The Choice of the Number of Varieties: Justifying Simple Mechanisms

Adam Chi Leung Wong*,†

August 7, 2014

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.

Keywords: Mechanism design, Nonlinear pricing, Short menu

JEL Classification Numbers: D42, D82, L15.

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

*School of International Business Administration, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, China 200433. Phone: +8621 65906560. Fax: +8621 65907458. Email: wongchileung@gmail.com.

†The author is grateful to three anonymous referees and the co-editor for their comments and suggestions.

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 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

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

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

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.

[3] However, the converse is true.

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, Shen, Xu, and Yeh (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, Nisan, and Segal (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.

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

[5] Bergemann, Shen, Xu, and Yeh (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, Nisan, and Segal (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, Nisan, and Segal (2007) is, however, general.

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 the multiplicative form in agents' types and the distributions of types satisfy some regularity conditions, a social planner can divide the agents of each side into only two classes such that the resulting "coarse matching" achieves at least half of the social gain achieved by the fully optimal matching. These results can be regarded as analogous to our two-thirds result.

My companion paper Wong (2012) considers nonlinear pricing and compares the maximum profits of different forms of pricing schemes (namely, bundling, incremental discounts, and all-units discounts) with any common complexity level. It complements this paper in that it sheds some light on how to choose among different forms of simple mechanisms.

Miravete (2007) uses a large sample of independent cellular telephone markets to structurally estimate a monopolistic nonlinear pricing model. His estimates suggest that "firms should only offer few tariff options if the product development costs of designing them are non-negligible." His findings empirically support our theoretical results, and our results provide the rationale for his.

The rest of the paper is organized as follows. Section 2 sets up the environment and provides the standard solution in the literature. Section 3 characterizes the optimal discrete offering and provides a preliminary analysis of the constrained profit conditional on a maximum number of varieties. Section 4 presents the main results on properties of the constrained profit sequence. Section 5 considers a discrete type version of the same model. The proofs we do not provide in the text are in the Appendix.

2 The Model

2.1 Environment

Consider the monopolistic quality differentiation environment of Mussa and Rosen (1978). A commodity can be produced by a monopolist in a spectrum of varieties. The hedonic attributes of the varieties are characterized by a one-dimensional nonnegative quality index q. The consumers have unit demand for the commodity; that is, every consumer chooses to buy either 0 or 1 unit. A consumer who decides to buy the commodity must pick one of the offered varieties. The higher the quality index of a variety, the higher is the willingness-to-pay of a consumer for the variety.

Consumers are heterogeneous in their types, indexed by t. The utility of a type t consumer who buys a variety with quality q and pays the price p is represented as $t \cdot v(q) - p$. If the consumer does not buy a unit, her utility is zero. The function $v : \mathbb{R}_+ \to \mathbb{R}_+$ is twice continuously differentiable with the properties

$$v(0) = 0, \quad v'(q) > 0, \quad v''(q) \leq 0 \; \forall q \geq 0. \tag{1}$$

The distribution of t is characterized by a cumulative distribution function F, whose support is a compact interval $[\underline{t}, \bar{t}]$. We assume that F admits a positive density function f on the support. The type of every consumer is the consumer's private information. The monopolist only knows the prior distribution F.

The unit cost of producing a variety with quality q is denoted as $c(q)$. The cost function $c : \mathbb{R}_+ \to \mathbb{R}_+$ is twice continuously differentiable with the properties

$$c(0) = 0, \quad c'(q) > 0 \; \forall q > 0, \quad c''(q) > 0 \; \forall q \geq 0. \tag{2}$$

We also assume that

$$\frac{c'(0)}{v'(0)} < \bar{t} < \lim_{q \to \infty} \frac{c'(q)}{v'(q)} \tag{3}$$

to guarantee that the market is nontrivial and the optimal quality for each type is bounded. Note that for each type t, the efficient allocation of quality is either 0 if $tv'(0) \leq c'(0)$, or the $q > 0$ that solves $tv'(q) = c'(q)$ if $tv'(0) > c'(0)$. (Note that not consuming a unit can be identified with consuming a variety with quality 0.)

2.2 Unconstrained solution

If the monopolist can costlessly establish as many varieties as it wants, the problem it faces is the standard problem studied by Mussa and Rosen (1978). That is, the monopolist solves

$$\max_{q(\cdot) \geq 0, p(\cdot)} \int_{\underline{t}}^{\bar{t}} [p(t) - c(q(t))] \, dF(t) \tag{4}$$

subject to

$$tv(q(t)) - p(t) \geq 0 \quad \forall t \in [\underline{t}, \bar{t}] \tag{5}$$

$$tv(q(t)) - p(t) \geq tv(q(t')) - p(t') \quad \forall t, t' \in [\underline{t}, \bar{t}]. \tag{6}$$

The functions $p(t)$ and $q(t)$ specify the monopolist's choice of price and quality for consumers with type t. The objective function in (4) is the (per consumer) profit. Constraint (5) is the individual rationality (IR) constraint, which exists because every consumer has the outside option of buying nothing, paying nothing and getting the reservation utility zero. Constraint (6) is the incentive compatibility (IC) constraint, which arises because the consumers' types are private information.

Since the problem above has no constraint on the number of varieties to be offered (as opposed to the problems in later sections) and the standard solution in general involves a continuum of varieties, we call it the *unconstrained problem* or *unconstrained program*; we call the maximized value the *unconstrained profit*, denoted as Π_∞.

Adopting the standard technique to solve this kind of problem[7], we reduce the unconstrained program to

$$\max_{q(\cdot)\geq 0} \int_{\underline{t}}^{\bar{t}} [K(t)v(q(t)) - c(q(t))]\, dF(t) \tag{7}$$

subject to the monotonicity constraint

$$q(t) \text{ is nondecreasing in } t, \tag{8}$$

where

$$K(t) \equiv t - \frac{1 - F(t)}{f(t)}$$

is the "virtual type function." It is common in the literature to assume that K is nondecreasing in order to make the monotonicity constraint (8) non-binding. However, we *do not* assume this until Subsection 4.3, where we make a stronger regularity assumption. Then "bunching" (i.e., different types served with the same quality) may occur and constraint (8) cannot be dropped. The modern method to solve problem (4) involves applying optimal control theory, regarding $q(t)$ as a state variable. While this method works for our unconstrained program, it does not work when we analyze our constrained program in the next section, because $q(t)$ is discontinuous there. In order to address this problem, we apply the classic ironing procedure due to Myerson (1981).

[7] See for example Fudenberg and Tirole (1991), Chapter 7.

Define

$$R(y) \equiv \int_0^y K(F^{-1}(\tilde{y}))d\tilde{y} \quad \forall y \in [0,1],$$

$$\hat{R}(y) \equiv \operatorname{conv} R(y) \quad \forall y \in [0,1],$$

where $\operatorname{conv} R$ denotes the convex hull of R (i.e., the greatest convex function below R; see Rockafellar (1970), p.36). Then define the "ironed virtual type function" $J(t)$ as the unique left-continuous function on $[\underline{t}, \bar{t}]$ such that

$$\hat{R}(y) \equiv \int_0^y J(F^{-1}(\tilde{y}))d\tilde{y} \quad \forall y \in [0,1].$$

It is not hard to see that this J function is fully characterized by the following four properties:

$$J \text{ is nondecreasing and left-continuous,} \tag{9}$$

$$\int_{\underline{t}}^{t} J(\tilde{t})dF(\tilde{t}) \leq \int_{\underline{t}}^{t} K(\tilde{t})dF(\tilde{t}) \quad \forall t \in [\underline{t}, \bar{t}], \tag{10}$$

$$(10) \text{ holds as equality at } t = \bar{t}, \tag{11}$$

$$\text{whenever } (10) \text{ is strict at } t,\ J \text{ is constant in some neighborhood of } t. \tag{12}$$

Then it can be shown that the profit-maximizing quality $q_\infty(t)$ for type t is such that

$$J(t)v'(q_\infty(t)) = c'(q_\infty(t)) \tag{13}$$

whenever type t is served, i.e., whenever $J(t)v'(0) > c'(0)$.

Proposition 1 (Mussa and Rosen, 1978) *The unconstrained profit* Π_∞ *can be written as*

$$\Pi_\infty = \max_{q(\cdot)\geq 0} \int_{\underline{t}}^{\bar{t}} [J(t)v(q(t)) - c(q(t))]\, dF(t) = \int_{\underline{t}}^{\bar{t}} [J(t)v(q_\infty(t)) - c(q_\infty(t))]\, dF(t) \tag{14}$$

where the optimal continuous offering is defined by $q_\infty(t) = 0$ *for* $\underline{t} \leq t < t_*$ *and by* (13) *for* $t_* \leq t \leq \bar{t}$*, where the lowest served type* t_* *is defined by*

$$t_* \equiv \inf \{t \in [\underline{t}, \bar{t}] : J(t)v'(0) > c'(0)\}.$$

Proof. Here we reprove this result by using Myerson's (1981) technique so that the proof can be adapted to the constrained program later.

The objective function in (7) can be written as

$$\int_{\underline{t}}^{\bar{t}} [J(t)v(q(t)) - c(q(t))]\, dF(t) + \int_{\underline{t}}^{\bar{t}} [K(t) - J(t)]\, v(q(t))dF(t).$$

The second term can be rewritten through integration by parts as

$$\begin{aligned} & \int_{\underline{t}}^{\bar{t}} v(q(t))d\left[R(F(t)) - \hat{R}(F(t))\right] \\ = \ & (R(1) - \hat{R}(1))v(q(\bar{t})) - (R(0) - \hat{R}(0))v(q(\underline{t})) - \int_{\underline{t}}^{\bar{t}} \left[R(F(t)) - \hat{R}(F(t))\right] dv(q(t)). \end{aligned}$$

Since $\hat{R}$ is the convex hull of R and R is continuous, we have $R(1) = \hat{R}(1)$ and $R(0) = \hat{R}(0)$. Therefore, the objective function in (7) can be rewritten as

$$\int_{\underline{t}}^{\bar{t}} [J(t)v(q(t)) - c(q(t))]\, dF(t) - \int_{\underline{t}}^{\bar{t}} \left[R(F(t)) - \hat{R}(F(t))\right] dv(q(t)). \tag{15}$$

Now, clearly $q_\infty(\cdot)$ is nondecreasing and it maximizes the first term. It remains to show that the second term is also maximized over nondecreasing $q(\cdot)$ at $q_\infty(\cdot)$ and that the resulting maximum value is zero. Indeed, the second term is nonpositive for any nondecreasing $q(\cdot)$, because $\hat{R}$ is the convex hull of R so that $\hat{R}(\cdot) \leq R(\cdot)$. Besides, the second term is zero at $q_\infty(\cdot)$. This is again because $\hat{R}$ is the convex hull of R, so that if $\hat{R}(F(t)) < R(F(t))$ then $\hat{R}'(F(\cdot))$ is constant in some neighborhood of t and hence so are $J(\cdot)$ and $q_\infty(\cdot)$. ■

For notational convenience, we assume that $J(\underline{t})v'(0) < c'(0)$ from now on, meaning that some very low types of consumers will not be served (i.e., $t_* > \underline{t}$).

3 The Constrained Program

Now, consider the situation where the monopolist restricts itself to offer at most n varieties. The IR, IC, and nonnegativity constraints still remain. Therefore, the monopolist solves the following *constrained problem* or *constrained program*:

$$\Pi_n = \max_{q(\cdot)\geq 0} \int_{\underline{t}}^{\bar{t}} [K(t)v(q(t)) - c(q(t))]\, dF(t) \tag{16}$$

subject to the monotonicity constraint (8) and the constraint that

$$q(\cdot) \text{ takes at most } n \text{ values except zero.} \tag{17}$$

This constrained program differs from the unconstrained program only in that it has the extra constraint (17). We call Π_n the *constrained profit* and $\Pi_\infty - \Pi_n$ the *uncaptured profit* given the number of varieties n.

3.1 Optimal discrete offering

Now, let us analyze the optimal discrete offering in the constrained program (16), given the number of varieties n. First, from the two constraints in the constrained program, the positive part of $q(\cdot)$ must be a nondecreasing n-step function. Let q_i $(i = 1, \ldots, n)$ be the quality serving the types in $[t_i, t_{i+1}]$ (with the convention that $t_{n+1} = \bar{t}$). Second, we will see that, as in the unconstrained program, the monotonicity constraint in the constrained program can be dropped if we replace the virtual type function K by its ironed counterpart J. That is, in the constrained program, the monotonicity constraint and constraint (17) do not interact. So let us consider the following $2n$-dimensional problem:

$$\max_{\substack{q_1,\ldots,q_n \in \mathbb{R}_+; \\ t_1,\ldots,t_n \in [\underline{t},\bar{t}]}} \left\{ \sum_{i=1}^{n} \int_{t_i}^{t_{i+1}} \left[J(t)v(q_i) - c(q_i) \right] dF(t) \text{ s.t. } t_1 \leq \cdots \leq t_n \right\}. \tag{18}$$

It can also be shown that any solution of (18) must involve $0 < q_1 < \cdots < q_n$ and $\underline{t} < t_1 < \cdots < t_n < \bar{t}$. Therefore, the first-order necessary conditions of (18) with respect to $(q_1, \ldots, q_n; t_1, \ldots, t_n)$ can be written as two first-order difference equations,

$$\int_{t_i}^{t_{i+1}} J(t) dF(t) \cdot v'(q_i) = c'(q_i) \cdot (F(t_{i+1}) - F(t_i)) \qquad \forall i = 1, \ldots, n, \tag{19}$$

$$J(t_i) \cdot (v(q_i) - v(q_{i-1})) = c(q_i) - c(q_{i-1}) \qquad \forall i = 1, \ldots, n, \tag{20}$$

and two boundary conditions,

$$q_0 = 0, \tag{21}$$

$$t_{n+1} = \bar{t}. \tag{22}$$

Proposition 2 *Given* n*, the constrained program* (16) *has a solution. The constrained profit can be written as*

$$\Pi_n = \max_{q(\cdot) \geq 0} \left\{ \int_{\underline{t}}^{\bar{t}} \left[J(t)v(q(t)) - c(q(t)) \right] dF(t) \ \ s.t. \ \ (17) \right\}. \tag{23}$$

Moreover, any $(q_1, \ldots, q_n; t_1, \ldots, t_n)$ *with* $q_1, \ldots, q_n \geq 0$ *and* $\underline{t} \leq t_1 \leq \cdots \leq t_n \leq \bar{t}$ *and*

$$\Pi_n = \sum_{i=1}^{n} \int_{t_i}^{t_{i+1}} [J(t)v(q_i) - c(q_i)]\, dF(t), \tag{24}$$

which we call optimal discrete offering, must involve $0 < q_1 < \cdots < q_n$ *and* $\underline{t} < t_1 < \cdots < t_n < \bar{t}$ *(i.e., constraint* (17) *is binding), and satisfy the difference equations* (19) *and* (20) *and the boundary conditions* (21) *and* (22).

Proof. *Step 1.* Any solution of (23) must involve n distinct nonzero values. To see this, first recall that J is left-continuous. Moreover, $J(\bar{t}) = \bar{t}$ and $J(t) < \bar{t}$ for all $t < \bar{t}$ because K has those properties. Therefore, J is strictly increasing on some neighborhood of $\bar{t}$. Since $J(\bar{t})\, v'(0) > c'(0)$, it follows that $q_\infty(\cdot)$ is strictly increasing and positive on some neighborhood of $\bar{t}$. Now, if a solution $q(\cdot)$ of (23) takes only $n-1$ nonzero values, some $\hat{q}(\cdot)$ function that takes n nonzero values can approximate $q_\infty(\cdot)$ better than $q(\cdot)$ does, i.e., $|q_\infty(t) - \hat{q}(t)| \leq |q_\infty(t) - q(t)|$ for all $t \in [\underline{t}, \bar{t}]$ and $|q_\infty(t) - \hat{q}(t)| < |q_\infty(t) - q(t)|$ for all t in some nonempty open subset of $[\underline{t}, \bar{t}]$. Since the objective function's integrand $J(t)\, v(q) - c(q)$ is single-peaked in q with the unique maximizer $q_\infty(t)$, we see that $\hat{q}(\cdot)$ yields a strictly higher profit than $q(\cdot)$ does, which is a contradiction.

Step 2. Any solution of (18) must involve $0 < q_1 < \cdots < q_n$ and $\underline{t} < t_1 < \cdots < t_n < \bar{t}$. To see this, let $(q_1, \ldots, q_n; t_1, \ldots, t_n)$ be a solution. Note that Step 1 already implies that $q_1, \ldots, q_n$ are distinct and positive, and that $t_1 < \cdots < t_n < \bar{t}$. Moreover, $t_1 > \underline{t}$ is guaranteed by $J(\underline{t})v'(0) < c'(0)$, because if $t_1 = \underline{t}$, the first-order derivative of the objective function with respect to t_1 is

$$-J(\underline{t})v(q_1) + c(q_1) = -q_1 \left[J(\underline{t}) \frac{v(q_1)}{q_1} - \frac{c(q_1)}{q_1} \right] > -q_1 \left[J(\underline{t})v'(0) - c'(0) \right] > 0,$$

which is a contradiction. Furthermore, suppose, by way of contradiction, that $i < i'$ and $q_i > q_{i'}$. Then, clearly we must have

$$J(t)\, v(q_i) - c(q_i) > J(t)\, v(q_{i'}) - c(q_{i'}) \text{ for almost all } t \in (t_i, t_{i+1}),$$

$$J(t')\, v(q_{i'}) - c(q_{i'}) > J(t')\, v(q_i) - c(q_i) \text{ for almost all } t' \in (t_{i'}, t_{i'+1}).$$

But they imply that $J(t) > J(t')$ for some $t < t'$, which contradicts the fact that J is nondecreasing. Thus, we have $q_1 < \cdots < q_n$.

Step 3. Problem (18) has a solution. To see this, note that the objective function in (18) is continuous in $(q_1, \ldots, q_n; t_1, \ldots, t_n)$. Moreover, every q_i can be without loss restricted to be below some large upper bound, because our assumption (3) implies that $\bar{t}v'(q) < c'(q)$ for all large enough q. Then, the constraint set is compact, and from Weierstrass Theorem a maximizer exists. Furthermore, the objective function is differentiable in $(q_1, \ldots, q_n; t_1, \ldots, t_n)$, and from Step 2, all the constraints are locally nonbinding. Therefore, any solution to (18) must satisfy (19), (20), (21), and (22).

Step 4. It remains to show that (16) and (23) have the same maximum value. The logic is similar to the one in the proof of Proposition 1. We only need to show that, for any solution $q(\cdot)$ to (23), if $R(F(t)) > \hat{R}(F(t))$ then $q(\cdot)$ is constant in some neighborhood of t. Equivalently, we need to show that, if $(q_1^*, \ldots, q_n^*; t_1^*, \ldots, t_n^*)$ is a solution to (18), then whenever J is constant in some neighborhood of t, we have $t \neq t_i^*$ for all $i = 1, \ldots, n$. Suppose, by way of contradiction, that J is constant in some neighborhood of t_i^*. Then, since $(q_1^*, \ldots, q_n^*; t_1^*, \ldots, t_n^*)$ satisfies condition (20), a small change in t_i^* does not change the value of the objective function. But after this change in t_i^*, condition (19) for $i = i^*$ must be violated, and so one can adjust q_i^* accordingly to increase the value of the objective function, which is a contradiction. Alternatively, one can check the second-order necessary condition with respect to (t_i, q_i):

$$\begin{vmatrix} \frac{\partial^2(\cdot)}{\partial t_i^2} & \frac{\partial^2(\cdot)}{\partial q_i \partial t_i} \\ \frac{\partial^2(\cdot)}{\partial t_i \partial q_i} & \frac{\partial^2(\cdot)}{\partial q_i^2} \end{vmatrix} \geq 0 \text{ at } (q_1^*, \ldots, q_n^*; t_1^*, \ldots, t_n^*),$$

where we have omitted the objective function. Note that $\partial^2(\cdot)/\partial t_i^2$ evaluated at $(q_1^*, \ldots, q_n^*; t_1^*, \ldots, t_n^*)$ is $-J'(t_i^*) = 0$, and $\partial^2(\cdot)/\partial q_i \partial t_i$ evaluated at $(q_1^*, \ldots, q_n^*; t_1^*, \ldots, t_n^*)$ is $-[J(t_i^*)v'(q_i^*) - c'(q_i^*)]f(t_i^*)$. Thus, the above second-order necessary condition reduces to $J(t_i^*)v'(q_i^*) = c'(q_i^*)$. But it does not hold because from (20) and $q_{i-1}^* < q_i^*$,

$$J(t_i^*)v'(q_i^*) < J(t_i^*)\frac{v(q_i^*) - v(q_{i-1}^*)}{q_i^* - q_{i-1}^*} = \frac{c(q_i^*) - c(q_{i-1}^*)}{q_i^* - q_{i-1}^*} < c'(q_i^*).$$

Therefore, any solution to (23), which exists from Step 3, is also a solution to (16). ■

Thus, the optimal discrete offering $(q_1, \ldots, q_n; t_1, \ldots, t_n)$ in a constrained program solves a system of two difference equations. This system does not have a closed-form solution except for special cases (see Example 1). It has at least one solution, but the solution may not be unique. In fact, problem (18) may have multiple local maximizers; therefore, the first-order conditions (even together with second-order conditions) are only necessary, but not sufficient.[8]

[8] To construct an example with multiple local maximizers, one only needs to consider the one-variety case, i.e., $n = 1$. There could be a local maximizer at which a high-quality high-price variety is offered resulting

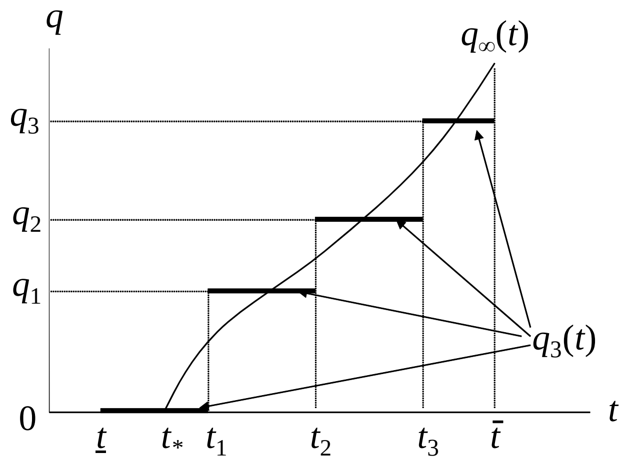

Figure 1: Comparison between the optimal continuous offering and the optimal discrete offering when the number of varieties is 3

Not surprisingly, both (19) and (20) converge to the unconstrained optimal offering formula $J(t)v'(q) = c'(q)$ as the consecutive t_i's and q_i's become closer and closer. The solution $q(\cdot)$ of the original constrained program (16) characterized by the above optimal discrete offering $(q_1, \ldots, q_n; t_1, \ldots, t_n)$ is denoted by $q_n(\cdot)$, which we also call the optimal discrete offering. From condition (20) and $q_{i-1} < q_i$, we must have

$$\frac{c'(q_{i-1})}{v'(q_{i-1})} < J(t_i) = \frac{c(q_i) - c(q_{i-1})}{v(q_i) - v(q_{i-1})} < \frac{c'(q_i)}{v'(q_i)},$$

and hence

$$q_{i-1} < q_\infty(t_i) < q_i.$$

Figure 1 sketches the pattern of an optimal discrete offering: the optimal discrete offering of a constrained program is a step-function approximation of the optimal continuous offering of the unconstrained program.

Given an optimal discrete offering $(q_1, \ldots, q_n; t_1, \ldots, t_n)$, the associated prices $p_1, \ldots, p_n$ of the n varieties can be easily computed as follows:

$$p_1 = t_1 q_1,$$

$$p_i = p_{i-1} + t_i(q_i - q_{i-1}) \quad \forall i = 2, \ldots, n.$$

in a small market coverage, and another one at which a low-quality low-price variety is offered resulting in a large market coverage. Such a situation can arise when J has a relatively flat part between two relatively steep parts (which easily occurs when K is not monotone), or when the function $q \mapsto c'(q)/v'(q)$ has a relatively steep part between two relatively flat parts.

The first equation is from the fact that a type t_1 consumer will be indifferent between buying the variety with quality q_1 and buying nothing. The second equation is from the fact that a type t_i consumer will be indifferent between buying the variety with quality q_i and quality q_{i-1}.

The following solvable "linear-quadratic-uniform" example reveals a number of features that we will generalize to our general setup.

Example 1 *Assume that $v(q) = q$ and $c(q) = q^2/2$, and that the distribution of consumers' types is uniform on the support $[0, 1]$, which implies that $J(t) = 2t - 1$. Then the optimal continuous offering is $q_\infty(t) = 2t - 1$ for $t \geq t_* = 1/2$. The unconstrained profit is $\Pi_\infty = 1/12$. If the monopolist restricts itself to offer at most n varieties, the optimal discrete offering $(q_1, \ldots, q_n; t_1, \ldots, t_n)$ can be computed using the first-order conditions* (19) *through* (22)*, namely,*

$$q_i = \frac{2i}{2n+1}, \quad t_i = \frac{n+i}{2n+1} \quad \forall i = 1, \ldots, n.$$

The corresponding constrained profit is[9]

$$\Pi_n = \frac{4n(n+1)}{(2n+1)^2}\Pi_\infty. \tag{25}$$

Thus, Π_n is monotonically increasing to Π_∞ as $n \to \infty$; the uncaptured profit

$$\Pi_\infty - \Pi_n = \frac{1}{(2n+1)^2}\Pi_\infty \tag{26}$$

is of order $1/n^2$. The ratio of constrained to unconstrained profit Π_n/Π_∞ is $24/25$ for $n = 2$. Therefore, a 2-variety menu can capture a large part of the unconstrained profit. The marginal benefit of adding one more variety

$$\Pi_{n+1} - \Pi_n = \frac{8(n+1)}{(2n+1)^2(2n+3)^2}\Pi_\infty$$

is monotonically decreasing to 0, and is of order $1/n^3$.[10] *The associated percentage change*

$$\frac{\Pi_{n+1} - \Pi_n}{\Pi_n} = \frac{2}{n(2n+3)^2}$$

[9] Equation (25) holds as long as consumers' utility is linear in quality, the seller's unit cost of production is quadratic in quality, and the distribution of consumers' types is uniform on an interval support. Hence, all the following results in this example generalize to the general "linear-quadratic-uniform" setting.

[10] One can also show that, in the linear-quadratic-uniform case, the social welfare (and consumers' surplus) converges at the same rate, i.e., the "uncaptured welfare" is of order $1/n^2$ and the "marginal welfare" is of order $1/n^3$. However, whether these convergence rates can be generalized to our general setup is unknown.

is approximately 8%, 2%, and 0.8% for $n = 1, 2,$ *and* 3*, respectively.*[11]

3.2 Visualization of constrained and unconstrained profits

In order to nicely visualize the constrained and unconstrained profits in our general setting and then prove our results, it is convenient to use a change of variable $x \equiv J(t)$, because the integrand in the constrained and unconstrained profits is linear in ironed virtual type $J(t)$. Define

$$H(q, x) \equiv x \cdot v(q) - c(q), \tag{27}$$

$$J^{-1}(x) \equiv \sup_t \{J(t) \geq x\}, \quad G(x) \equiv F(J^{-1}(x)) \quad \forall x \in [x_*, \bar{t}],$$

where

$$x_* \equiv J(t_*) = \frac{c'(0)}{v'(0)}.$$

$H(q, x)$ is the ironed virtual surplus at quality q and ironed virtual type x, and G is the distribution of ironed virtual types. Now, the unconstrained profit can be written as

$$\Pi_\infty = \int_{x_*}^{\bar{t}} \max_{q \geq 0} H(q, x)\, dG(x) = \int_{x_*}^{\bar{t}} H\left(q_\infty(J^{-1}(x)), x\right) dG(x). \tag{28}$$

Given n and the corresponding optimal discrete offering $(q_1, \ldots, q_n; t_1, \ldots, t_n)$, the constrained profit can be written as

$$\Pi_n = \sum_{i=1}^{n} \int_{x_i}^{x_{i+1}} H(q_i, x)\, dG(x), \tag{29}$$

where $x_i \equiv J(t_i)$ for $i = 1, \ldots, n$, and $x_{n+1} \equiv \bar{t}$.

Equation (20) can be written as

$$H(q_i, x_i) = H(q_{i-1}, x_i) \quad \forall i = 1, \cdots, n. \tag{30}$$

Fixing any q_i, the slope of $H(q_i, x)$ with respect to x is

$$\frac{\partial H(q_i, x)}{\partial x} = v(q_i) > 0.$$

Now, we can nicely visualize Π_∞ and Π_n in a single diagram.[12] First, note that for any i,

[11] The 2009 version of this paper also contains numerical simulations for other examples. In each of these examples, Π_3/Π_∞ is more than 97% and $(\Pi_4 - \Pi_3)/\Pi_3$ is less than 1%. See Wong (2009).

[12] Analyzing the social planner's problems or the perfect information monopolist problems (both constrained and unconstrained) amounts to only replacing $J(t)$ by t, or replacing G by F. All our results can be easily adapted there.

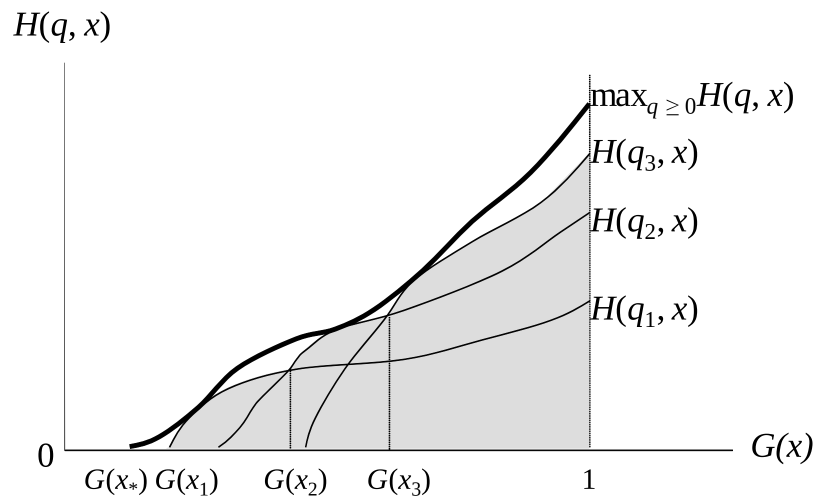

Figure 2: Visualizations of the unconstrained and constrained profits

the two curves $H(q_i, x)$ and $H(q_{i-1}, x)$ plotted against x must cross only once at $x = x_i$, because $H_{12} > 0$. When plotted against $G(x)$, they must cross only once at $G(x) = G(x_i)$. Moreover, the curve $\max_{q \geq 0} H(q, x)$ plotted against x or against $G(x)$ is the upper envelope of all the curves $H(q, x)$ with various values of q. The ideas are shown in Figure 2, where $n = 3$. From (28), it is clear that Π_∞ is the area below the bold curve $\max_{q \geq 0} H(q, x)$ in Figure 2. Moreover, from (29) and (30), Π_n is represented as the shaded area.

An important insight from Figure 2 is that each of the varieties offered helps capture the unconstrained profit Π_∞ in a first-order sense, since the slope of $H(q, x)$ with respect to quality q is flat at the ideal second best quality. The uncaptured profit is therefore of second or higher order. We will use this idea and apply Taylor's Theorem to show that the uncaptured profit is of order no more than $1/n^2$ in Subsection 4.2.

4 Properties of the Constrained Profit Sequence

4.1 Basic properties and diminishing marginal benefit

The sequence of constrained profit $\{\Pi_n\}_{n=0}^{\infty}$ has the following properties. (Note that Π_0 is also well defined: when $n = 0$, no variety can be offered; so, $q_0(t) = 0$ and $\Pi_0 = 0$. Of course, the analyses in Subsections 3.1 and 3.2 are only for $n \geq 1$.)

Proposition 3 *(i)* $\Pi_n > \Pi_{n-1}$ *for every* $n = 1, 2, \ldots$; *(ii)* $\lim_{n\to\infty} \Pi_n = \Pi_\infty$, *where* Π_∞ *is characterized in Proposition 1; and (iii)* $\lim_{n\to\infty} (\Pi_{n+1} - \Pi_n) = 0$.

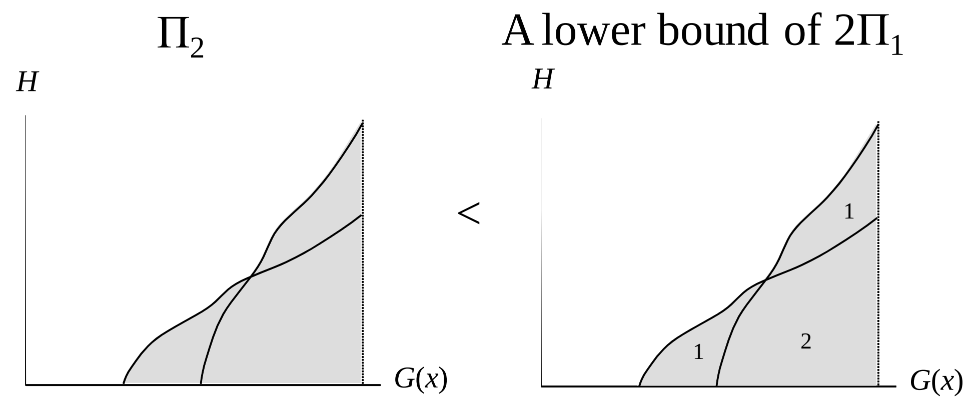

Figure 3: Comparing Π_2 and $2\Pi_1$

Proof. Obviously, $\Pi_1 > 0 = \Pi_0$. For $n \geq 2$, we again have $\Pi_n > \Pi_{n-1}$ because we know from Proposition 2 that constraint (17) in the n-variety constrained program is binding. This proves (i). Now, the sequence $\{\Pi_n\}_{n=0}^{\infty}$ is increasing and bounded from above by Π_∞. Thus, it has a finite limit, and $\lim_{n\to\infty} \Pi_n \leq \Pi_\infty$. To see the equality in (ii), note that as an increasing function, q_∞ can be arbitrarily well approximated by nondecreasing finite-step functions. Finally, since any convergent sequence in Euclidean space is also a Cauchy sequence, $\Pi_{n+1} - \Pi_n$ converges to zero as n goes to infinity. This proves (iii). ■

We also prove that the marginal constrained profit $\Pi_{n+1} - \Pi_n$ is strictly decreasing in n. Intuitively, as the number of varieties offered becomes larger, the space for improving profit by adding an extra variety becomes smaller and the effectiveness of the extra variety becomes less. However, this "diminishing marginal benefit property" is far from trivial, because adding one more variety would lead to an optimal adjustment of all the previously offered varieties. Although $\Pi_{n+1} - \Pi_n$ must ultimately diminish, it is rather surprising that the property holds for every n. This diminishing marginal benefit property is not only interesting on its own, but also crucial to our analysis of the rates of convergence of this marginal benefit and the monopoly choice of varieties, and to our result on Π_2 in the linear-quadratic model.

Theorem 1 (Diminishing marginal benefit) *The increments of* $\{\Pi_n\}_{n=0}^{\infty}$ *are decreasing, i.e.,* $\Pi_{n+1} - \Pi_n < \Pi_n - \Pi_{n-1}$ *for every* $n = 1, 2, \ldots$.

The proof of Theorem 1 is in the Appendix. To explain the idea behind the proof, we sketch the proof of the first two inequalities $\Pi_2 - \Pi_1 < \Pi_1 - \Pi_0$ and $\Pi_3 - \Pi_2 < \Pi_2 - \Pi_1$. The first one is equivalent to $\Pi_2 < 2\Pi_1$. The left panel of Figure 3 visualizes Π_2. Let $q_{1,2}$

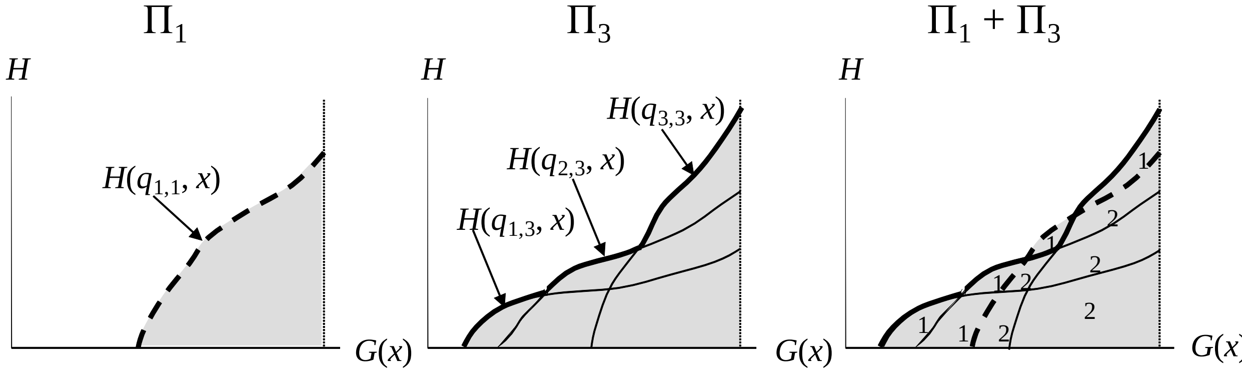

Figure 4: Visualization of $\Pi_1 + \Pi_3$

and $q_{2,2}$ be the two quality levels involved in the optimal discrete offering when $n = 2$. Now, imagine two plans of the monopolist: the first plan is to offer only one variety with quality $q_{1,2}$, and the second plan is to offer only one variety with quality $q_{2,2}$. The total profit from these two plans is visualized in the right panel of Figure 3, where a "2" in an area indicates that the area should be counted twice because that area accounts for profits from both plans, and a "1" in an area indicates that the area should be counted only once because that area accounts for profits from only one of the two plans. Now, $2\Pi_1$ must be greater than the total profit from the two plans; further, from Figure 3, it is obvious that the two plans make a total profit greater than Π_2. This proves that $\Pi_2 < 2\Pi_1$.

The second inequality, which is equivalent to $\Pi_1 + \Pi_3 < 2\Pi_2$, is harder. The shaded areas in the left and middle panels of Figure 4 visualize Π_1 and Π_3, respectively. The right panel of Figure 4 visualizes the sum $\Pi_1 + \Pi_3$. In this right panel, Π_1 is the area below the dashed curve and Π_3 the area below the bold solid curve. A "2" in an area indicates that the area should be counted twice because that area occurs in both Π_1 and Π_3. Similarly, a "1" in an area indicates that the area should be counted only once because that area occurs in either Π_1 or Π_3, not in both. Next, we show that $2\Pi_2$ must be greater than the area indicating $\Pi_1 + \Pi_3$. For this, we only need to construct two (suboptimal) 2-variety menus such that the sum of the two corresponding (suboptimal) 2-variety profits is larger than $\Pi_1 + \Pi_3$. The following procedure can do the job. First, let the optimal quality involved when $n = 1$ be $q_{1,1}$ and the optimal qualities involved when $n = 3$ be $q_{1,3}$, $q_{2,3}$, and $q_{3,3}$. Rank all the qualities involved in the above two constrained programs. For the example in Figure 4, this ranking is $q_{1,3} < q_{2,3} < q_{1,1} < q_{3,3}$. Then, collect those with odd ranking in one menu and those with even ranking in another menu. For the current example, the two menus are $(q_{1,3}, q_{1,1})$ and $(q_{2,3}, q_{3,3})$. While $\Pi_1 + \Pi_3$ is visualized in the left panel of Figure 5, the sum of the

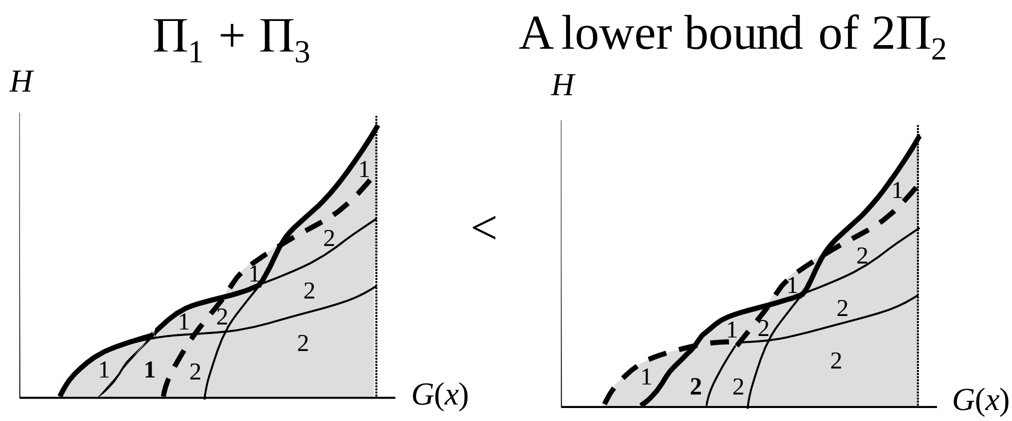

Figure 5: Comparing $\Pi_1 + \Pi_3$ and $2\Pi_2$

two 2-variety profits is visualized in the right panel of Figure 5. It is now easy to see that $\Pi_1 + \Pi_3 < 2\Pi_2$ because the two panels of Figure 5 are the same except that a "1" in the left panel is replaced by a "2" in the right panel.

The above logic is valid in general and is the intuition behind the proof of Theorem 1.

Corollary 1 *For every $n = 1, 2, \ldots,$*

$$\frac{\Pi_n - \Pi_{n-1}}{\Pi_n} \leq \frac{1}{n}.$$

Proof. Theorem 1 implies that, for every $n = 1, 2, \ldots,$

$$\Pi_n = (\Pi_n - \Pi_{n-1}) + (\Pi_{n-1} - \Pi_{n-2}) + \cdots (\Pi_1 - \Pi_0) \geq n\,(\Pi_n - \Pi_{n-1})\,.$$

■

4.2 Rate of convergence results

In this subsection, we provide our rate of convergence results for the uncaptured profit $\Pi_\infty - \Pi_n$, the marginal profit $\Pi_{n+1} - \Pi_n$, and the monopoly choice of the number of varieties.

Theorem 2 (Quadratic rate result) *There exists a finite constant $M_0 < \infty$, not depending on n, such that for every $n = 0, 1, 2, \ldots,$*

$$\Pi_\infty - \Pi_n \leq \frac{M_0}{(2n+1)^2}. \tag{31}$$

Thus, $\Pi_\infty - \Pi_n = O(1/n^2)$*. Moreover, if the distribution* G *of ironed virtual types, given by* $G(J(t)) = F(t)$*, admits a density* g *over the support* $[J(t_*), \bar{t}]$ *that is bounded from above by* $\bar{g}$*, then a tight bound is given by*

$$M_0 = \frac{M_1 \bar{g}}{6} \cdot \left(\bar{t} - \frac{c'(0)}{v'(0)} \right)^3, \tag{32}$$

where

$$M_1 \equiv \sup_{q \in [0, q(\bar{t})]} \left\{ \frac{(v'(q))^3}{v'(q)c''(q) - c'(q)v''(q)} \right\} < \infty.$$

Thus, the uncaptured profit $\Pi_\infty - \Pi_n$ is of order no more than $1/n^2$. From Example 1, this convergence rate is tight. The bound given by (31) and (32) is also tight. To see this, note that in Example 1, M_1 is 1 and $\bar{g}$ is $1/2$, and so the bound of $\Pi_\infty - \Pi_n$ given by Theorem 2 is exactly the same as the explicit solution of $\Pi_\infty - \Pi_n$ given by (26).

The proof of Theorem 2 is in the Appendix. The underlying reason for this quadratic rate result has already been explained in Introduction and Subsection 3.2. The main point is that the slope of the ironed virtual surplus function $J(t)v(q) - c(q)$ with respect to quality q is flat at its maximizer $q_\infty(t)$. A Taylor expansion argument shows that the uncaptured profit is of order no more than $1/n^2$ as long as the ironed virtual surplus function is smooth enough in q. The smoothness condition required is precisely the twice continuous differentiability of the value function v and cost function c, which is standard in the literature. Moreover, as the proof shows, this convergence rate of Π_n can be attained by a simple offering rule, which involves only a uniformly distributed set of (suboptimal) varieties.

Theorem 3 (Cubic rate result) *For every* $n = 1, 2, \ldots,$

$$\Pi_{n+1} - \Pi_n < \frac{27}{8n^2(2n+3)} M_0,$$

where M_0 *is a bound of* $(2n+1)^2 (\Pi_\infty - \Pi_n)$*, which exists from Theorem 2. Thus,* $\Pi_{n+1} - \Pi_n = O(1/n^3)$*.*

Proof. For any positive integers n, i with $n \geq i \geq 1$, we have

$$\Pi_\infty - \Pi_i = (\Pi_\infty - \Pi_{n+1}) + (\Pi_{n+1} - \Pi_n) + \cdots + (\Pi_{i+1} - \Pi_i). \tag{33}$$

From Theorem 2, the left-hand side of (33) is bounded by

$$\Pi_\infty - \Pi_i \leq \frac{M_0}{(2i+1)^2}.$$

The right-hand side of (33) is bounded by

$$(\Pi_\infty - \Pi_{n+1}) + (\Pi_{n+1} - \Pi_n) + \cdots + (\Pi_{i+1} - \Pi_i) > (n - i + 1)\left(\Pi_{n+1} - \Pi_n\right),$$

from Theorem 1 and the fact that $\Pi_{n+1} < \Pi_\infty$. Combining the above results, we have

$$\Pi_{n+1} - \Pi_n < \frac{M_0}{(n-i+1)(2i+1)^2} \quad \forall i = 1, \ldots, n.$$

It remains to find a tight lower bound for

$$\max_{i \in \{1,\ldots,n\}} (n-i+1)(2i+1)^2.$$

Since the unique maximizer of $(n-i+1)(2i+1)^2$ on $\mathbb{R}_+$ is $(4n+3)/6$, let us define $i^*(n)$ as the largest integer that does not exceed $(4n+3)/6$. Then,

$$\frac{4n-3}{6} < i^*(n) \le \frac{4n+3}{6},$$

$$\begin{aligned}\max_{i \in \{1,\ldots,n\}} (n-i+1)(2i+1)^2 &\ge (n - i^*(n) + 1)(2i^*(n)+1)^2 \\ &> \left(n - \frac{4n+3}{6} + 1\right)\left(2\left(\frac{4n-3}{6}\right) + 1\right)^2 \\ &= \frac{8}{27} n^2 \left(2n+3\right).\end{aligned}$$

Therefore,

$$\Pi_{n+1} - \Pi_n < \frac{M_0}{\max_{i \in \{1,\ldots,n\}} (n-i+1)(2i+1)^2} < \frac{27}{8n^2\left(2n+3\right)} M_0.$$

■

Theorems 1 and 3 together tell us that the marginal constrained profit $\Pi_n - \Pi_{n-1}$ monotonically converges to its limit zero at the cubic rate.

If there is a (per consumer) fixed cost $k > 0$ of developing each variety and the monopolist can freely choose the number of varieties, then the optimal number of varieties n^* is the maximizer of the profit net of the cost of developing varieties. As k goes to zero, the optimal number of varieties $n^*(k)$ certainly goes to infinity. But the following corollary of Theorems 1 and 3 tells us that $n^*(k)$ goes to infinity at an extremely slow rate; more precisely, it is of order no more than $1/k^{1/3}$.

Corollary 2 *Let $n^*(k)$ be an (usually unique) optimal number of varieties when the per consumer cost of developing every variety is $k > 0$, i.e., $n^*(k)$ maximizes $\Pi_n - nk$ over $n \in \{0, 1, 2, \ldots\}$. Then, $n^*(k) = O(1/k^{1/3})$ as $k \to 0$.*

Proof. If k is too large, $n^*(k) = 0$. Suppose $k > 0$ is not so large, such that $n^*(k) > 0$. Since $\Pi_{n+1} - \Pi_n$ is decreasing in n from Theorem 1, $n^*(k)$ is characterized by $\Pi_{n^*(k)+1} - \Pi_{n^*(k)} \le k$ and $\Pi_{n^*(k)} - \Pi_{n^*(k)-1} \ge k$. It follows from Theorem 3 that

$$k^{1/3} n^*(k) \le \left(\Pi_{n^*(k)} - \Pi_{n^*(k)-1}\right)^{1/3} n^*(k) = O(1).$$

■

4.3 Performance of 2-variety menus in linear-quadratic model

In this subsection we show that under the following assumptions, a 2-variety menu can capture a large part of unconstrained profit. (Recall that $\Pi_2 = \frac{24}{25}\Pi_\infty$ in Example 1.) First, consumers' utility is linear in quality and the seller's unit cost of production is quadratic in quality. That is, the functions v and c are

$$v(q) = A_0 q, \quad c(q) = A_1 q + \frac{A_2}{2} q^2, \tag{34}$$

where A_0 and A_2 are positive constants and A_1 is a nonnegative constant. (Our assumptions (3) and $J(\underline{t})v'(0) < c'(0)$ reduce to $(2\underline{t} - \bar{t})\, A_0 < A_1 < \bar{t} A_0$.) Second, the distribution G of ironed virtual types, given by $G(J(t)) \equiv F(t)$, has a positive density g on the support $[x_*, \bar{t}]$, where $x_* \equiv J(t_*) = c'(0)/v'(0)$.[13] Third, the distribution G satisfies the following regularity condition:

$$\frac{G(x) - G(x_*)}{g(x)} \text{ and } -\frac{1 - G(x)}{g(x)} \text{ are nondecreasing in } x \text{ on } [x_*, \bar{t}]. \tag{35}$$

[14]

Consider an auxiliary problem, which is the constrained program with two modifications. First, the types below t_* (where $J(t_*)A_0 = A_1$) are thrown out. Second, all types above t_* have to be covered (i.e., the outside option of buying nothing is not available). For any

[13] This in particular implies that J is strictly increasing (and hence $J = K$) on $[t_*, \bar{t}]$, i.e., there is no bunching in the optimal continuous offering.

[14] Equivalently, it says $\frac{F(t)}{f(t)} J'(t)$ and $-\frac{1-F(t)}{f(t)} J'(t)$ are nondecreasing in t on $[t_*, \bar{t}]$.

positive integer n or $n = \infty$, the maximum profit of the auxiliary problem is then[15]

$$\begin{aligned}\hat{\Pi}_n &= \max_{q(\cdot)\geq 0} \int_{t_*}^{\bar{t}} [J(t)v(q(t)) - c(q(t))]\, dF(t) \\ &= \max_{q(\cdot)\geq 0} \int_{t_*}^{\bar{t}} \left[(J(t)A_0 - A_1)\, q(t) - \frac{A_2}{2} (q(t))^2 \right] dF(t)\end{aligned}$$

subject to

$$q(\cdot) \text{ takes at most } n \text{ values (not except zero).}$$

We also define $\hat{\Pi}_0 \equiv 0$.

There are two remarks to make. First, $\hat{\Pi}_n < \Pi_n$ for any $n \notin \{0, \infty\}$, and $\hat{\Pi}_\infty = \Pi_\infty$. Second, it can be easily seen that all of our previous results for $\{\Pi_n\}_{n=0}^{\infty}$ also hold for $\{\hat{\Pi}_n\}_{n=0}^{\infty}$. In particular, the diminishing marginal benefit result can be proved in exactly the same way. (Of course, the bounds in our rate of convergence results need to be modified. For example, the factor $(2n+1)^2$ in Theorem 2 should be replaced by $(2n)^2$.)

Lemma 1 *Considering the auxiliary problem, the benefit of going from offering one variety to two varieties accounts for at least half of the total benefit of going all the way to the optimal continuous offering, i.e.,*

$$\hat{\Pi}_2 - \hat{\Pi}_1 \geq \Pi_\infty - \hat{\Pi}_2.$$

The proof of Lemma 1, which uses the technique in McAfee (2002), is in the Appendix. Basically, the proof considers a suboptimal 2-variety menu, with one variety serving the consumers with below-mean virtual types, and the other serving the consumers with above-mean virtual types. It can be shown that the sum of Π_∞ and $\hat{\Pi}_1$, which have explicit forms under linear-quadratic model, cannot be larger than two times the profit attained by the above suboptimal 2-variety menu.

Lemma 1, together with diminishing marginal benefit, implies that offering only two varieties can capture more than two-thirds of the unconstrained maximum profit.

Theorem 4 (Two-thirds result)

$$\Pi_2 > \hat{\Pi}_2 > \frac{2}{3}\Pi_\infty.$$

Proof. As noted previously, $\{\hat{\Pi}_n\}_{n=0}^{\infty}$ has the diminishing marginal benefit property. In

[15] Unlike our original constrained program, the integrand in $\hat{\Pi}_n$ evaluated at the optimal discrete offering is typically negative for very low types.

particular, $\hat{\Pi}_2 < 2\hat{\Pi}_1$. This, together with Lemma 1, implies that

$$\begin{aligned}\Pi_\infty &= \left(\Pi_\infty - \hat{\Pi}_2\right) + \left(\hat{\Pi}_2 - \hat{\Pi}_1\right) + \hat{\Pi}_1 \\ &\leq 2\left(\hat{\Pi}_2 - \hat{\Pi}_1\right) + \hat{\Pi}_1 < 3\hat{\Pi}_1.\end{aligned}$$

Thus, $\hat{\Pi}_1 > \Pi_\infty/3$, and hence

$$\Pi_\infty - \hat{\Pi}_2 \leq \hat{\Pi}_2 - \hat{\Pi}_1 < \hat{\Pi}_2 - \Pi_\infty/3$$

$$2\hat{\Pi}_2 > 4\Pi_\infty/3.$$

■

5 Discrete Type Model

In this section, we consider a discrete type version of the above model and show how our results are adapted. Since there are only a finite number of consumers' types now, the fully optimal profit can be attained by offering a finite number of varieties, and the rate of convergence is no more a concern. But one might wonder whether the diminishing marginal benefit property and the two-thirds result still hold. It turns out that the diminishing marginal benefit property can be proved in the same manner as before. However, the two-thirds result holds only approximately, and the adaptation of regularity condition (35) is not completely immediate. All the proofs for this section are in the Appendix.

Suppose that there are $M < \infty$ possible types of consumers. For each $m = 1, \ldots, M$, a type m consumer would have utility $t_m \cdot v(q) - p$ if he buys a variety with quality $q \geq 0$ and pays the price p, where $0 < t_1 < \cdots < t_M$. The seller believes that each consumer is of type m with probability $f_m > 0$. Also define $F_m \equiv \sum_{m'=1}^{m} f_{m'}$ for each $m = 1, \ldots, M$. The other modeling details are the same as in Subsection 2.1.

For each $m = 1, \ldots, M$, define

$$K_m \equiv t_m - (t_{m+1} - t_m)\frac{1 - F_m}{f_m}$$

as the "virtual type" of type m consumers.[16] Now, given a maximum number of varieties n, the constrained program and constrained profit (or unconstrained program and profit if

[16] How t_{M+1} is defined is not important since K_M is anyhow equal to t_M.

$n \geq M$) can be written as

$$\Pi_n \equiv \max_{q_1,\ldots,q_M \in \mathbb{R}_+} \sum_{m=1}^{M} \left[K_m v(q_m) - c(q_m)\right] f_m \tag{36}$$

subject to

$$q_1 \leq q_2 \leq \cdots \leq q_M, \tag{37}$$

$$(q_1, \ldots, q_M) \text{ takes at most } n \text{ values except zero.} \tag{38}$$

It is not hard to see that the constrained program has a solution so that the constrained profit Π_n is well defined.

As in Subsection 2.2, the virtual types can be ironed so that the above monotonicity constraint can be dropped. Here we bypass the convex hull-type construction but directly define a vector $(J_1, \ldots, J_M)$ of "ironed virtual types" by its characterizing properties, which are discrete type versions of (9) through (12). Formally, we define $(J_1, \ldots, J_M)$ as the unique vector that satisfies the following properties:

$$J_1 \leq J_2 \leq \cdots \leq J_M, \tag{39}$$

$$\hat{R}_m \equiv \sum_{m'=1}^{m} J_{m'} f_{m'} \leq \sum_{m'=1}^{m} K_{m'} f_{m'} \equiv R_m \quad \forall m = 1, \ldots, M, \tag{40}$$

$$(40) \text{ holds as equality at } m = M, \tag{41}$$

$$\text{whenever } (40) \text{ is strict at } m,\ J_m = J_{m+1}. \tag{42}$$

Proposition 4 *In the discrete type model, the constrained profit given n (or unconstrained profit if $n \geq M$) can be written as*

$$\Pi_n = \max_{q_1,\ldots,q_M \in \mathbb{R}_+} \left\{ \sum_{m=1}^{M} \left[J_m v(q_m) - c(q_m)\right] f_m \ \ s.t.\ (38) \right\}. \tag{43}$$

Define M_* as the lowest type m such that $J_m v'(0) > c'(0)$. Clearly, offering $M - M_* + 1$ varieties for types $M_*, \ldots, M$ is enough to attain the fully optimal profit, i.e., $\Pi_n = \Pi_{M-M_*+1}$ for all $n \geq M - M_* + 1$.

Theorem 5 *In the discrete type model, the diminishing marginal benefit property holds, i.e. $\Pi_{n+1} - \Pi_n \leq \Pi_n - \Pi_{n-1}$ for every $n = 1, 2, \ldots$. Moreover, if (i) the consumers' utility and the seller's unit cost of production take the form as in* (34) *(i.e., linear-quadratic in quality),* *and (ii)*

$$(J_{m+1} - J_m)\,\frac{F_m - F_{M_*-1}}{f_m}, \quad (J_{m+1} - J_m)\,\frac{F_m - F_{M_*-1}}{f_{m+1}},$$
$$-\,(J_{m+1} - J_m)\,\frac{1 - F_m}{f_m}, \quad -\,(J_{m+1} - J_m)\,\frac{1 - F_m}{f_{m+1}}$$

are all nondecreasing in m, then

$$\Pi_2 \geq \frac{2}{3}\Pi_{M-M_*+1} - \frac{A_0^2\,(1 - F_{M_*-1})\,\delta}{3A_2},$$

where

$$\delta \equiv \min_{\hat{m}} \left(J_{\hat{m}} - \frac{\sum_{m=M_*}^{M} J_m f_m}{1 - F_{M_*-1}} \right)^2.$$

Appendix

In order to prove Theorem 1, we first introduce a simple lemma.

Lemma 2 *Let S_1 and S_2 be two vectors of real numbers. (Their dimensions could be different.) Then*

$$\max\nolimits_{(2)}(S_1, S_2) \geq \min\{\max(S_1), \max(S_2)\},$$

where $\max_{(2)}(S)$ denotes the second largest element in S.

Proof. If $\max(S_1, S_2)$ is in S_1, then $\max_{(2)}(S_1, S_2) \geq \max(S_2)$. If $\max(S_1, S_2)$ is in S_2, then $\max_{(2)}(S_1, S_2) \geq \max(S_1)$. In both cases, our claim is true. ■

Proof of Theorem 1. Let the optimal discrete offering given any number n of varieties be $(q_{1,n}, \ldots, q_{n,n}; t_{1,n}, \ldots, t_{n,n})$, and let $x_{i,n} \equiv J(t_{i,n})$. Note that, for all $i, m \in \{1, 2, \ldots, n\}$ and all $x \in [x_{i,n}, x_{i+1,n}]$, we have $H(q_{i,n}, x) \geq H(q_{m,n}, x)$ (see Figure 2). Therefore, it follows from (29) that, for any n,

$$\Pi_n = \int_{x_*}^{\bar{t}} \max\left(0, H(q_{1,n}, x), H(q_{2,n}, x), \ldots, H(q_{n,n}, x)\right) dG(x).$$

Now, for every x and n define the vector $S_n(x)$ as

$$S_n(x) \equiv \left(0, H(q_{1,n}, x), H(q_{2,n}, x), \ldots, H(q_{n,n}, x)\right).$$

Then,

$$\Pi_{n-1} = \int_{x_*}^{\bar{t}} \max(S_{n-1}(x))dG(x).$$

Similarly,

$$\Pi_{n+1} = \int_{x_*}^{\bar{t}} \max(S_{n+1}(x))dG(x).$$

Now, let us construct two suboptimal n-variety menus as follows. First, arrange the $2n$ numbers $q_{1,n-1}, q_{2,n-1}, \ldots, q_{n-1,n-1}, q_{1,n+1}, q_{2,n+1}, \ldots, q_{n+1,n+1}$ in ascending order and denote the ordered numbers as $q_{(1)}, q_{(2)}, \ldots, q_{(2n)}$, with the convention that $q_{(i)}$ is the i-th smallest number. Let us also define $q_{(0)}$ as 0.

Construct the first menu to include $q_{(1)}, q_{(3)}, q_{(5)}, \ldots, q_{(2n-1)}$, and the second menu to include $q_{(2)}, q_{(4)}, q_{(6)}, \ldots, q_{(2n)}$. The corresponding x_i's are constructed optimally given the quality offers, so that the corresponding profits $\hat{\Pi}_n^{odd}$ and $\hat{\Pi}_n^{even}$ for these two menus are

$$\hat{\Pi}_n^{odd} = \int_{x_*}^{\bar{t}} \max\left(0, H(q_{(1)}, x), H(q_{(3)}, x), \ldots, H(q_{(2n-1)}, x)\right) dG(x),$$

$$\hat{\Pi}_n^{even} = \int_{x_*}^{\bar{t}} \max\left(0, H(q_{(2)}, x), H(q_{(4)}, x), \ldots, H(q_{(2n)}, x)\right) dG(x).$$

Note that for any x, if

$$\max\left(0, 0, H(q_{(1)}, x), H(q_{(2)}, x), H(q_{(3)}, x), \ldots, H(q_{(2n)}, x)\right)$$

is $H(q_{(i)}, x)$, then the corresponding second largest element

$$\max_{(2)}\left(0, 0, H(q_{(1)}, x), H(q_{(2)}, x), H(q_{(3)}, x), \ldots, H(q_{(2n)}, x)\right)$$

must be either $H(q_{(i-1)}, x)$ or $H(q_{(i+1)}, x)$, because $H(q, x)$ satisfies increasing differences (see Figure 2). Therefore,

$$\begin{aligned}
\hat{\Pi}_n^{odd} + \hat{\Pi}_n^{even} &= \int_{x_*}^{\bar{t}} \Big[\max\left(0, 0, H(q_{(1)}, x), H(q_{(2)}, x), H(q_{(3)}, x), \ldots, H(q_{(2n)}, x)\right) \\
&\quad + \max_{(2)}\left(0, 0, H(q_{(1)}, x), H(q_{(2)}, x), H(q_{(3)}, x), \ldots, H(q_{(2n)}, x)\right)\Big] dG(x) \\
&= \int_{x_*}^{\bar{t}} \max(S_{n-1}(x), S_{n+1}(x))dG(x) + \int_{x_*}^{\bar{t}} \max_{(2)}(S_{n-1}(x), S_{n+1}(x))dG(x).
\end{aligned}$$

On the other hand,

$$\begin{aligned}\Pi_{n-1}+\Pi_{n+1} &= \int_{x_*}^{\bar{t}}[\max(S_{n-1}(x))+\max(S_{n+1}(x))]dG(x)\\ &= \int_{x_*}^{\bar{t}}[\max\{\max(S_{n-1}(x)),\max(S_{n+1}(x))\}\\ &\quad+\min\{\max(S_{n-1}(x)),\max(S_{n+1}(x))\}]dG(x)\\ &= \int_{x_*}^{\bar{t}}[\max(S_{n-1}(x),S_{n+1}(x))+\min\{\max(S_{n-1}(x)),\max(S_{n+1}(x))\}]dG(x).\end{aligned}$$

Therefore,

$$\begin{aligned}&\left(\hat{\Pi}_n^{odd}+\hat{\Pi}_n^{even}\right)-(\Pi_{n-1}+\Pi_{n+1})\\ =&\int_{x_*}^{\bar{t}}[\max_{(2)}(S_{n-1}(x),S_{n+1}(x))-\min\{\max(S_{n-1}(x)),\max(S_{n+1}(x))\}]dG(x)\\ \geq&\ 0.\end{aligned}$$

The last inequality is from Lemma 2.

Now, we have

$$2\Pi_n-(\Pi_{n-1}+\Pi_{n+1})\geq\left(\hat{\Pi}_n^{odd}+\hat{\Pi}_n^{even}\right)-(\Pi_{n-1}+\Pi_{n+1})\geq 0, \tag{44}$$

which implies that $\Pi_n-\Pi_{n-1}\geq\Pi_{n+1}-\Pi_n$.

It remains to show that at least one inequality in (44) is strict. The second inequality in (44) is strict unless

$$\max_{(2)}(S_{n-1}(x),S_{n+1}(x))=\min\{\max(S_{n-1}(x)),\max(S_{n+1}(x))\} \text{ for almost all } x.$$

By the definitions of $S_{n-1}(x)$ and $S_{n+1}(x)$ and from the fact that $S_{n+1}(x)$ has two more elements than $S_{n-1}(x)$, the above cannot hold unless

$$\begin{aligned}q_{1,n-1} &= q_{2,n+1}\\ q_{2,n-1} &= q_{3,n+1}\\ &\vdots\\ q_{n-2,n-1} &= q_{n-1,n+1}\\ q_{n-1,n-1} &= q_{n,n+1}.\end{aligned}$$

But given the above, the constructed varieties for $\hat{\Pi}_n^{odd}$ and $\hat{\Pi}_n^{even}$ are $(q_{1,n+1}, q_{2,n+1}, \ldots, q_{n,n+1})$ and $(q_{2,n+1}, q_{3,n+1}, \ldots, q_{n+1,n+1})$. Therefore, at least one inequality in (44) is strict unless there exist some optimal offers $q_{1,n+1}, \ldots, q_{n+1,n+1}$ for the $(n+1)$-variety problem such that

1. $q_{1,n+1}, q_{2,n+1}, \ldots, q_{n,n+1}$ are some optimal offers for the n-variety problem;
2. $q_{2,n+1}, q_{3,n+1}, \ldots, q_{n+1,n+1}$ are some optimal offers for the n-variety problem; and
3. $q_{2,n+1}, q_{3,n+1}, \ldots, q_{n-1,n+1}$ are some optimal offers for the $(n-1)$-variety problem.

However, from our first-order conditions/difference equations (19) and (20), these are impossible. ■

Proof of Theorem 2. From (28),

$$\Pi_\infty = \int_{x_*}^{\bar{t}} \max_{q \geq 0} H(q, x)\, dG(x) = \int_{x_*}^{\bar{t}} H(\tilde{q}(x), x)\, dG(x),$$

where $\tilde{q}(x) \equiv q_\infty(J^{-1}(x))$ for every $x \in [x_*, \bar{t}]$, or equivalently

$$xv'(\tilde{q}(x)) = c'(\tilde{q}(x)).$$

Given some number of varieties n, any discrete offering can be expressed in terms of ironed virtual types instead of types. Let q_i be the quality to serve consumers with ironed virtual types $x \in [x_i, x_{i+1}]$. Now consider the (suboptimal) discrete offering characterized by

$$x_i = \frac{c'(0)}{v'(0)} + \frac{2i-1}{2n+1}\left(\bar{t} - \frac{c'(0)}{v'(0)}\right),$$

$$q_i = \tilde{q}\left(\frac{x_i + x_{i+1}}{2}\right).$$

Then the corresponding (suboptimal) n-variety profit is

$$\hat{\Pi}_n = \sum_{i=1}^{n} \int_{x_i}^{x_{i+1}} H(q_i, x) dG(x).$$

Subtracting $\hat{\Pi}_n$ from Π_∞ and noticing that $x_* = c'(0)/v'(0) < x_1 < \cdots < x_n < x_{n+1} = \bar{t}$, we have

$$\Pi_\infty - \hat{\Pi}_n = \int_{x_*}^{x_1} H(\tilde{q}(x), x) dG(x) + \sum_{i=1}^{n} \int_{x_i}^{x_{i+1}} \left[H(\tilde{q}(x), x) - H(q_i, x)\right] dG(x). \tag{45}$$

The integrand of the first term of the right-hand side of (45), as a twice continuously differentiable function of x, can be Taylor expanded around x_* as

$$x_* v(\tilde{q}(x_*)) - c(\tilde{q}(x_*)) + (x - x_*)\, v(\tilde{q}(x_*)) + \frac{1}{2}(x - x_*)^2\, v'(\tilde{q}(\hat{x}_0))\tilde{q}'(\hat{x}_0)$$

for some $\hat{x}_0 \in [x_*, x_1]$. Because $\tilde{q}(x_*) = \tilde{q}(c'(0)/v'(0)) = 0$, this can be further simplified as

$$\frac{1}{2}\left(x - \frac{c'(0)}{v'(0)}\right)^2 v'(\tilde{q}(\hat{x}_0))\tilde{q}'(\hat{x}_0).$$

The integrand of other terms of the right-hand side of (45), as a twice continuously differentiable function of x, can be Taylor expanded around $(x_i + x_{i+1})/2$ as

$$\begin{aligned} & \left(\frac{x_i + x_{i+1}}{2} v(q_i) - c(q_i)\right) - \left(\frac{x_i + x_{i+1}}{2} v(q_i) - c(q_i)\right) \\ & + \left(x - \frac{x_i + x_{i+1}}{2}\right)(v(q_i) - v(q_i)) + \frac{1}{2}\left(x - \frac{x_i + x_{i+1}}{2}\right)^2 v'(\tilde{q}(\hat{x}_i))\tilde{q}'(\hat{x}_i) \\ = & \ \frac{1}{2}\left(x - \frac{x_i + x_{i+1}}{2}\right)^2 v'(\tilde{q}(\hat{x}_i))\tilde{q}'(\hat{x}_i) \end{aligned}$$

for some $\hat{x}_i \in [x_i, x_{i+1}]$. By definition of $\tilde{q}(\cdot)$, for any $x \in [x_*, \bar{t}]$,

$$v'(\tilde{q}(x)) + x v''(\tilde{q}(x))\tilde{q}'(x) = c''(\tilde{q}(x))\tilde{q}'(x)$$

and hence

$$\begin{aligned} v'(\tilde{q}(x))\tilde{q}'(x) &= \frac{(v'(\tilde{q}(x)))^2}{c''(\tilde{q}(x)) - x v''(\tilde{q}(x))} \\ &\leq \sup_{x \in [x_*, \bar{t}]} \left\{ \frac{(v'(\tilde{q}(x)))^2}{c''(\tilde{q}(x)) - x v''(\tilde{q}(x))} \right\} \\ &= \sup_{t \in [t_*, \bar{t}]} \left\{ \frac{(v'(q_\infty(t)))^2}{c''(q_\infty(t)) - J(t) v''(q_\infty(t))} \right\}. \end{aligned}$$

Since $J(t) = c'(q_\infty(t))/v'(q_\infty(t))$,

$$\begin{aligned} v'(\tilde{q}(x))\tilde{q}'(x) &\leq \sup_{t \in [t_*, \bar{t}]} \left\{ \frac{(v'(q_\infty(t)))^3}{v'(q_\infty(t)) c''(q_\infty(t)) - c'(q_\infty(t)) v''(q_\infty(t))} \right\} \\ &= \sup_{q \in [0, q_\infty(\bar{t})]} \left\{ \frac{(v'(q))^3}{v'(q) c''(q) - c'(q) v''(q)} \right\} \equiv M_1. \end{aligned}$$

Note that $M_1 < \infty$, because it is the supremum of a continuous function over a compact set.

Now, the first term of the right-hand side of (45) is

$$\begin{aligned} & \int_{c'(0)/v'(0)}^{x_1} \frac{1}{2}\left(x - \frac{c'(0)}{v'(0)}\right)^2 v'(\tilde{q}(\hat{x}_0))\tilde{q}'(\hat{x}_0)dG(x) \\ \leq & \int_{c'(0)/v'(0)}^{x_1} \frac{1}{2}\left(x - \frac{c'(0)}{v'(0)}\right)^2 M_1 dG(x) \\ < & \int_{c'(0)/v'(0)}^{x_1} \frac{1}{2}\left(x_1 - \frac{c'(0)}{v'(0)}\right)^2 M_1 dG(x) \\ = & \frac{M_1}{2(2n+1)^2}\left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^2 \left(G\left(x_1\right) - G(x_*)\right). \end{aligned} \tag{46}$$

The other terms of the right-hand side of (45) are

$$\begin{aligned} & \int_{x_i}^{x_{i+1}} \frac{1}{2}\left(x - \frac{x_i + x_{i+1}}{2}\right)^2 v'(\tilde{q}(\hat{x}_i))\tilde{q}'(\hat{x}_i)dG(x) \\ \leq & \int_{x_i}^{x_{i+1}} \frac{1}{2}\left(x - \frac{x_i + x_{i+1}}{2}\right)^2 M_1 dG(x) \\ < & \int_{x_i}^{x_{i+1}} \frac{1}{2}\max\left\{\left(x_i - \frac{x_i + x_{i+1}}{2}\right)^2, \left(x_{i+1} - \frac{x_i + x_{i+1}}{2}\right)^2\right\} M_1 dG(x) \\ = & \frac{M_1}{2(2n+1)^2}\left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^2 \left(G\left(x_{i+1}\right) - G(x_i)\right). \end{aligned} \tag{47}$$

Therefore,

$$\begin{aligned} \Pi_\infty - \Pi_n & \leq \Pi_\infty - \hat{\Pi}_n \\ & < \frac{M_1}{2(2n+1)^2}\left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^2 \left(G\left(x_1\right) - G(x_*) + \sum_{i=1}^{n}\left(G\left(x_{i+1}\right) - G(x_i)\right)\right) \\ & = \frac{M_1}{2(2n+1)^2}\left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^2 \left(G\left(\bar{t}\right) - G(x_*)\right) \\ & = \frac{M_1}{2(2n+1)^2}\left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^2 \left(1 - F(t_*)\right). \end{aligned}$$

This proves (31) with

$$M_0 \equiv \frac{M_1}{8} \cdot \left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^2 \left(1 - F(t_*)\right).$$

Moreover, if $G(\cdot)$ admits a density $g(\cdot)$ over the support $[x_*, \bar{t}]$ that is bounded from above by $\bar{g}$, then line (46) is at most

$$\begin{aligned}
\frac{1}{2} M_1 \bar{g} \int_{c'(0)/v'(0)}^{x_1} \left(x - \frac{c'(0)}{v'(0)}\right)^2 dx &= \frac{1}{6} M_1 \bar{g} \left(x_1 - \frac{c'(0)}{v'(0)}\right)^3 \\
&= \frac{M_1 \bar{g}}{6\,(2n+1)^3} \left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^3,
\end{aligned}$$

and line (47) is at most

$$\begin{aligned}
\frac{1}{2} M_1 \bar{g} \int_{x_i}^{x_{i+1}} \left(x - \frac{x_i + x_{i+1}}{2}\right)^2 dx &= \frac{1}{6} M_1 \bar{g} \left[\left(\frac{x_{i+1} - x_i}{2}\right)^3 - \left(\frac{x_i - x_{i+1}}{2}\right)^3\right] \\
&= \frac{2 M_1 \bar{g}}{6\,(2n+1)^3} \left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^3.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\Pi_\infty - \Pi_n &\leq \Pi_\infty - \hat{\Pi}_n \\
&\leq \frac{(2n+1)\, M_1 \bar{g}}{6\,(2n+1)^3} \left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^3 \\
&= \frac{M_1 \bar{g}}{6\,(2n+1)^2} \left(\bar{t} - \frac{c'(0)}{v'(0)}\right)^3.
\end{aligned}$$

■

Proof of Lemma 1. The maximized values Π_∞, $\hat{\Pi}_1$, and $\hat{\Pi}_2$ can be solved as

$$\Pi_\infty = \frac{1}{2A_2} \int_{t_*}^{\bar{t}} (J(t)A_0 - A_1)^2 \, dF(t) = \frac{1}{2A_2} \int_{x_*}^{\bar{t}} (xA_0 - A_1)^2 \, dG(x),$$

$$\hat{\Pi}_1 = \frac{1}{2A_2} \frac{\left[\int_{x_*}^{\bar{t}} (xA_0 - A_1)\, dG(x)\right]^2}{1 - G(x_*)},$$

$$\hat{\Pi}_2 = \frac{1}{2A_2} \max_{\hat{x} \in [x_*, \bar{t}]} \left\{ \frac{\left[\int_{x_*}^{\hat{x}} (xA_0 - A_1)\, dG(x)\right]^2}{G(\hat{x}) - G(x_*)} + \frac{\left[\int_{\hat{x}}^{\bar{t}} (xA_0 - A_1)\, dG(x)\right]^2}{1 - G(\hat{x})} \right\}.$$

For any $\hat{x} \in [x_*, \bar{t}]$,

$$\begin{aligned}
2A_2\Pi_\infty &= \int_{x_*}^{\hat{x}} (xA_0 - A_1)^2 \, dG(x) + \int_{\hat{x}}^{\bar{t}} (xA_0 - A_1)^2 \, dG(x) \\
&= (\hat{x}A_0 - A_1)^2 (G(\hat{x}) - G(x_*)) - 2A_0 \int_{x_*}^{\hat{x}} \frac{G(x) - G(x_*)}{g(x)} (xA_0 - A_1) \, dG(x) \\
&\quad + (\hat{x}A_0 - A_1)^2 (1 - G(\hat{x})) + 2A_0 \int_{\hat{x}}^{\bar{t}} \frac{1 - G(x)}{g(x)} (xA_0 - A_1) \, dG(x) \\
&= (\hat{x}A_0 - A_1)^2 (1 - G(x_*)) \\
&\quad -2A_0 (G(\hat{x}) - G(x_*)) \int_{x_*}^{\hat{x}} \frac{G(x) - G(x_*)}{g(x)} (xA_0 - A_1) \frac{dG(x)}{G(\hat{x}) - G(x_*)} \\
&\quad +2A_0 (1 - G(\hat{x})) \int_{\hat{x}}^{\bar{t}} \frac{1 - G(x)}{g(x)} (xA_0 - A_1) \frac{dG(x)}{1 - G(\hat{x})}. \qquad (48)
\end{aligned}$$

Since both $(G(x) - G(x_*))/g(x)$ and $xA_0 - A_1$ are nondecreasing in x, they have nonnegative covariance when x is randomly drawn from $[x_*, \hat{x}]$. Thus, the integral in the second line of (48) is at least

$$\int_{x_*}^{\hat{x}} \frac{G(x) - G(x_*)}{g(x)} \frac{dG(x)}{G(\hat{x}) - G(x_*)} \cdot \int_{x_*}^{\hat{x}} (xA_0 - A_1) \frac{dG(x)}{G(\hat{x}) - G(x_*)}.$$

Since $(1 - G(x))/g(x)$ is nonincreasing and $xA_0 - A_1$ is nondecreasing in x, they have nonpositive covariance when x is randomly drawn from $[\hat{x}, \bar{t}]$. Thus, the integral in the third line of (48) is at most

$$\int_{\hat{x}}^{\bar{t}} \frac{1 - G(x)}{g(x)} \frac{dG(x)}{1 - G(\hat{x})} \cdot \int_{\hat{x}}^{\bar{t}} (xA_0 - A_1) \frac{dG(x)}{1 - G(\hat{x})}.$$

Therefore,

$$\begin{aligned}
2A_2\Pi_\infty &\leq (\hat{x}A_0 - A_1)^2 (1 - G(x_*)) \\
&\quad - \frac{2A_0}{G(\hat{x}) - G(x_*)} \int_{x_*}^{\hat{x}} \frac{G(x) - G(x_*)}{g(x)} dG(x) \cdot \int_{x_*}^{\hat{x}} (xA_0 - A_1) \, dG(x) \\
&\quad + \frac{2A_0}{1 - G(\hat{x})} \int_{\hat{x}}^{\bar{t}} \frac{1 - G(x)}{g(x)} dG(x) \cdot \int_{\hat{x}}^{\bar{t}} (xA_0 - A_1) \, dG(x). \qquad (49)
\end{aligned}$$

The first integral in the second line of (49) can be written as

$$\hat{x} (G(\hat{x}) - G(x_*)) - \int_{x_*}^{\hat{x}} x dG(x),$$

and the first integral in the third line of (49) can be written as

$$-\hat{x}\left(1-G(\hat{x})\right)+\int_{\hat{x}}^{\bar{t}} x dG(x).$$

Therefore,

$$\begin{aligned}
2A_2\Pi_\infty \;\leq\;& \left(\hat{x}A_0-A_1\right)^2\left(1-G(x_*)\right) \\
& -2A_0\left(\hat{x}-\frac{\int_{x_*}^{\hat{x}} x dG(x)}{G(\hat{x})-G(x_*)}\right)\cdot\int_{x_*}^{\hat{x}}\left(xA_0-A_1\right)dG(x) \\
& -2A_0\left(\hat{x}-\frac{\int_{\hat{x}}^{\bar{t}} x dG(x)}{1-G(\hat{x})}\right)\cdot\int_{\hat{x}}^{\bar{t}}\left(xA_0-A_1\right)dG(x). \qquad (50)
\end{aligned}$$

The second line of (50) can be written as

$$-2\left(\hat{x}A_0-A_1\right)\int_{x_*}^{\hat{x}}\left(xA_0-A_1\right)dG(x)+\frac{2\left[\int_{x_*}^{\hat{x}}\left(xA_0-A_1\right)dG(x)\right]^2}{G(\hat{x})-G(x_*)}$$

and the third line of (50) can be written as

$$-2\left(\hat{x}A_0-A_1\right)\int_{\hat{x}}^{\bar{t}}\left(xA_0-A_1\right)dG(x)+\frac{2\left[\int_{\hat{x}}^{\bar{t}}\left(xA_0-A_1\right)dG(x)\right]^2}{1-G(\hat{x})},$$

and

$$\frac{\left[\int_{x_*}^{\hat{x}}\left(xA_0-A_1\right)dG(x)\right]^2}{G(\hat{x})-G(x_*)}+\frac{\left[\int_{\hat{x}}^{\bar{t}}\left(xA_0-A_1\right)dG(x)\right]^2}{1-G(\hat{x})}\leq 2A_2\hat{\Pi}_2.$$

Therefore,

$$2A_2\Pi_\infty\leq\left(\hat{x}A_0-A_1\right)^2\left(1-G(x_*)\right)-2\left(\hat{x}A_0-A_1\right)\int_{x_*}^{\bar{t}}\left(xA_0-A_1\right)dG(x)+4A_2\hat{\Pi}_2.$$

Since the above holds for any $\hat{x}\in[x_*,\bar{t}]$, let us from now on take the unique $\hat{x}$ that solves

$$\hat{x}A_0-A_1=\frac{\int_{x_*}^{\bar{t}}\left(xA_0-A_1\right)dG(x)}{1-G(x_*)}.$$

Then,

$$2A_2\Pi_\infty\leq-\frac{\left[\int_{x_*}^{\bar{t}}\left(xA_0-A_1\right)dG(x)\right]^2}{1-G(x_*)}+4A_2\hat{\Pi}_2=-2A_2\hat{\Pi}_1+4A_2\hat{\Pi}_2.$$

■

Proof of Proposition 4. The objective function in (36) can be written as

$$\sum_{m=1}^{M}\left[J_m v(q_m)-c(q_m)\right]f_m+\sum_{m=1}^{M}\left(K_m-J_m\right)v(q_m)f_m.$$

The above second term can be rewritten as

$$\begin{aligned}
& \sum_{m=1}^{M}\left(R_m-\hat{R}_m\right)v(q_m)-\sum_{m=2}^{M}\left(R_{m-1}-\hat{R}_{m-1}\right)v(q_m)\\
= & \sum_{m=1}^{M}\left(R_m-\hat{R}_m\right)v(q_m)-\sum_{m=1}^{M-1}\left(R_m-\hat{R}_m\right)v(q_{m+1})\\
= & \left(R_M-\hat{R}_M\right)v(q_M)-\sum_{m=1}^{M-1}\left(R_m-\hat{R}_m\right)\left(v(q_{m+1})-v(q_m)\right)\\
= & -\sum_{m=1}^{M-1}\left(R_m-\hat{R}_m\right)\left(v(q_{m+1})-v(q_m)\right).
\end{aligned}$$

The last equality is from (41). Therefore, the objective function in (36) can be rewritten as

$$\sum_{m=1}^{M}\left[J_m v(q_m)-c(q_m)\right]f_m-\sum_{m=1}^{M-1}\left(R_m-\hat{R}_m\right)\left(v(q_{m+1})-v(q_m)\right). \tag{51}$$

Let $(q_1^*,\ldots,q_M^*)$ be the solution to (43). Thus, it maximizes the first term of (51) subject to the nonnegativity constraint and constraint (38). $(q_1^*,\ldots,q_M^*)$ is nondecreasing due to (39). Moreover, from (40), the second term is nonpositive for any nondecreasing vector $(q_1,\ldots,q_M)$.

Now, it remains to show that, when evaluated at $(q_1^*,\ldots,q_M^*)$, the second term of (51) attains zero, its highest possible value. Suppose not. Then there is some $\hat{m}\in\{1,\ldots,M-1\}$ such that $\hat{R}_{\hat{m}}<R_{\hat{m}}$ and $q_{\hat{m}+1}^*>q_{\hat{m}}^*$. We also have $J_{\hat{m}}=J_{\hat{m}+1}$ from (42). Let $\bar{m}$ be the largest index such that $q_{\bar{m}}^*=q_{\hat{m}+1}^*$. That is, the set of types served with quality $q_{\hat{m}+1}^*$ is $\{\hat{m}+1,\ldots,\bar{m}\}$. Since $(q_1^*,\ldots,q_M^*)$ solves (43), we must have

$$J_{\hat{m}}v(q_{\hat{m}})-c(q_{\hat{m}})=J_{\hat{m}+1}v(q_{\hat{m}+1})-c(q_{\hat{m}+1}), \tag{52}$$

for otherwise the objective of (43) can be raised by either letting type $\hat{m}$ be served by $q^*_{\hat{m}+1}$ or letting type $\hat{m}+1$ be served by $q^*_{\hat{m}}$ (since $J_{\hat{m}} = J_{\hat{m}+1}$). Moreover,

$$q^*_{\hat{m}+1} \in \arg\max_{q \geq 0} \sum_{m'=\hat{m}+1}^{\bar{m}} \left[J_{m'} v(q) - c(q) \right] f_{m'}, \tag{53}$$

for otherwise the objective of (43) can be raised by changing the quality serving $\{\hat{m}+1, \ldots, \bar{m}\}$. From (52) we can extend the coverage of $q^*_{\hat{m}+1}$ to $\{\hat{m}, \ldots, \bar{m}\}$ without changing the objective of (43). But then (53) does not hold anymore after replacing $\sum_{m'=\hat{m}+1}^{\bar{m}}$ by $\sum_{m'=\hat{m}}^{\bar{m}}$. Thus, the objective of (43) can be raised by changing the quality serving $\{\hat{m}, \ldots, \bar{m}\}$, which is a contradiction. ■

Proof of Theorem 5. The proof of the diminishing marginal benefit property is easily adapted from its counterpart in the continuous type model, and is omitted here. (Note that in the discrete type model, we only claim the weak inequality version.) We prove the discrete type version of the two-thirds result in the following.

Let $v(q) = A_0 q$ and $c(q) = A_1 q + \frac{A_2}{2} q^2$. Define $\hat{\Pi}_0 \equiv 0$, and for $n \geq 1$,

$$\hat{\Pi}_n \equiv \max_{q_{M_*}, \ldots, q_M \in \mathbb{R}_+} \sum_{m=1}^{M} \left[J_m v(q_m) - c(q_m) \right] f_m$$

subject to

$$(q_1, \ldots, q_M) \text{ takes at most } n \text{ values (not except zero).}$$

Clearly, $\hat{\Pi}_n \leq \Pi_n$ for every n; and $\hat{\Pi}_n = \Pi_n = \Pi_{M-M_*+1}$ whenever $n \geq M - M_* + 1$. The sequence $\left\{ \hat{\Pi}_n \right\}_{n=0}^{\infty}$ has the diminishing marginal benefit property.

Let $Y_m \equiv J_m A_0 - A_1$. The maximized values Π_{M-M_*+1}, $\hat{\Pi}_1$ and $\hat{\Pi}_2$ can be solved as

$$\Pi_{M-M_*+1} = \frac{1}{2A_2} \sum_{m=M_*}^{M} Y_m^2 f_m,$$

$$\hat{\Pi}_1 = \frac{1}{2A_2} \frac{\left[\sum_{m=M_*}^{M} Y_m f_m \right]^2}{1 - F_{M_*-1}},$$

$$\hat{\Pi}_2 = \frac{1}{2A_2} \max_{\hat{m} \in \{M_*, \ldots, M\}} \left\{ \frac{\left[\sum_{m=M_*}^{\hat{m}} Y_m f_m \right]^2}{F_{\hat{m}} - F_{M_*-1}} + \frac{\left[\sum_{m=\hat{m}+1}^{M} Y_m f_m \right]^2}{1 - F_{\hat{m}}} \right\}.$$

For any $\hat{m} \in \{M_*, \ldots, M\}$, $2A_2\Pi_{M-M_*+1}$ is equal to

$$
\begin{aligned}
& 2A_2\Pi_{M-M_*+1} \\
= \ & \sum_{m=M_*}^{\hat{m}} Y_m^2 \left[(F_m - F_{M_*-1}) - (F_{m-1} - F_{M_*-1})\right] + \sum_{m=\hat{m}+1}^{M} Y_m^2 \left[(1 - F_{m-1}) - (1 - F_m)\right] \\
= \ & \sum_{m=M_*}^{\hat{m}} Y_m^2 \left(F_m - F_{M_*-1}\right) - \sum_{m=M_*+1}^{\hat{m}} Y_m^2 \left(F_{m-1} - F_{M_*-1}\right) \\
& + \sum_{m=\hat{m}+1}^{M} Y_m^2 \left(1 - F_{m-1}\right) - \sum_{m=\hat{m}+1}^{M-1} Y_m^2 \left(1 - F_m\right) \\
= \ & Y_{\hat{m}}^2 \left(1 - F_{M_*-1}\right) - \sum_{m=M_*}^{\hat{m}} \left(Y_m^2 - Y_{m-1}^2\right) \left(F_{m-1} - F_{M_*-1}\right) \\
& + \sum_{m=\hat{m}+1}^{M} \left(Y_m^2 - Y_{m-1}^2\right) \left(1 - F_{m-1}\right). \qquad (54)
\end{aligned}
$$

Let $\alpha_m \equiv (Y_m - Y_{m-1})(F_{m-1} - F_{M_*-1})/f_m$ and $\beta_m \equiv (Y_{m+1} - Y_m)(F_m - F_{M_*-1})/f_m$. The second term of (54) is

$$
\begin{aligned}
& - \sum_{m=M_*}^{\hat{m}} Y_m \left(Y_m - Y_{m-1}\right) \left(F_{m-1} - F_{M_*-1}\right) - \sum_{m=M_*}^{\hat{m}} Y_{m-1} \left(Y_m - Y_{m-1}\right) \left(F_{m-1} - F_{M_*-1}\right) \\
= \ & - \left(F_{\hat{m}} - F_{M_*-1}\right) \sum_{m=M_*}^{\hat{m}} Y_m \alpha_m \frac{f_m}{F_{\hat{m}} - F_{M_*-1}} - \left(F_{\hat{m}-1} - F_{M_*-1}\right) \sum_{m=M_*}^{\hat{m}-1} Y_m \beta_m \frac{f_m}{F_{\hat{m}-1} - F_{M_*-1}}. \\
\leq \ & - \frac{\left(\sum_{m=M_*}^{\hat{m}} Y_m f_m\right)\left(\sum_{m=M_*}^{\hat{m}} \alpha_m f_m\right)}{F_{\hat{m}} - F_{M_*-1}} - \frac{\left(\sum_{m=M_*}^{\hat{m}-1} Y_m f_m\right)\left(\sum_{m=M_*}^{\hat{m}-1} \beta_m f_m\right)}{F_{\hat{m}-1} - F_{M_*-1}}. \qquad (55)
\end{aligned}
$$

The last inequality follows from the fact that Y_m, α_m, and β_m are all nondecreasing in m under our assumptions. After some simplifications, we can rewrite (55) as

$$
- \left(\sum_{m=M_*}^{\hat{m}} Y_m f_m\right) Y_{\hat{m}} - \left(\sum_{m=M_*}^{\hat{m}-1} Y_m f_m\right) Y_{\hat{m}} + \frac{\left(\sum_{m=M_*}^{\hat{m}} Y_m f_m\right)^2}{F_{\hat{m}} - F_{M_*-1}} + \frac{\left(\sum_{m=M_*}^{\hat{m}-1} Y_m f_m\right)^2}{F_{\hat{m}-1} - F_{M_*-1}}.
$$

Similarly, one can show under our assumptions that the third term of (54) is at most

$$
- \left(\sum_{m=\hat{m}+1}^{M} Y_m f_m\right) Y_{\hat{m}} - \left(\sum_{m=\hat{m}}^{M} Y_m f_m\right) Y_{\hat{m}} + \frac{\left(\sum_{m=\hat{m}+1}^{M} Y_m f_m\right)^2}{1 - F_{\hat{m}}} + \frac{\left(\sum_{m=\hat{m}}^{M} Y_m f_m\right)^2}{1 - F_{\hat{m}-1}}.
$$

It follows that

$$
\begin{aligned}
2A_2\Pi_{M-M_*+1} &\leq Y_{\hat{m}}^2(1-F_{M_*-1}) - 2\left(\sum_{m=M_*}^{M} Y_m f_m\right) Y_{\hat{m}} + 4A_2\hat{\Pi}_2 \\
&= \left[Y_{\hat{m}}^2 - 2\left(\frac{\sum_{m=M_*}^{M} Y_m f_m}{1-F_{M_*-1}}\right) Y_{\hat{m}} + \left(\frac{\sum_{m=M_*}^{M} Y_m f_m}{1-F_{M_*-1}}\right)^2\right](1-F_{M_*-1}) \\
&\quad -2A_2\hat{\Pi}_1 + 4A_2\hat{\Pi}_2
\end{aligned}
$$

$$
\begin{aligned}
2A_2\left(\Pi_{M-M_*+1} + \hat{\Pi}_1\right) &\leq \left(Y_{\hat{m}} - \frac{\sum_{m=M_*}^{M} Y_m f_m}{1-F_{M_*-1}}\right)^2 (1-F_{M_*-1}) + 4A_2\hat{\Pi}_2 \\
&= A_0^2\left(J_{\hat{m}} - \frac{\sum_{m=M_*}^{M} J_m f_m}{1-F_{M_*-1}}\right)^2 (1-F_{M_*-1}) + 4A_2\hat{\Pi}_2.
\end{aligned}
$$

Since the above is true for any $\hat{m} \in \{M_*, \ldots, M\}$, we minimize the right-hand side over $\hat{m}$ to obtain

$$2A_2\left(\Pi_{M-M_*+1} + \hat{\Pi}_1\right) \leq A_0^2(1-F_{M_*-1})\,\delta + 4A_2\hat{\Pi}_2$$

$$\Pi_{M-M_*+1} + \hat{\Pi}_1 \leq \frac{A_0^2(1-F_{M_*-1})\,\delta}{2A_2} + 2\hat{\Pi}_2$$

$$\Pi_{M-M_*+1} \leq 2(\hat{\Pi}_2 - \hat{\Pi}_1) + \hat{\Pi}_1 + \frac{A_0^2(1-F_{M_*-1})\,\delta}{2A_2}.$$

By the diminishing marginal benefit property, we have $\hat{\Pi}_2 - \hat{\Pi}_1 \leq \hat{\Pi}_1$, and hence

$$\Pi_{M-M_*+1} \leq 3\hat{\Pi}_1 + \frac{A_0^2(1-F_{M_*-1})\,\delta}{2A_2}$$

$$\hat{\Pi}_1 \geq \frac{1}{3}\left(\Pi_{M-M_*+1} - \frac{A_0^2(1-F_{M_*-1})\,\delta}{2A_2}\right).$$

It follows that

$$
\begin{aligned}
\Pi_{M-M_*+1} &\leq 2\hat{\Pi}_2 - \hat{\Pi}_1 + \frac{A_0^2(1-F_{M_*-1})\,\delta}{2A_2} \\
&\leq 2\hat{\Pi}_2 - \frac{1}{3}\left(\Pi_{M-M_*+1} - \frac{A_0^2(1-F_{M_*-1})\,\delta}{2A_2}\right) + \frac{A_0^2(1-F_{M_*-1})\,\delta}{2A_2}
\end{aligned}
$$

$$2\hat{\Pi}_2 \geq \frac{4}{3}\Pi_{M-M_*+1} - \frac{4}{3}\left(\frac{A_0^2(1-F_{M_*-1})\,\delta}{2A_2}\right)$$

$$\hat{\Pi}_2 \geq \frac{2}{3}\Pi_{M-M_*+1} - \frac{A_0^2\left(1-F_{M_*-1}\right)\delta}{3A_2}.$$

■

References

Bergemann, D., J. Shen, Y. Xu, and E. Yeh (2012): "Multi-dimensional Mechanism Design with Limited Information," in *Proceedings of the 13th ACM Conference on Electronic Commerce*, pp. 162–178. ACM.

Bergemann, D., J. Shen, Y. Xu, and E. M. Yeh (2011): "Mechanism Design with Limited Information: The Case of Nonlinear Pricing," 2nd International ICST Conference on Game Theory for Networks, Shanghai, 2011.

Blumrosen, L., N. Nisan, and I. Segal (2007): "Auctions with Severely Bounded Communication," *Journal of Artificial Intelligence Research*, 28, 233–266.

Chu, L. Y., and D. E. M. Sappington (2007): "Simple Cost-Sharing Contracts," *American Economic Review*, pp. 419–428.

Fudenberg, D., and J. Tirole (1991): *Game Theory*. Mit Press.

Guesnerie, R., and J.-J. Laffont (1984): "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(3), 329–369.

Kos, N. (2012): "Communication and Efficiency in Auctions," *Games and Economic Behavior*, 75(1), 233–249.

Maskin, E., and J. Riley (1984): "Monopoly with Incomplete Information," *RAND Journal of Economics*, 15(2), 171–196.

McAfee, R. P. (2002): "Coarse Matching," *Econometrica*, 70(5), 2025–2034.

Miravete, E. J. (2007): "The Limited Gains From Complex Tariffs," http://www.eugeniomiravete.com/papers/EJM-Gains.pdf.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(2), 175–208.

Mussa, M., and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18(2), 301–317.

Myerson, R. B. (1981): “Optimal Auction Design,” *Mathematics of Operations Research*, 6(1), 58–73.

Rockafellar, R. (1970): *Convex Analysis*. Princeton University Press.

Rogerson, W. P. (2003): “Simple Menus of Contracts in Cost-based Procurement and Regulation,” *American Economic Review*, 93(3), 919–926.

Wilson, R. (1989): “Efficient and Competitive Rationing,” *Econometrica*, 57(1), 1–40.

——— (1993): *Nonlinear Pricing*. Oxford University Press New York.

Wong, A. C. L. (2009): “The Choice of the Number of Varieties: Justifying Simple Mechanisms,” http://ssrn.com/abstract=1366228.

——— (2012): “Menu Implementability and Practical Pricing Schemes,” http://ssrn.com/abstract=1915752.