Journal of Mathematical Economics 54 (2014) 7–21

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

The choice of the number of varieties: Justifying simple mechanisms

Adam Chi Leung Wong *

School of International Business Administration, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, 200433, China

HIGHLIGHTS

- We study monopolistic quality differentiation with constrained number of varieties.
- The marginal benefit of adding variety is diminishing.
- The loss from restricting to n varieties is of order no more than $1/n^2$.
- The marginal benefit of adding one more variety from n is of order no more than $1/n^3$.
- Offering only two varieties can lead to more than two-thirds of the second best profit.

ARTICLE INFO

Article history:
Received 16 January 2014
Received in revised form
28 July 2014
Accepted 6 August 2014
Available online 14 August 2014

Keywords:
Mechanism design
Nonlinear pricing
Short menu

ABSTRACT

We study a mechanism designer's trade-off between the complexity level and optimality level of a mechanism. While our techniques apply to a much larger class of mechanism design problems, we focus on the quality differentiation model of Mussa and Rosen (1978), restricting the monopolist to menus with at most a finite number n of varieties. We prove that (i) the marginal benefit of adding one more variety is diminishing in n; (ii) the loss from restricting the number of varieties is of order no more than $1/n^2$; (iii) the marginal benefit of adding one more variety is of order no more than $1/n^3$; and (iv) offering only two varieties can lead to more than two-thirds of the potential profit from the second best offering. Our analysis suggests that the monopolist would probably offer only a small number of varieties in the menu.

1. Introduction

Mechanism design theory has now become a classic and far-reaching branch in economics. It has been used to derive optimal income taxation schemes (Mirrlees, 1971), optimal nonlinear pricing schemes (Maskin and Riley, 1984), and optimal quality differentiation (Mussa and Rosen, 1978), among many others. While these theoretical solutions of optimal mechanisms have been well known, people generally tend to embrace much simpler mechanisms in reality, like an income taxation scheme with a few tax bands and marginal tax rates, a multipart tariff with a small number of "parts", and a quality–price scheme with only a few quality-differentiated varieties. How well can a suboptimal but simpler mechanism perform relative to the fully optimal mechanism? If complicating the mechanism is costly, how should the mechanism designer choose the optimal "complexity level" of the mechanism?

We will consider the monopolistic quality differentiation framework of Mussa and Rosen (1978). In this framework, a monopolist is uninformed about its customers' preferences over quality (or types), but it can produce and offer a spectrum of quality-differentiated varieties to separate different types of customers. The optimal spectrum involves a continuum of quality-differentiated varieties, tailor-made for each consumer type. However, if the monopolist decides, for practical considerations, to offer at most a finite number n of varieties only, it would design a discrete offering (i.e., a menu of a finite number of quality–price choices) in order to maximize its profit subject to the maximum number of varieties n. There would then be a "constrained profit" Π_n for each n. Our main task is to characterize the properties of the constrained profit sequence $\{\Pi_n\}_{n=0}^{\infty}$. We also consider the setting with a fixed cost of developing each variety, which endogenizes the number of varieties. Our analysis suggests that the monopolist would probably offer only a small number of varieties in the menu.[1]

While we restrict our study to the monopolistic quality differentiation problem for a concrete presentation, we emphasize that the techniques developed in this paper can be applied to other

* Tel.: +86 21 65906560; fax: +86 21 65907458.
E-mail address: wongchileung@gmail.com.

[1] This is true for a social planner's welfare maximization problem as well.

http://dx.doi.org/10.1016/j.jmateco.2014.08.002
0304-4068/

mechanism design (or principal–agent) problems, where there is one principal and one agent, and the agent has one-dimensional private information.[2] The number n should be thought of as a measure of a mechanism's complexity level, which could be interpreted in different ways for different kinds of problems. For example, n could be reinterpreted as the number of two-part tariffs offered by a seller to consumers in the context of nonlinear pricing, or the number of possible messages that can be sent from the agent to the principal in a principal–agent model with limited communication.

The "constrained program", that is, seeking the optimal discrete offering subject to a maximum number of varieties, has no explicit solution except for special cases (e.g., Example 1). However, we can show a number of qualitative features of an optimal discrete offering and the constrained profit sequence. First, it is not hard to show that an optimal discrete offering (given any n) must be a step function fluctuated around the optimal continuous offering (or second best offering), and Π_n monotonically converges to the fully optimal profit (or second best profit) Π_∞ as n becomes large.

If adding every extra variety to the offered menu is costly and the marginal benefit of adding one more variety $\Pi_{n+1} - \Pi_n$ is diminishing, the monopolist should optimally choose the number of varieties that approximately equalizes the marginal benefit and marginal cost of adding one more variety. Our first main result is that the marginal benefit $\Pi_{n+1} - \Pi_n$ is really diminishing in n. To the best of my knowledge, this is the first diminishing marginal benefit result in any similar context. Intuitively, as the number of varieties offered gets larger, the space for improving profit by adding one more variety becomes smaller and hence the effectiveness of the extra variety becomes less. However, this "diminishing marginal benefit property" is far from trivial, because adding one more variety would cause an optimal adjustment of all previously offered varieties. Although $\Pi_{n+1} - \Pi_n$ must ultimately diminish, it is rather surprising that the property holds for every n in a general setup. This diminishing marginal benefit property is not only interesting on its own, but also crucial to proving many of our other results. Moreover, the proof of the diminishing marginal benefit property, which involves comparing different constrained profits and suboptimal profits in graphs, is interesting on its own.

Our second main result is what we call the "quadratic rate result", that is, the "uncaptured profit" $\Pi_\infty - \Pi_n$ is of order no more than $1/n^2$. The intuition is that the slope of the virtual surplus with respect to quality is flat at the ideal second best quality. Hence, the loss from deviating from the second best quality due to discrete offering is of second or higher order, but not of first order. In a discrete offering with n varieties, although different types of consumers have to be pooled and served with a single quality, the distance between the quality serving a particular type and the second best quality for that type is approximately proportional to $1/n$. A Taylor expansion argument shows that the uncaptured profit is of order no more than $1/n^2$. Moreover, this convergence rate can be attained by a simple offering rule involving a uniformly distributed set of (suboptimal) varieties. Furthermore, the bound we provide for $\Pi_\infty - \Pi_n$ is tight.

Our third main result is what we call the "cubic rate result", that is, the marginal benefit of adding one more variety $\Pi_{n+1} - \Pi_n$ is of order no more than $1/n^3$. As a matter of mathematical fact, the aforementioned quadratic rate result alone does not imply the cubic rate result.[3] The latter is an implication of the quadratic rate result *and* the diminishing marginal benefit property. Intuitively, the diminishing marginal benefit property ensures that the uncaptured profit would be captured to a large extent by the earlier extra varieties. Hence, the convergence rate of the marginal benefit $\Pi_{n+1} - \Pi_n$ would be faster than that of the uncaptured profit $\Pi_\infty - \Pi_n$. As yet another implication of the cubic rate result and the diminishing marginal benefit property, the existence of a moderate marginal cost k of developing extra varieties (cost of complexity) can plausibly justify the optimal number of varieties (optimal complexity level) being quite small. More precisely, the optimal number of varieties is of order no more than $1/k^{1/3}$.

Our fourth main result is what we call the "two-thirds result". It says that the monopolist can earn more than two-thirds of the unconstrained profit by offering only two varieties, that is, $\Pi_2 > 2\Pi_\infty/3$.[4] The literature has results of this kind derived in the context of procurement and regulation, and matching (see below), but to the best of my knowledge, this is the first result of this kind in any nonlinear pricing-type context. Most, if not all, of this kind of results in the literature need to assume specific functional forms. The same applies to ours. For our two-thirds result to hold, we need to assume that the consumers' utility is linear and the production cost quadratic in quality (the so-called linear-quadratic model, an extensively studied one in the literature), and the distribution of virtual types satisfies a regularity condition. Once again, the diminishing marginal benefit property plays a major role in the proof.

While the above results are obtained in a continuous type model, we also consider a discrete type version of the same model and adapt our results there.

The most related paper in the literature is the one concurrently written by Bergemann et al. (2011). It proves the quadratic rate result in the context of nonlinear pricing. However, its analysis, which applies the quantization theory, works only for the linear-quadratic model.[5] On the other hand, Wilson (1989), Wilson (1993), and Blumrosen et al. (2007) also provide quadratic rate results in contexts that are mathematically different from ours, namely, the efficient rationing of services, Ramsey pricing, and auctions with bounded communication, respectively.[6] However, they do not analyze the marginal benefit of complicating the mechanism (which is crucial to the optimal choice of complexity level), nor the performance of a simple mechanism relative to that of the second best or first best.

In the context of procurement contracting, Rogerson (2003) considers the "Fixed Price Cost Reimbursement (FPCR) menus", that is, two-item menus in which one item is a cost-reimbursement contract and the other a fixed-price contract, of which the principal allows the agent to pick one. He shows that, if the agent's utility is quadratic and the agent's type is uniformly distributed, then "the optimal FPCR menu always captures at least three-quarters of the gain that the optimal complex menu achieves". Chu and Sappington (2007) allow a more general family of power distributions and show that a menu of two options, namely, a cost-reimbursement contract and a linear cost sharing contract, can always secure at least 73% of the gain. McAfee (2002) shows that in the context of two-sided matching, if the matching surplus takes

2 For the optimal solution of this kind of problems, see Fudenberg and Tirole (1991, Chapter 7), or Guesnerie and Laffont (1984).

3 However, the converse is true.

4 Of course, this, together with the diminishing marginal benefit property, implies $\Pi_1 > \Pi_\infty/3$.

5 Bergemann et al. (2012) consider a linear-quadratic nonlinear pricing model with multidimensional agents' types and choices, but restrict their attention to social welfare maximization problems and hence effectively assume away the incentive compatibility constraints, the central difficulty of multidimensional problems. They show that the convergence rate of welfare loss is slower than quadratic.

6 In Blumrosen et al. (2007), the characterization of optimal auctions under communication restriction is only for cases with either two bidders or two possible bids. Kos (2012) generalizes the analysis by allowing for a finite number of bidders and possible bids. The quadratic rate result in Blumrosen et al. (2007) is, however, general.