *Journal of Mathematical Economics*, 28: 155-186, 1997.
- [19] H.D. Sherali M. S. Bazaraa and C.M. Shetty. *Nonlinear Programming: Theory and Algorithms*. Wiley Interscience, New Jersey, 2006.
- [20] E. Maskin and J. Riley. *Monopoly with incomplete information*. *The Rand Journal of Economics*, 15: 171-196, 1984.
- [21] J. Mirrlees. *An exploration in the theory of optimum income taxation*. *Review of Economic Studies*, 38: 175-208, 1971.
- [22] M. Mussa and S. Rosen. *Monopoly and product quality*. *Journal of Economic Theory*, 18: 301-317, 1978.
- [23] J. V. Outrata and W. Romisch. *On optimality conditions for some nonsmooth optimization problems over L^p spaces*. *journal of optimization theory and applications*, 126: 411438, 2005.

- [24] L. S. Pontryagin, V.G. Boltayanskii, R.V. Gamkrelidze, and E.F. Mishchenko. *The Mathematical Theory of Optimal Processes*. Wiley, New York, 1962.
- [25] J. C. Rochet and P. Chon. *Ironing, sweeping, and multidimensional screening*. *Econometrica*, 66: 753-826, 1998.
- [26] W. Rudin. *Real and Complex Analysis*. McGraw-Hill book Co., Singapore, 1986.
- [27] B. Salanie. *The Economics of Contracts*. The MIT Press, London, England, 1998.
- [28] P. Schleiter and E. Morgan-Jones. *Citizens, presidents and assemblies: the study of semi-presidentialism beyond Duverger and Linz*. *British Journal of Political Science*, 39: 871-892, 2009.
- [29] M. Spence. *Job market signaling*. *Quarterly Journal of Economics*, 87: 355-74, 1973.
- [30] R. B. Wilson. *Nonlinear Pricing*. Oxford University Press, New York, 1993.