Multi-Product Price Optimization and Competition under the Nested Logit Model with Product-Differentiated Price Sensitivities

Guillermo Gallego
Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10027, gmg2@columbia.edu

Ruxian Wang
The Johns Hopkins Carey Business School, Baltimore, MD 21202, ruxian.wang@jhu.edu

We study firms that sell multiple differentiated substitutable products and customers whose purchase behavior follows a Nested Logit model, of which the Multinomial Logit model is a special case. Customers make purchasing decision sequentially under the Nested Logit model: they first select a nest of products and subsequently purchase a product within the selected nest. We consider the general Nested Logit model with product-differentiated price sensitivities and general nest coefficients. The problem is to price the products to maximize the expected total profit. We show that the *adjusted markup*, defined as price minus cost minus the reciprocal of price sensitivity, is constant for all products within a nest at optimality. This reduces the problem's dimension to a single variable per nest. We also show that each nest has an *adjusted nest-level markup* that is nest invariant, which further reduces the problem to a single variable optimization of a continuous function over a bounded interval. We provide conditions for this function to be uni-modal. We also use this result to simplify the oligopolistic multi-product price competition and characterize the Nash equilibrium. Furthermore, we extend to more general attraction functions including the linear utility and the multiplicative competitive interaction model, and show that the same technique can significantly simplify the multi-product pricing problems under the Nested Attraction model.

Key words: multi-product pricing; Attraction model; Nested Logit model; Multinomial Logit model; product-differentiated price sensitivity; substitutable products

History: first version in December 2011; this version on August 1, 2013

1. Introduction

Firms offering a menu of differentiated substitutable products face the problem of pricing them to maximize profits. This becomes more complicated with rapid technology development as new products are constantly introduced into the market and typically have a short life cycle. In this paper we are concerned with the problem of maximizing expected profits when customers follow a nested choice model where they first select a nest of products and then a product within the nest. The selection of nests and products depend on brand, product features, quality and price. The Nested Logit (NL) model and its special case the Multinomial Logit (MNL) model are among the most popular models to study purchase behavior of customers who face multiple substitutable products. The main contribution of this paper is to find very efficient solutions for a very general class of NL models and to explore the implications for oligopolistic competition and dynamic pricing.

The MNL model has received significant attention by researchers from economics, marketing, transportation science and operations management, and has motivated tremendous theoretical research and empirical validations in a large range of applications since it was first proposed by McFadden (1974), who was later awarded the 2000 Nobel Prize in Economics. The MNL model has been derived from an underlying random utility model, which is based on a probabilistic model of individual customer utility. Probabilistic choice can model customers with inherently unpredictable behavior that shows probabilistic tendency to prefer one alternative to another. When there is a random component in a customer' utility or a firm has only probabilistic information on the utility function of any given customer, the MNL model describes customers' purchase behavior very well.

The MNL model has been widely used as a model of customer choice, but it severely restricts the correlation patterns among choice alternatives and may behave badly under certain conditions (Williams and Ortuzar 1982), in particular when alternatives are correlated. This restrictive property is known as the *independence of irrelevant alternatives* (IIA) property (see Luce 1959). If the choice set contains alternatives that can be grouped such that alternatives within a group are more similar than alternatives outside the group, the MNL model is not realistic because adding new alternative reduces the probability of choosing similar alternatives more than dissimilar alternatives. This is often explained with the famous "red-bus/blue-bus" paradox (see Debreu 1952).

The NL model has been developed to relax the assumption of independence between all the alternatives, modeling the "similarity" between "nested" alternatives through correlation on utility components, thus allowing differential substitution patterns within and between nests. The NL model has become very useful on contexts where certain options are more similar than others, although the model lacks computational and theoretical simplicity. Williams (1977) first formulated the NL model and introduced structural conditions associated with its inclusive value parameters, which are necessary for the compatibility of the NL model with utility maximizing theory. He formally derived the NL model as a descriptive behavioral model completely coherent with basic micro-economic concepts. McFadden (1980) generated the NL model as a particular case of the generalized extreme value (GEV) discrete-choice model family and showed that it is numerically equivalent to Williams (1977). The NL model can also be derived from Gumbel marginal functions. Later on, Daganzo and Kusnic (1993) pointed out that although the conditional probability may be derived from a logit form, it is not necessary that the conditional error distribution be Gumbel. To keep consistent with micro-economic concepts, like random utility maximization, certain restrictions on model parameters that control the correlation among unobserved attributes have to be satisfied. One of the restrictions is that nest coefficients are required to lie within the unit interval.

Multi-product price optimization under the NL model and the MNL model has been the subject of active research since the models were first developed. Hanson and Martin (1996) show that the profit function of multiple differentiated substitutable products under the MNL model is not jointly concave with respect to the price vector. While the objective function is not concave in prices, it turns out to be concave with respect to the market share vector, which is in one-to-one correspondence with the price vector. To the best of our knowledge, this result is first established by Song and Xue (2007) and Dong et al. (2009) in the MNL model and by Li and Huh (2011) in the NL model. In all of their models, the price-sensitivity parameters are assumed identical for all the products within a nest and the nest coefficients are restricted to be in the unit interval. Empirical studies have shown that the product-differentiated price sensitivity may vary widely and the importance of allowing different price sensitivities in the MNL model (see Berry et al. 1995 and Erdem et al. 2002) has been recognized. Borsch-Supan (1990) points out that the restriction for nest coefficients in the unit interval leads too often rejection of the NL model. Unfortunately, the concavity with respect to the market share vector is lost when price-sensitivity parameters are product-differentiated or nest coefficients are greater than one as shown through an example in Appendix A.

Under the MNL model with identical price-sensitivity parameters, it has been observed that the *markup*, defined as price minus cost, is constant across all the products of the firm at optimality (see Anderson and de Palma 1992, Aydin and Ryan 2000, Hopp and Xu 2005 and Gallego and Stefanescu 2011). The profit function is uni-modal and its unique optimal solution can be found by solving the first order conditions (see Aydin and Porteus 2008, Akcay et al. 2010 and Gallego and Stefanescu 2011). In this paper, we consider the general NL model with product-differentiated price-sensitivity parameters and general nest coefficients. We show that the *adjusted markup*, which is defined as price minus cost minus the reciprocal of the price sensitivity, is constant across all the products in each nest at optimal (locally or globally) prices. When optimizing multiple nests of products, the *adjusted nest-level markup*, which is an adjusted average markup for all the products in the same nest, is also constant for each nest. By using this result, the multi-product and the multi-nest optimizations can be reduced to a single-dimensional problem of maximizing a continuous function over a bounded interval. We also provide mild conditions under which the single-dimensional problem is uni-modal, further simplifying the problem.

In a game-theoretic decentralized framework, the existence and uniqueness of a pure Nash equilibrium in a price competition model depend fundamentally on the demand functions as well as the cost structure. Milgrom and Roberts (1990) identify a rich class of demand functions, including the MNL model, and point out that the price competition game is supermodular, which guarantees the existence of a pure Nash equilibrium. Bernstein and Federgruen (2004) and Federgruen and Yang (2009) extend this result for a generalization of the MNL model referred to as the attraction model. Gallego et al. (2006) provide sufficient conditions for the existence and uniqueness of a Nash equilibrium under the cost structure that is increasing convex in the sale volume. Liu (2006), Cachon and Kok (2007) and Kok and Xu (2011) consider the NL model with identical price sensitivities for the products of the same firm and have characterized the Nash equilibrium. Moreover, Li and Huh (2011) study the same model with nest coefficients restricted in the unit interval and derive the unique equilibrium in a closed-form expression involving the Lambert W function (see Corless et al. 1996). In all these models, the price-sensitivity parameters for the products of the same firm are assumed identical. This paper considers competition under the general NL model and shows that the multi-product price competition is equivalent to a log-supermodular game in a single-dimensional strategy space.

The remainder of this paper is organized as follows. In Section 2, we consider the general Nested Logit model and show that the *adjusted markup* is constant across all the products of a nest. Moreover, the *adjusted nest-level markup* is also constant for each nest in a multi-nest optimization problem. In Section 3, we investigate the oligopolistic price competition problem, where each firm controls a nest of substitutable products. A Nash equilibrium exists for the general NL model and sufficient conditions for the uniqueness of the equilibrium are also provided. In Section 4, we consider an extension to other Nested Attraction models and conclude with a summary of our main results and useful management insights for application in business.

2. Price Optimization under the Nested Logit Model

Suppose that the products in consideration are substitutable, and they constitute n nests and nest i has m_i products. Customers' product selection behavior follows the NL model: they first select a nest and then choose a product within their selected nest. Let $Q_i(\mathbf{p})$ be the probability that a customer selects nest i at the upper stage; and let $q_{j|i}(\mathbf{p}_i)$ denote the probability that product j of nest i is selected at the lower stage, given that the customer selects nest i at the upper stage, where $\mathbf{p}_i = (p_{i1}, p_{i2}, \ldots, p_{im_i})$ is the price vector for the products in nest i, and $\mathbf{p} = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$ is the price matrix for all the products in the n nests. Following Williams (1977), McFadden (1980) and Greene (2007), $Q_i(\mathbf{p})$ and $q_{j|i}(\mathbf{p}_i)$ are defined as follows:

$$Q_i(\mathbf{p}) = \frac{\big(a_i(\mathbf{p}_i)\big)^{\gamma_i}}{1 + \sum_{l=1}^{n} \big(a_l(\mathbf{p}_l)\big)^{\gamma_l}}, \tag{1}$$

$$q_{j|i}(\mathbf{p}_i) = \frac{e^{\alpha_{ij} - \beta_{ij} p_{ij}}}{\sum_{s=1}^{m_i} e^{\alpha_{is} - \beta_{is} p_{is}}}, \tag{2}$$

where α_{is} can be interpreted as the "quality" of product s in nest i, $\beta_{is} \geq 0$ is the product-differentiated price sensitivity, $a_l(\mathbf{p}_l) = \sum_{s=1}^{m_l} e^{\alpha_{ls} - \beta_{ls} p_{ls}}$ represents the attractiveness of nest l (Anderson et al. 1992 show that the expected value of the maximum utility among all the products in nest l is equal to $\log(a_l(\mathbf{p}_l))$), and nest coefficient γ_i can be viewed as the degree of inter-nest heterogeneity. When $0 < \gamma_i < 1$, products are more similar within nest i than cross nests; when $\gamma_i = 1$, products in nest i have the same degree of similarity as products in other nests, and the NL model degenerates to the standard MNL model; when $\gamma_i > 1$, products are more similar to the ones in other nests. The probability that a customer will select product j of nest i is equal to

$$\pi_{ij}(\mathbf{p}) = Q_i(\mathbf{p}) \cdot q_{j|i}(\mathbf{p}_i). \tag{3}$$

Apparently, $\sum_{j=1}^{m_i} q_{j|i}(\mathbf{p}_i) = 1$ and $\sum_{j=1}^{m_i} \pi_{ij}(\mathbf{p}) = Q_i(\mathbf{p})$.

Without loss of generality, assume that the market size is normalized to 1. For the NL model, the monopolist's problem is to determine prices for all the products to maximize the expected total profit $R(\mathbf{p})$, which is expressed as follows

$$R(\mathbf{p}) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} (p_{ij} - c_{ij}) \pi_{ij}(\mathbf{p}). \tag{4}$$

The profit function $R(\mathbf{p})$ is high-dimensional and is hard directly to optimize. Hanson and Martin (1996) provide an example showing that $R(\mathbf{p})$ is not quasi-concave in $\mathbf{p}$ even under the MNL model, so other researchers, including Song and Xue (2007), Dong et al. (2009), take another approach. They express the profit as a function of the market-share vector and show that it is jointly concave with respect to market shares. Li and Huh (2011) extend to the NL model with nest coefficient $\gamma_i \leq 1$ and identical price-sensitivity parameters within each firm (may different across firms). However, the profit function may not be jointly concave when the price sensitivities are allowed to be product-differentiated, as we do in this paper, within each nest. Appendix A analyzes the problem and shows that the objective function fails to be jointly concave with respect to the market shares through an example. We will next take a different approach to consider the multi-product pricing problem under the general NL model.

Although the *markup* is no longer constant for the NL model with product-differentiated price-sensitivity parameters, we can obtain a similar result, which is crucial in simplifying the high-dimensional pricing problem.

Theorem 1 *The adjusted markup, defined as price minus cost minus the reciprocal of price sensitivity, is constant at optimality for all the products in each given nest.*

Aydin and Ryan (2000), Hopp and Xu (2005) and Gallego and Stefanescu (2011) observe that the *markup*, defined as price minus cost, is constant for all the products under the standard MNL model with identical price-sensitivity parameters. Li and Huh (2011) extend it to the NL model but the price sensitivities are still identical for all the products within the same nest although they may be different across nests.

Let θ_i denote the constant *adjusted markup* for all products in nest i, i.e.,

$$\theta_i = p_{ij} - c_{ij} - 1/\beta_{ij}. \tag{5}$$

For the sake of notation simplicity, let $Q_i(\boldsymbol{\theta})$ be the probability that a customer selects nest i at the upper stage, where $\boldsymbol{\theta}$ is the vector of *adjusted markups* for all the nests, i.e., $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$; and let $q_{j|i}(\theta_i)$ denote the probability that product j of nest i is selected at the lower stage, given that the customer selects nest i at the upper stage, where the prices in each nest satisfy the constant *adjusted markup* as shown in equation (5). Plugging equation (5) into the probabilities defined in equations (1) and (2) results in

$$\begin{aligned} Q_i(\boldsymbol{\theta}) &= \frac{\big(a_i(\theta_i)\big)^{\gamma_i}}{1 + \sum_{l=1}^{n} \big(a_l(\theta_l)\big)^{\gamma_l}}, \\ q_{j|i}(\theta_i) &= \frac{e^{\widetilde{\alpha}_{ij} - \beta_{ij}\theta_i}}{\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is} - \beta_{is}\theta_i}}, \end{aligned}$$

where $a_l(\theta_l) = \sum_{s=1}^{m_l} e^{\widetilde{\alpha}_{ls} - \beta_{ls}\theta_l}$ and $\widetilde{\alpha}_{ls} = \alpha_{ls} - \beta_{ls}c_{ls} - 1$ for each l and s. Then, the total probability that a customer will select product j of nest i is equal to $\pi_{ij}(\boldsymbol{\theta}) = Q_i(\boldsymbol{\theta}) \cdot q_{j|i}(\theta_i)$.

Note that the average profit of nest i can be expressed by $\sum_{j=1}^{m_i} (\theta_i + 1/\beta_{ij}) q_{j|i}(\theta_i) = \theta_i + w_i(\theta_i)$, where $w_i(\theta_i) = \sum_{j=1}^{m_i} 1/\beta_{ij} \cdot q_{j|i}(\theta_i)$. Then, the total expected profit corresponding to prices such that the *adjusted markup* is equal to θ_i for all the products in each nest i, can be rewritten as follows

$$R(\boldsymbol{\theta}) = \sum_{i=1}^{n} Q_i(\boldsymbol{\theta})\big(\theta_i + w_i(\theta_i)\big). \tag{6}$$

Then, high-dimensional price optimization problem is equivalent to determining the *adjusted markup* for each nest, which significantly reduces the dimension of the search space.

We remark that the optimal *adjusted markup* θ_i^* does not have to be positive in general, but it must be strictly positive when the nest coefficient is less than one for nest i (i.e., $\gamma_i \le 1$) because the total profit can also be expressed by $\theta_i^* + (1 - 1/\gamma_i) w_i(\theta_i^*)$ as shown below. Consequently, when $\gamma_i > 1$, it may be optimal to include "loss-leaders" as part of the optimal pricing strategy. More specifically, it may be optimal to include products with negative *adjusted markups* or even negative margins for the purpose of attracting attention to the nest.

Example 1. To demonstrate the "loss-leaders" phenomenon, we construct a simple example with a single nest containing two products. The parameters in the NL models are: $\boldsymbol{\alpha} = (0.8122, 0.4687)$, $\beta = (0.0039, 0.7637)$ and $\gamma = 1.4536$. The costs are $\mathbf{c} = (10, 10)$. At the upper stage of the NL model, the customer chooses an option between "purchase" and "non-purchase"; then she selects one of the two products if choosing the "purchase" option at the previous stage. The problem is to determine the prices for the two products to maximize the total profit assuming customers' purchase behavior follows the NL model.

By Theorem 1, the two-dimensional problem can be simplified to a single-dimensional problem of maximizing the total profit with respect to the *adjusted markup*. It is easy to obtain the unique optimal *adjusted markup* $\theta^* = -10.2063$, which is negative. Then, the optimal prices are $\mathbf{p} = \theta^* + \mathbf{c} + 1/\boldsymbol{\beta} = (256.2040, 1.1031)$. The total profit is equal to 33.7106. Note that the *markup* of the second product is negative, which is surprising at the first glance. In contrast, all the products will be sold at finite prices with positive margins under the MNL model.

Next, what if we do not offer the second product or equivalently set its price infinite? Consider the pricing problem only for the first product under the NL model. It is straightforward to find the optimal price 218.0770, which is lower than its optimal price when the second product is also offered at a finite price. The profit is 31.6803, which is 6% lower than that of offering the two products.

If the second product is offered with a negative margin, the attraction of the nest containing the two product is higher and more customers will select the "purchase" option at the upper stage of the NL model. If the second product is not offered, the nest attraction is lower and more customers will decide not to purchase. Although the second product contributes negative profit, the nest has higher attraction and more customers will select the "purchase" option at the upper stage. It results that the market share is higher, the total profit can be higher as well if the additional profit of the first product outperforms the loss from the sales of the second product. □

We now state our main condition for multi-product price optimization under the general NL model:

Condition 1 *For each nest i, $\gamma_i \geq 1$ or $\frac{\max_s \beta_{is}}{\min_s \beta_{is}} \leq \frac{1}{1-\gamma_i}$.*

Both the standard MNL model ($\gamma_i = 1$ for each nest i) and the NL model with identical price-sensitivity parameters and $\gamma_i < 1$ satisfy Condition 1. When $\gamma_i > 1$, it corresponds to the scenario where products are more similar cross nests; when $0 < \gamma_i < 1$, it refers to the case where products within the same nest are more similar, so the price coefficients of the products in the same nest should not vary too much and it is reasonable to require $\max_s \beta_{is}/\min_s \beta_{is} \leq 1/(1-\gamma_i)$. Condition 1 will be used later to establish important structural results.

We remark that when Condition 1 is satisfied, it requires either $\gamma_i \geq 1$ or $\max_s \beta_{is}/\min_s \beta_{is} \leq 1/(1-\gamma_i)$ for nest i. More specifically, it may happen that $\gamma_i \geq 1$ for nest i, and $\gamma_{i'} < 1$ and $\max_s \beta_{i's}/\min_s \beta_{i's} \leq 1/(1-\gamma_{i'})$ for another nest i'. Recall that $\theta_i + w_i(\theta_i)$ is the average markup for all the products in nest i, so we call $\theta_i + (1-1/\gamma_i)w_i(\theta_i)$ the *adjusted nest-level markup* for nest i.

Theorem 2 *Under Condition 1, the adjusted nest-level markup, defined as $\theta_i + (1-1/\gamma_i)w_i(\theta_i)$ for nest i, is constant for each nest.*

Then, the multi-product price optimization problem can be reduced to maximizing $R(\phi)$ with respect to the *adjusted nest-level markup* ϕ in a single-dimensional space, i.e.,

$$R(\phi) = \sum_{i=1}^{n} Q_i(\boldsymbol{\theta})\big(\theta_i + w_i(\theta_i)\big), \tag{7}$$

where for each nest i the *adjusted markup* θ_i is the uniquely determined by

$$\theta_i + (1-1/\gamma_i)w_i(\theta_i) = \phi. \tag{8}$$

Profit $R(\phi)$ is an implicit function expressed in terms of θ_i, as there is a one-to-one mapping between θ_i and ϕ for each i according to equation (8) under Condition 1

$$\frac{\partial}{\partial\theta_i}\Big(\theta_i + (1-1/\gamma_i)w_i(\theta_i)\Big) = \big(1-(1-\gamma_i)w_i(\theta_i)v_i(\theta_i)\big)/\gamma_i > 0,$$

where $v_i(\theta_i) = \sum_{j=1}^{m_i} \beta_{ij} \cdot q_{j|i}(\theta_i)$. The inequality holds because $w_i(\theta_i)v_i(\theta_i) \leq \max_s \beta_{is}/\min_s \beta_{is}$ as shown in Lemma 1 in the Appendix.

Corollary 1 *Under Condition 1, $R(\phi)$ is strictly uni-modal in ϕ. Moreover, $R(\phi)$ takes its maximum value at its unique fixed point, denoted by ϕ^*.*

Then, the optimal *adjusted markup* θ_i^* for nest i is the unique corresponding solution to equation (8). Thus, the optimal price for product j of nest i is equal to$p_{ij}^* = \theta_i^* + c_i + 1/\beta_{ij}$.

The multi-product price optimization can also be transformed to an optimization problem with respect to the total market share. Let $R(\rho)$ be the maximum achievable total expected profit given that the aggregate market share is equal to ρ, i.e., $\sum_{i=1}^n Q_i(\mathbf{p}) = \rho$.

$$\begin{aligned} R(\rho) := \max_{\mathbf{p}} & \sum_{i=1}^n \sum_{j=1}^{m_i} (p_{ij} - c_{ij}) \pi_{ij}(\mathbf{p}) \\ s.t., & \sum_{i=1}^n Q_i(\mathbf{p}) = \rho. \end{aligned} \tag{9}$$

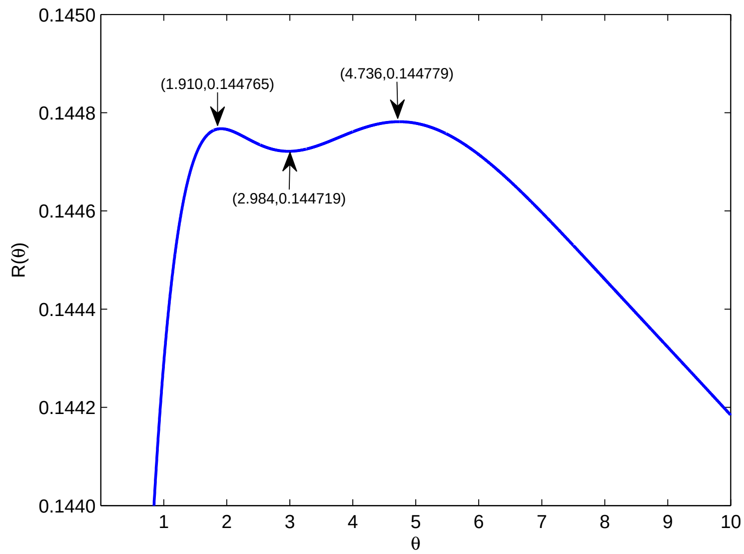
Although in general the total profit is not jointly concave with respect to the market-share matrix, it has a nice structure in the aggregate market share.

Corollary 2 *Under Condition 1, $R(\rho)$ is strictly concave with respect to the aggregate market share ρ.*

If Condition 1 is satisfied, the profit function $R(\phi)$ is uni-modal in ϕ and $R(\rho)$ is concave in ρ, so the first order condition is sufficient to determine the optimal prices, which can be easily found by several well known algorithms for uni-modal or concave functions, e.g., the binary search method and golden section search algorithm; if Condition 1 is not satisfied, $R(\phi)$ may not be uni-modal as illustrated in the following example.

Example 2. Consider the NL model with a single nest containing five products. The parameters for the NL model are $\tilde{\boldsymbol{\alpha}} = (-4.8150, -6.2897, -6.1610, -6.1906, -6.7078)$, $\boldsymbol{\beta} = (0.6720, 1.1249, 1.0247, 0.7968, 0.0150)$ and the nest coefficient $\gamma = 0.9150$. Note that $\frac{\max_s \beta_{is}}{\min_s \beta_{is}} = 1.1249/0.0150 = 74.99 > 1/(1-\gamma_i) = 11.76$, which violates Condition 1. By Theorem 1, the five products should be priced such that the *adjusted markup* is constant. We use $R(\theta)$ to represent the total profit of the five products corresponding to the *adjusted markup* θ. Figure 1 clearly shows that $R(\theta)$ is not uni-modal with respect to θ and there are three stationary points in the interval $(1, 10)$: $(1.910, 0.144765)$, $(2.984, 0.144719)$ and $(4.736, 0.144779)$. Observe that the relative difference of profits is very small: $(0.144779 - 0.144719)/0.144719 = 0.04\%$. This suggests that $R(\theta)$ is very flat at the peak and any solution to the first order condition gives a good approximation. $\square$

Figure 1 **Non-unimodality of $R(\boldsymbol{\theta})$**



3. Oligopolistic Competition

We will next consider oligopolistic price competition where each firm controls a nest of multiple products. This is consistent with an NL model where customers first select a brand and then a product within a brand. The oligopolistic price (Bertrand) competition with multiple products under the standard MNL model has been widely examined and the existence and uniqueness of Nash equilibrium have been established (see Gallego et al. 2006, Allon et al. 2011). Liu (2006) and Li and Huh (2011) have studied price competition under the NL model with identical price sensitivities for all the products of each firm. However, their approach cannot easily extend to the general NL model with product-differentiated price sensitivities. To the best of our knowledge, our paper is the first to study oligopolistic competition with multiple products under the general NL model with product dependent price-sensitivities and arbitrary nest coefficients.

In the price competition game, the expected profit for firm i is

$$\textbf{Game I:} \qquad R_i(\mathbf{p}_i, \mathbf{p}_{-i}) = \sum_{k=1}^{m_i} (p_{ik} - c_{ik}) \cdot \pi_{ik}(\mathbf{p}_i, \mathbf{p}_{-i})$$

where $\mathbf{p}_i$ is the price vector of firm i, i.e., $\mathbf{p}_i = (p_{i1}, p_{i2}, \ldots, p_{im_i})$, and $\mathbf{p}_{-i}$ is the price vectors for all other firms except firm i, i.e., $\mathbf{p}_{-i} = (\mathbf{p}_1, \ldots, \mathbf{p}_{i-1}, \mathbf{p}_{i+1}, \ldots, \mathbf{p}_n)$.

By Theorem 1, the multi-product pricing problem for each firm can be reduced to a problem with a single decision variable of the *adjusted markup* as follows

$$\textbf{Game II:} \qquad R_i(\theta_i, \theta_{-i}) = Q_i(\boldsymbol{\theta})(\theta_i + w_i(\theta_i)).$$

We remark that $R_i(\theta_i, \theta_{-i})$ is log-separable. Because the profit function $R_i(\theta_i, \theta_{-i})$ is uni-modal with respect to θ_i under Condition 1, then it is also quasi-concave in θ_i because quasi-concavity and uni-modality are equivalent in a single-dimensional space. The quasi-concavity can guarantee the existence of the Nash equilibrium (see, e.g., Nash 1951 and Anderson et al. 1992), but there are some stronger results without requiring Condition 1 because of the special structure of the NL model.

Theorem 3 *(a)* **Game I** *is equivalent to* **Game II**, *i.e., they have the same equilibria.*

(b) **Game II** *is strictly* log-*supermodular; the equilibrium set is a nonempty complete lattice and, therefore, has the componentwise largest and smallest elements, denoted by* $\overline{\boldsymbol{\theta}}^*$ *and* $\underline{\boldsymbol{\theta}}^*$ *respectively. Furthermore, the largest equilibrium* $\overline{\boldsymbol{\theta}}^*$ *is preferred by all the firms.*

The multi-product price competition game has been reduced to an equivalent game with a single *adjusted markup* for each firm. The existence of Nash equilibrium has been guaranteed and the largest one is a Pareto improvement among the equilibrium set, preferred by each firm.

To examine the uniqueness of the Nash equilibrium, we will concentrate on **Game II**, which is equivalent to **Game I** by Theorem 3. We first consider a special case: the symmetric game, and leave the discussion for the general case in the Appendix. Suppose that there are n firms and that all the parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)$ in the NL model including the cost vector $\mathbf{c}$ are the same for each firm i. Note that the firm index is omitted in this symmetric game but the parameters in the NL model including costs may be product-dependent within a firm. Some properties about the equilibrium set can be further derived.

Condition 2 $\gamma \geq \frac{n}{n-1}$ *or* $\frac{\max_s \beta_s}{\min_s \beta_s} \leq \frac{1}{1-\frac{n-1}{n}\cdot\gamma}$.

We remark that Condition 2 is a bit stronger than Condition 1 and they are closer for larger n (they coincide when n goes infinite).

Theorem 4 *(a) Only symmetric equilibria exist for the symmetric game discussed above.*

(b) The equilibrium is unique under Condition 2.

Under Condition 2, the unique equilibrium is the solution to the equation

$$\theta + \left(1 - \frac{1}{\gamma_i(1 - Q_i(\theta))}\right) w_i(\theta) = 0, \tag{10}$$

where $Q_i(\theta)$ is the probability to select each nest i given that the *adjusted markup* is θ for each firm, i.e.,

$$Q_i(\theta) = \frac{\left(\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is} - \beta_{is}\theta_i}\right)^{\gamma_i}}{1 + n\left(\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is} - \beta_{is}\theta_i}\right)^{\gamma_i}}.$$

It refers to the proof of Theorem 4 in the Appendix for details.

4. Extension: Nested Attraction Model

The attraction models have received increasing attention in the marketing literature, as it specifies that a market share of a firm is equal to its attraction divided by the total attraction of all the firms in the market, including the non-purchase option, where a firm's attraction is a function of the values of its marketing instruments, e.g., brand value, advertising, product features and variety, etc. As an extension, we will consider the generalized Nested Attraction models, of which the NL model discussed above is a special case.

In this section, we extend to the general Nested Attraction model:

$$\begin{aligned} Q_i(\mathbf{p}) &= \frac{\big(a_i(\mathbf{p}_i)\big)^{\gamma_i}}{1+\sum_{l=1}^{n}\big(a_l(\mathbf{p}_l)\big)^{\gamma_l}}, \\ q_{j|i}(\mathbf{p}_i) &= \frac{a_{ij}(p_{ij})}{\sum_{s=1}^{m_i} a_{is}(p_{is})}, \\ \pi_{ij}(\mathbf{p}) &= Q_i(\mathbf{p})\cdot q_{j|i}(\mathbf{p}_i), \end{aligned}$$

where $a_{is}(p_{is})$ is the attractiveness of product s of nest i at price p_{is} and is twice-differentiable with respect to p_{is}, and $a_l(\mathbf{p}_l)=\sum_{s=1}^{m_l} a_{ls}(p_{ls})$ is the total attractiveness of nest l. Note that $a_{is}(p_{is}) = e^{\alpha_{is}-\beta_{is}p_{is}}$ for the NL model discussed above; for the linear model $a_{is}(p_{is})=\alpha_{is}-\beta_{is}p_{is}$, $\alpha_{is},\beta_{is}>0$; for the multiplicative competitive interaction (MCI) model, $a_{is}(p_{is})=\alpha_{is}p_{is}^{-\beta_{is}}$, $\alpha_{is}>0,\beta_{is}>1$.

Condition 3 *(a)* $a'_{ij}(p_{ij})\le 0$, $2(a'_{ij}(p_{ij}))^2 > a_{ij}(p_{ij})a''_{ij}(p_{ij})\ \forall j,\ p_{ij}$.
(b) That $a'_{ij}(p_{ij})=0$ *implies that* $(p_{ij}-c_{ij})a_{ij}(p_{ij})=0$.

That $a'_{ij}(p_{ij})\le 0$ says that each product's attractiveness is decreasing in its price; that $2(a'_{ij}(p_{ij}))^2 > a_{ij}(p_{ij})a''_{ij}(p_{ij})$ can be implied by a stronger condition that $a_{ij}(p_{ij})$ is log-concave or concave in p_{ij}. Condition 3(b) requires that $a_{ij}(p_{ij})$ converges to zero at a faster rate than linear functions when $a'_{ij}(p_{ij})$ converges to zero. In other words, when $a'_{ij}(p_{ij})=0$, product j of nest i does not contribute any profit so it can be eliminated from the profit function.

Theorem 5 *Under Condition 3, the prices at optimality satisfy that* $(p_{ij}-c_{ij})+a_{ij}(p_{ij})/a'_{ij}(p_{ij})$ *is constant for each product* j *in nest* i.

It is straightforward to verify that the MNL model, the Nested linear attraction model and the Nested MCI model all satisfy Condition 3. The Corollary follows immediately for the special cases.

Corollary 3 *The following quantities are constant at optimal prices for the Nested linear attraction model and the Nested MCI model, respectively:*

$$2p_{ij}-c_{ij}-\frac{\alpha_{ij}}{\beta_{ij}},\quad (1-\frac{1}{\beta_{ij}})p_{ij}-c_{ij}.$$

Denote the constant quantity by θ_i for each nest i, i.e., $\theta_i = (p_{ij} - c_{ij}) + a_{ij}(p_{ij})/a'_{ij}(p_{ij})$ for each product j of nest i. There is a one-to-one mapping between θ_i and p_{ij} for each product j under Condition 3 because it holds that

$$\frac{\partial}{\partial p_{ij}}\left((p_{ij} - c_{ij}) + \frac{a_{ij}(p_{ij})}{a'_{ij}(p_{ij})}\right) = \frac{2(a'_{ij}(p_{ij}))^2 - a_{ij}(p_{ij})a''_{ij}(p_{ij})}{a'_{ij}(p_{ij})} < 0.$$

Then, the multi-product pricing problem under the general Nested Attraction model can be reduced to maximizing the total profit $R(\boldsymbol{\theta})$ with respect to θ_i for each nest i, where $R(\boldsymbol{\theta})$ under the Nested Attraction model is defined as follows

$$R(\boldsymbol{\theta}) = \sum_{i=1}^{n}\sum_{k=1}^{m_i}(p_{ik} - c_{ik})\pi_{ik}(\mathbf{p}), \tag{11}$$

where price p_{ij} for each product j of each nest i is uniquely determined by

$$(p_{ij} - c_{ij}) + \frac{a_{ij}(p_{ij})}{a'_{ij}(p_{ij})} = \theta_i. \tag{12}$$

Furthermore, the results in price competition can also be obtained under the Nested Attraction model.

5. Concluding Remark

Discrete choice modeling has become a popular vehicle to study purchase behavior of customers who face multiple substitutable products. The MNL discrete choice model has been well studied and widely used in marketing, economics, transportation science and operations management, but it suffers the IIA property, which limits its application and acceptance, especially in the scenarios with correlated products. The generalized NL model with a two-stage structure can alleviate the IIA property. Empirical studies have shown that the NL model works well in the environment with differentiated substitutable products.

This paper considers price optimization and competition with multiple substitutable products under the general NL model with product-differentiated price-sensitivity parameters and general nest coefficients. Our analysis shows that the *adjusted markup*, defined as price minus cost minus the reciprocal of price sensitivity, is constant for all products of each nest at optimality. In addition, the optimal *adjusted nest-level markup* is also constant for each nest. By using this result, the multi-product and multi-nest optimization problems is reduced to a single-dimensional maximization of a continuous function over a a bounded interval. Mild conditions are provided for this function to be uni-modal. We also use this result to characterize the Nash equilibrium for the price competition under the NL model.

We also study the general Nested Attraction model, of which the NL model and the MNL model are special cases, and show how it can be transformed to an optimization problem in a single-dimensional space. The two-stage model can alleviate the IIA property and derive high acceptance and wide use in practice. In the future, the research and practice on customers' selection behavior with three or even higher stages may attract more attention because it may be closer to the rationality of the decision process. Another future research direction may consider the heterogeneity of customers and investigate the discrete choice model in the context with multiple heterogenous market segments.

Appendix A: Non-concavity of Market Share Transformation

In this section, we will first express the profit in terms of the market-share vector and then show it is no long jointly concave under the NL model with product-differentiated price sensitivities.

From equation (1),

$$\frac{Q_i}{1-\sum_{l=1}^{n} Q_l} = \left(\sum_{s=1}^{m_i} e^{\alpha_{is}-\beta_{is}p_{is}}\right)^{\gamma_i}.$$

Combining with equation (2) results in

$$e^{\alpha_{ij}-\beta_{ij}p_{ij}} = \frac{\pi_{ij}}{Q_i}\cdot\left(\frac{Q_i}{1-Q_i}\right)^{\frac{1}{\gamma_i}}.$$

Then, p_{ij} can be expressed in terms of $\boldsymbol{\pi}_i := (\pi_{i1}, \pi_{i2}, \ldots, \pi_{im_i})$ as follows

$$p_{ij}(\boldsymbol{\pi}_i) = \frac{1}{\beta_{ij}}\big(\log Q_i - \log \pi_{ij}\big) + \frac{1}{\beta_{ij}\gamma_i}\big(\log(1-Q_i) - \log Q_i\big) + \frac{\alpha_{ij}}{\beta_{ij}}. \tag{13}$$

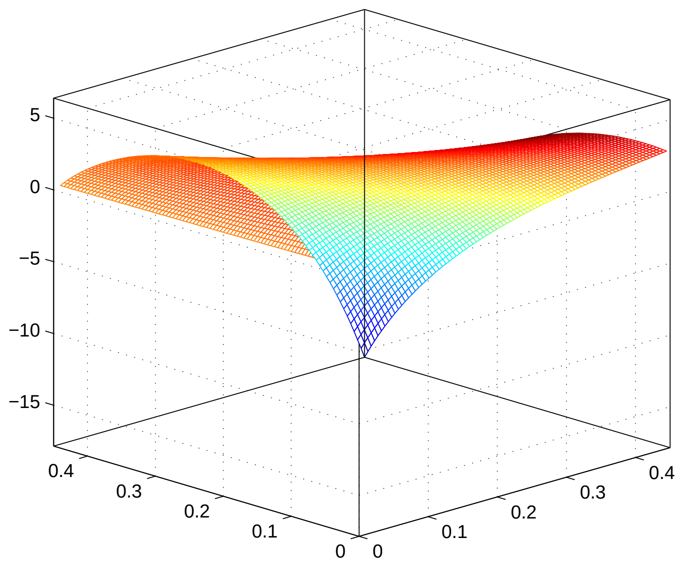
The total profit can be rewritten as a function of the market-share matrix:

$$R(\boldsymbol{\pi}) = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\left(\frac{1}{\beta_{ij}}\big(\log Q_i - \log \pi_{ij}\big) + \frac{1}{\beta_{ij}\gamma_i}\big(\log(1-Q_i) - \log Q_i\big) - \widetilde{c}_{ij}\right)\cdot \pi_{ij}, \tag{14}$$

where $Q_i = \sum_{s=1}^{m_i} \pi_{is}$ and $\widetilde{c}_{ik} = c_{ik} - \frac{\alpha_{ik}}{\beta_{ik}}$.

Consider an NL model with a single nest consisting of two products with product-differentiated price coefficients $\boldsymbol{\beta}_i = (0.9, 0.1)$ and nest coefficient $\gamma_i = 0.1$. Figure 2 demonstrates that $R(\boldsymbol{\pi})$ is not jointly concave with respect to the market-share vector $\boldsymbol{\pi}$.

Figure 2 Non-concavity of $R(\pi)$ w.r.t. the Market-share Vector



Appendix B: Proofs

Proof of Theorem 1. Consider the first order condition of $R(\mathbf{p})$ with respect to price p_{ij} in a given nest i:

$$\frac{\partial R(\mathbf{p})}{\partial p_{ij}} = \pi_{ij}(\mathbf{p}) \cdot \left[1 - \beta_{ij}(p_{ij} - c_{ij}) + \beta_{ij}(1-\gamma_i)\sum_{s=1}^{m_i}(p_{is} - c_{is})q_{s|i}(\mathbf{p}_i) + \beta_{ij}\gamma_i \sum_{l=1}^{n}\sum_{s=1}^{m_i}(p_{is} - c_{is})\pi_{is}(\mathbf{p})\right] = 0.$$

Roots of the above FOC can be found by either setting $\pi_{ij}(\mathbf{p}) = 0$, which requires $p_{ij} = \infty$ or letting the inner term of the square bracket equal 0, which is equivalent to

$$p_{ij} - c_{ij} - \frac{1}{\beta_{ij}} = (1-\gamma_i)\sum_{s=1}^{m_i}(p_{is} - c_{is})q_{s|i}(\mathbf{p}_i) + \gamma_i \sum_{l=1}^{n}\sum_{s=1}^{m_l}(p_{ls} - c_{ls})\pi_{ls}(\mathbf{p}). \tag{15}$$

The right hand side (RHS) is independent product index j in nest i, so $p_{ij} - c_{ij} - \frac{1}{\beta_{ij}}$, the so-called *adjusted markup*, is constant for all the products with finite prices in nest i. We will next show that *all* the products should be charged finite prices such that the *adjusted markup* is constant.

Let F_i be the set of products in nest i offered at finite prices. Denote the *adjusted markup* of each product $k \in F_i$ by θ_i. Suppose the prices of all products in other nests are fixed. Note that the total attractiveness of all other nests but nest i is $\sum_{l\neq i}\left(a_l(\mathbf{p}_l)\right)^{\gamma_l} = \sum_{l\neq i}\left(\sum_{s=1}^{m_l} e^{\alpha_{ls}-\beta_{ls}p_{ls}}\right)^{\gamma_l}$. Let ρ be the total market share, which can be expressed as follows

$$\rho = \sum_{k=1}^{n} Q_k(\theta_i, \mathbf{p}_{-i}) = \frac{\sum_{l\neq i}\left(a_l(\mathbf{p}_l)\right)^{\gamma_l} + \left(\sum_{s\in F_i} e^{\alpha_{is}-\beta_{is}\theta_i}\right)^{\gamma_i}}{1 + \sum_{l\neq i}\left(a_l(\mathbf{p}_l)\right)^{\gamma_l} + \left(\sum_{s\in F_i} e^{\alpha_{is}-\beta_{is}\theta_i}\right)^{\gamma_i}}.$$

Then, the total attractiveness of nest i can be expressed in terms of ρ and a_j, $j \neq i$,

$$\left(\sum_{s\in F_i} e^{\alpha_{is}-\beta_{is}\theta_i}\right)^{\gamma_i} = \frac{1}{1-\rho} - \left(1 + \sum_{l\neq i}\left(a_l(\mathbf{p}_l)\right)^{\gamma_l}\right). \tag{16}$$

Given that the total market share is ρ and the offered product set in nest i is F_i, the *adjusted markup* is the unique solution to equation (16), denoted by $\theta_i^{F_i}$. The total profit can be expressed as follows

$$\begin{aligned}
R^{F_i}(\rho) &= Q_i(\theta_i^{F_i}, \mathbf{p}_{-i})(\theta_i^{F_i} + w_i(\theta_i^{F_i})) + \sum_{j\neq i} Q_j(\theta_i^{F_i}, \mathbf{p}_{-i}) \sum_{s=1}^{m_j} (p_{js} - c_{js}) q_{s|j}(\mathbf{p}_j) \\
&= \left(1 - (1-\rho)\Big(1 + \sum_{l\neq i} \big(a_l(\mathbf{p}_l)\big)^{\gamma_l}\Big)\right)\left(\theta_i^{F_i} + \frac{\sum_{s\in F_i} e^{\alpha_{is} - \beta_{is}\theta_i^{F_i}}/\beta_{is}}{\big(1/(1-\rho) - (1 + \sum_{l\neq i}\big(a_l(\mathbf{p}_l)\big)^{\gamma_l})\big)^{1/\gamma_i}}\right) \\
&\quad + (1-\rho)\sum_{l\neq i}\big(a_l(\mathbf{p}_l)\big)^{\gamma_l}\sum_{s=1}^{m_j}(p_{js} - c_{js}) q_{s|j}(\mathbf{p}_j).
\end{aligned}$$

Let $H^{F_i}(\theta_i)$ be the RHS of the above equation with $\theta_i^{F_i}$ replaced by a free variable θ_i, i.e.,

$$\begin{aligned}
H^{F_i}(\theta_i) &= \left(1 - (1-\rho)\Big(1 + \sum_{l\neq i} \big(a_l(\mathbf{p}_l)\big)^{\gamma_l}\Big)\right)\left(\theta_i^{F_i} + \frac{\sum_{s\in F_i} e^{\alpha_{is} - \beta_{is}\theta_i^{F_i}}/\beta_{is}}{\big(1/(1-\rho) - (1 + \sum_{l\neq i}\big(a_l(\mathbf{p}_l)\big)^{\gamma_l})\big)^{1/\gamma_i}}\right) \\
&\quad + (1-\rho)\sum_{l\neq i}\big(a_l(\mathbf{p}_l)\big)^{\gamma_l}\sum_{s=1}^{m_j}(p_{js} - c_{js}) q_{s|j}(\mathbf{p}_j). \qquad (17)
\end{aligned}$$

Considering its derivative at $\theta_i = \theta_i^{F_i}$ results in

$$\left.\frac{\partial H^{F_i}(\theta_i)}{\partial \theta_i}\right|_{\theta_i = \theta_i^{F_i}} = \left(1 - (1-\rho)\Big(1 + \sum_{l\neq i} \big(a_l(\mathbf{p}_l)\big)^{\gamma_l}\Big)\right)\left.\left(1 - \frac{\sum_{s\in F_i} e^{\alpha_{is} - \beta_{is}\theta_i}}{\big(1/(1-\rho) - (1 + \sum_{l\neq i}(a_l(\mathbf{p}_l))^{\gamma_l})\big)^{1/\gamma_i}}\right)\right|_{\theta_i = \theta_i^{F_i}} = 0.$$

The second equality holds because of equation (16). Recalling that $H^{F_i}(\theta_i)$ is convex in θ_i, $R^{F_i}(\rho)$ is the minimum of $H^{F_i}(\theta_i)$ with respect to θ_i, i.e., $R^{F_i}(\rho) = \min_{\theta_i} H^{F_i}(\theta_i) = H^{F_i}(\theta_i^{F_i})$.

Suppose that another product z is added to set F_i and denote $F_i^+ := F_i \cup \{z\}$. We will next show that $R^{F_i^+}(\rho) > R^{F_i}(\rho)$ for any $0 < \rho < 1$. Similarly, we have $R^{F_i^+}(\rho) = \min_{\theta_i} H^{F_i^+}(\theta_i) = H^{F_i^+}(\theta_i^{F_i^+})$, where $H^{F_i^+}(\theta_i)$ is defined in function (25) and $\theta_i^{F_i^+}$ is the unique solution to equation (16) corresponding to offer set F_i^+.

It is apparent that $H^{F_i^+}(\theta_i) > H^{F_i}(\theta_i)$ for any θ_i. Then,

$$R^{F_i^+}(\rho) = H^{F_i^+}(\theta_i^{F_i^+}) > H^{F_i}(\theta_i^{F_i^+}) > H^{F_i}(\theta_i^{F_i}) = R^{F_i}(\rho).$$

The second inequality holds because $\theta_i^{F_i}$ is the minimizer of $H^{F_i}(\theta_i)$ with respect to θ_i. Therefore, $R^{F_i}(\rho)$ is strictly increasing in F_i for any $0 < \rho < 1$ and it is optimal to offer all the products at prices such that the *adjusted markup* is constant in each nest. $\square$

Lemma 1 *Define $w_i(\theta_i) = \sum_{k=1}^{m_i} \frac{1}{\beta_{ik}} \cdot q_{k|i}(\theta_i)$ and $v_i(\theta_i) = \sum_{k=1}^{m_i} \beta_{ik} \cdot q_{k|i}(\theta_i)$. The following monotonic properties hold.*

(a) $w_i(\theta_i)$ is increasing in θ_i and $\frac{1}{\max_s \beta_{is}} \le w_i(\theta_i) \le \frac{1}{\min_s \beta_{is}}$.

(b) $v_i(\theta_i)$ is decreasing in θ_i and $\min_s \beta_{is} \le v_i(\theta_i) \le \max_s \beta_{is}$. Furthermore, $w_i(\theta_i)v_i(\theta_i) \ge 1$ for any θ_i, and all the inequalities become equalities when β_{is} is identical for all $s \in F_i$.

Proof of Lemma 1. (a) Consider the first order derivative of $w_i(\theta_i)$. Then

$$\frac{\partial w_i(\theta_i)}{\partial \theta_i} = -1 + \frac{\sum_{k=1}^{m_i} e^{\widetilde{\alpha}_{ik}-\beta_{ik}\theta_i}/\beta_{ik} \sum_{s=1}^{m_i} \beta_{is} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}}{\left(\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^2}.$$

That $\frac{\partial w_i(\theta_i)}{\partial \theta_i} \geq 0$ can be shown by Cauchy-Schwarz inequality that is $\sum_{i=1}^m x_i \sum_{i=1}^m y_i \geq \left(\sum_{i=1}^m \sqrt{x_i y_i}\right)^2$ for any $x_i, y_i \geq 0$. Because

$$\sum_{k=1}^{m_i} \frac{e^{\widetilde{\alpha}_{ik}-\beta_{ik}\theta_i}}{\max_s \beta_{is}} \leq \sum_{k=1}^{m_i} \frac{e^{\widetilde{\alpha}_{ik}-\beta_{ik}\theta_i}}{\beta_{ik}} \leq \sum_{k=1}^{m_i} \frac{e^{\widetilde{\alpha}_{ik}-\beta_{ik}\theta_i}}{\min_s \beta_{is}}$$

then, $\frac{1}{\max_s \beta_{is}} \leq w_i(\theta_i) \leq \frac{1}{\min_s \beta_{is}}$ and the inequalities become equalities when β_{is} is constant for all $s = 1, \ldots, m_i$.

(b) Consider the first order derivative of $v_i(\theta_i)$. Then

$$\frac{\partial v_i(\theta_i)}{\partial \theta_i} = \frac{-\sum_{k=1}^{m_i} \beta_{ik}^2 e^{\widetilde{\alpha}_{ik}-\beta_{ik}\theta_i} \cdot \sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i} + \left(\sum_{s=1}^{m_i} \beta_{is} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^2}{\left(\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^2}.$$

It can be shown that $\frac{\partial v_i(\theta_i)}{\partial \theta_i} \leq 0$ by a similar argument to part (a).

$$w_i(\theta_i) v_i(\theta_i) = \frac{\sum_{k=1}^{m_i} e^{\widetilde{\alpha}_{ik}-\beta_{ik}\theta_i}/\beta_{ik} \sum_{s=1}^{m_i} \beta_{is} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}}{\left(\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^2} = \frac{\partial w_i(\theta_i)}{\partial \theta_i} + 1 \geq 1$$

The inequality holds because of part (a). □

Proof of Theorem 2. For the general NL model with product-differentiated price-sensitivity parameters, the FOC of the total profit $R(\boldsymbol{\theta})$ is

$$\frac{\partial R(\boldsymbol{\theta})}{\partial \theta_i} = \gamma_i Q_i(\boldsymbol{\theta}) v_i(\theta_i) \left[\sum_{l=1}^n Q_l(\boldsymbol{\theta})(\theta_l + w_l(\theta_l)) - \left(\theta_i + (1 - \frac{1}{\gamma_i}) w_i(\theta_i)\right)\right] = 0.$$

Again, because $v_i(\theta_i) \geq \min_s \beta_{is} > 0$, the solutions to the above FOC can be found by either letting $Q_i(\boldsymbol{\theta}) = 0$, which requires $\theta_i = \infty$ or setting the inner term of the square bracket equal to zero, which is equivalent to

$$\theta_i + (1 - \frac{1}{\gamma_i}) w_i(\theta_i) = \sum_{l=1}^n Q_l(\boldsymbol{\theta})(\theta_l + w_l(\theta_l)). \tag{18}$$

The RHS of equation (18) is independent of nest index i, so $\theta_i + (1 - \frac{1}{\gamma_i}) w_i(\theta_i)$ is invariant for each nest i, denoted by ϕ. We will next show that each nest should be charged finite *adjusted markup* and the so-called *adjusted nest-level markup* $\theta_i + (1 - \frac{1}{\gamma_i}) w_i(\theta_i)$ is constant all the nests.

Let E be the set of nests whose *adjusted markup* satisfies equation (18). The total market share can be expressed as follows

$$\rho = \sum_{i \in E} Q_i(\boldsymbol{\theta}) = \frac{\sum_{i \in E} \left(\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^{\gamma_i}}{1 + \sum_{i \in E} \left(\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^{\gamma_i}}, \tag{19}$$

Then,

$$\sum_{i\in E}\left(\sum_{s=1}^{m_i} e^{\tilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^{\gamma_i} = \frac{\rho}{1-\rho}, \tag{20}$$

$$\theta_i + (1-\frac{1}{\gamma_i})w_i(\theta_i) = \phi,\ \forall i\in E. \tag{21}$$

The solution to (20) and (21) is unique, which will be shown later, denoted by ϕ^E and $\boldsymbol{\theta}^E = (\theta_1^E,\dots,\theta_n^E)$. The total profit can be expressed

$$\begin{aligned} R^E(\rho) &= \sum_{i\in E} Q_i(\boldsymbol{\theta}^E)\left(\theta_i^E + w_i(\theta_i^E)\right) = \sum_{i\in E} Q_i(\boldsymbol{\theta}^E)\left(\phi^E + \frac{w_i(\theta_i^E)}{\gamma_i}\right) \\ &= \rho\phi^E + \sum_{i\in E}\frac{\left(\sum_{s=1}^{m_i} e^{\tilde{\alpha}_{is}-\beta_{is}\theta_i^E}\right)^{\gamma_i}\cdot w_i(\theta_i^E)}{\gamma_i/(1-\rho)}, \end{aligned}$$

Define a new function $H^E(\phi)$ with ϕ^E replaced by a free variable ϕ but keeping the relation between ϕ and θ_i as shown in (21), i.e.,

$$H^E(\phi) = \rho\phi + \sum_{i\in E}\frac{\left(\sum_{s=1}^{m_i} e^{\tilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^{\gamma_i}\cdot w_i(\theta_i)}{\gamma_i/(1-\rho)}. \tag{22}$$

where θ_i satisfies (21) for each $i\in E$. We will next show that $H^E(\phi)$ is convex in ϕ under Condition 1 for each i.

$$\frac{\partial H^E(\phi)}{\partial\phi} = \rho + \sum_{i\in E}\frac{\partial G^E(\phi)/\partial\theta_i}{\partial\phi/\partial\theta_i} = \rho - (1-\rho)\sum_{i\in E}\left(\sum_{s=1}^{m_i} e^{\tilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^{\gamma_i},$$

where $G^E(\phi) = \sum_{i\in E}\frac{\left(\sum_{s=1}^{m_i} e^{\tilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^{\gamma_i}\cdot w_i(\theta_i)}{\gamma_i/(1-\rho)}$. The second equality holds because

$$\begin{aligned} \frac{\partial G^E(\phi)}{\partial\theta_i} &= -\frac{1-\rho}{\gamma_i}\cdot\left(1-(1-\gamma_i)w_i(\theta_i)v_i(\theta_i)\right)\left(\sum_{s=1}^{m_i} e^{\tilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^{\gamma_i}, \\ \frac{\partial\phi}{\partial\theta_i} &= \frac{1}{\gamma_i}\left(1-(1-\gamma_i)w_i(\theta_i)v_i(\theta_i)\right). \end{aligned}$$

Moreover,

$$\left.\frac{\partial H^E(\phi)}{\partial\phi}\right|_{\phi=\phi^E} = 0.$$

The above equality holds because of equation (20). The second order derivative of $H^E(\phi)$ is

$$\frac{\partial^2 H^E(\phi)}{\partial\phi^2} = \sum_{i\in E}\frac{\frac{\partial}{\partial\theta_i}(\partial H^E(\phi)/\partial\phi)}{\partial\phi/\partial\theta_i} = \sum_{i\in E}\frac{\gamma_i^2(1-\rho)v_i(\theta_i)\left(\sum_{s=1}^{m_i} e^{\tilde{\alpha}_{is}-\beta_{is}\theta_i}\right)^{\gamma_i}}{1-(1-\gamma_i)w_i(\theta_i)v_i(\theta_i)} > 0.$$

The inequality holds because $1-(1-\gamma_i)w_i(\theta_i)v_i(\theta_i) > 0$ under Condition 1.

Thus, $H^E(\phi)$ is convex in ϕ, and the solution to (20) and (21) is unique. Moreover, $R^E(\rho)$ is the minimum of $H^E(\phi)$ with respect to ϕ, i.e., $R^E(\rho) = \min_\phi H^E(\phi)$. Let E^+ be the new nest set if the *adjusted markup* of another nest satisfies equation (18). We will next show that $R^{E^+}(\rho) > R^E(\rho)$ for any $0<\rho<1$.

It is apparent that $H^{E^+}(\phi) > H^E(\phi)$ for any ϕ, where $H^{E^+}(\phi)$ is the function defined in (22) corresponding to offer set E^+. Then,

$$R^{E^+}(\rho) = H^{E^+}(\phi^{E^+}) > H^E(\phi^{E^+}) > H^E(\phi^E) = R^E(\rho).$$

The second inequality holds because ϕ^E is the minimizer of $H^E(\phi)$. Therefore, $R^E(\rho)$ is strictly increasing in E for any $0 < \rho < 1$ and it is optimal to offer all the products at prices such that the *adjusted nest-level markup* is constant for all the nests. $\square$

Proof of Coronary 1. Consider the FOC of $R(\phi)$,

$$\frac{\partial R(\phi)}{\partial \phi} = \sum_{i=1}^{n} \frac{\partial R(\boldsymbol{\theta})/\partial \theta_i}{\partial \phi/\partial \theta_i} = (R(\phi) - \phi) \sum_{i=1}^{n} \frac{\gamma_i^2 Q_i(\boldsymbol{\theta}) v_i(\theta_i)}{1 - (1-\gamma_i) w_i(\theta_i) v_i(\theta_i)} = 0, \tag{23}$$

where θ_i is the solution to equation (21). Because $\sum_{i=1}^{n} \frac{\gamma_i^2 Q_i(\boldsymbol{\theta}) v_i(\theta_i)}{1-(1-\gamma_i) w_i(\theta_i) v_i(\theta_i)} > 0$ under Condition 1 for each i, then, $R(\phi)$ is increasing (decreasing) in ϕ if and only if $R(\phi) \geq (\leq) \phi$.

(i) Case I: there is only one solution to equation (23), denoted by ϕ^*. Apparently $R(\phi)$ is increasing in ϕ for $\phi \leq \phi^*$ and is decreasing in ϕ for $\phi > \phi^*$.

(ii) Case II: there are multiple solutions to equation (23). Suppose that there are two consecutive solutions $\phi_1 = R(\phi_1) < \phi_2 = R(\phi_2)$ and there is no solution to equation (23) between ϕ_1 and ϕ_2. It must hold that $R(\phi) < R(\phi_2)$ for any $\phi_1 < \phi < \phi_2$; otherwise, there must be another solution to equation (23) between ϕ_1 and ϕ_2, which contradicts that ϕ_1 and ϕ_2 are two consecutive solutions. We claim that $R(\phi)$ is increasing in ϕ for $\phi \in [\phi_1, \phi_2]$. Assume there are two points $\phi_1 < \phi_1' < \phi_2' < \phi_2$ such that $R(\phi_1') > R(\phi_2')$. Then, there must be a solution to equation (23) between ϕ_1' and ϕ_2, which also contradicts that ϕ_1 and ϕ_2 are two consecutive solutions. Thus, $R(\phi)$ is increasing between any two solutions to equation (23) and $R(\phi)$ may be decreasing after the largest solution. Therefore, $R(\phi)$ is unimodual with respect to ϕ under Condition 1. $\square$

Proof of Coronary 2. Let $\rho(\phi) = \sum_{i=1}^{n} Q_i(\boldsymbol{\theta})$, where θ_i is the solution to $\theta_i + (1 - \frac{1}{\gamma_i}) w_i(\theta_i) = \phi$, for each $i = 1, 2, \ldots, n$. Then,

$$\begin{aligned}
\frac{\partial \rho(\phi)}{\partial \phi} &= \sum_{i=1}^{n} \frac{\partial \rho(\phi)/\partial \theta_i}{\partial \phi/\partial \theta_i} = -Q_0(\boldsymbol{\theta}) \sum_{i=1}^{n} \frac{\gamma_i^2 Q_i(\boldsymbol{\theta}) v_i(\theta_i)}{1 - (1-\gamma_i) w_i(\theta_i) v_i(\theta_i)}, \\
\frac{\partial R(\phi)}{\partial \rho} &= \frac{\partial R(\phi)/\partial \phi}{\partial \rho(\phi)/\partial \phi} = -\frac{R(\boldsymbol{\theta}) - \phi}{1-\rho}.
\end{aligned}$$

We can easily show that $R(\rho)$ is unimodual in ρ by a similar argument to part (c). Moreover, we consider the second order derivative under Condition 1 for all i,

$$\frac{\partial^2 R(\rho)}{\partial \rho^2} = -\frac{\partial}{\partial \rho} \cdot \frac{R(\boldsymbol{\theta}) - \phi}{1-\rho} = -\frac{R(\boldsymbol{\theta}) - \phi}{(1-\rho)^2} + \frac{1}{1-\rho} \cdot \frac{\frac{\partial R(\boldsymbol{\theta})}{\partial \phi} - 1}{\frac{\partial \rho}{\partial \phi}}$$

$$= -\frac{1}{(1-\rho)^2 \sum_{i=1}^{n} \frac{\gamma_i^2 Q_i(\boldsymbol{\theta}) v_i(\theta_i)}{1-(1-\gamma_i) w_i(\theta_i) v_i(\theta_i)}} < 0.$$

The last equality hold because $(1-\gamma_i)w_i(\theta_i)v_i(\theta_i) < 1$ for all θ_i and each i under Condition 1. Therefore, $R(\rho)$ is concave in ρ under Condition 1. $\square$

Proof of Theorem 3. (a) Suppose that $(\mathbf{p}_i^*, \mathbf{p}_{-i}^*)$ is an equilibrium of **Game I**. From Theorem 1, the *adjusted markup* is constant for all the products of each firm, i.e., $p_{ij} - c_{ij} - \frac{1}{\beta_{ij}}$ is constant for all j, denoted by θ_i^*. We will argue that $(\theta_i^*, \theta_{-i}^*)$ must be the equilibrium of **Game II**. If firm i is better-off to deviate to $\hat{\theta}_i$, then firm i will also be better-off to deviate to $\hat{\mathbf{p}}_i$ in **Game I**, where $\widehat{\mathbf{p}}_i = (\widehat{p}_{i1}, \ldots, \widehat{p}_{im_i})$ and $\widehat{p}_{ij} = \widehat{\theta}_i + c_{ij} + \frac{1}{\beta_{ij}}$. It contradicts that $(\mathbf{p}_i^*, \mathbf{p}_{-i}^*)$ is an equilibrium of **Game I**.

Suppose that $(\theta_i^*, \theta_{-i}^*)$ is an equilibrium of **Game II**. We will argue that $(\mathbf{p}_i^*, \mathbf{p}_{-i}^*)$ is an equilibrium of **Game I**, where $p_{ij} = \theta_i^* + c_{ij} + \frac{1}{\beta_{ij}}$ for each j. If firm i is better-off to deviate to $\widehat{\mathbf{p}}_i := (\widehat{p}_{i1}, \widehat{p}_{i2}, \ldots, \widehat{p}_{im_i})$ in **Game I**, $\widehat{p}_{ij} - c_{ij} - \frac{1}{\beta_{ij}}$ must be constant for each j by Theorem 1, denoted by $\widehat{\theta}_i$. Then, firm i must be better-off to deviate to $\hat{\theta}_i$ in **Game II**, which contradicts that $(\theta_i^*, \theta_{-i}^*)$ is an equilibrium of **Game II**.

(b) Consider the derivatives of $\log R_i(\theta_i, \theta_{-i})$:

$$\begin{aligned}
\frac{\partial \log R_i(\theta_i, \theta_{-i})}{\partial \theta_i} &= -\gamma_i(1 - Q_i(\theta_i, \theta_{-i}))v_i(\theta_i) + \frac{w_i(\theta_i)v_i(\theta_i)}{\theta_i + w_i(\theta_i)}, \\
\frac{\partial \log R_i(\theta_i, \theta_{-i})}{\partial \theta_j} &= \gamma_j Q_j(\theta_j, \theta_{-j}) v_j(\theta_j) \geq 0, \ \forall j \neq i, \\
\frac{\partial^2 \log R_i(\theta_i, \theta_{-i})}{\partial \theta_i \partial \theta_j} &= \gamma_i \gamma_j Q_i(\theta_i, \theta_{-i}) Q_j(\theta_j, \theta_{-j}) v_i(\theta_i) v_j(\theta_j) \geq 0, \ \forall j \neq i.
\end{aligned}$$

Then, **Game II** is a log-supermodular game. Note that the strategy space for each firm is the real line. From Topkis (1998) and Vives (2001), the equilibrium set is a nonempty complete lattice and, therefore, has the componentwise largest element $\overline{\boldsymbol{\theta}}^*$ and smallest element $\underline{\boldsymbol{\theta}}^*$, respectively.

For any equilibrium $\boldsymbol{\theta}^*$, it holds that $\overline{\boldsymbol{\theta}}^* \geq \boldsymbol{\theta}^* \geq \underline{\boldsymbol{\theta}}^*$ and

$$\log R_i(\theta_i^*, \theta_{-i}^*) \leq \log R_i(\theta_i^*, \overline{\theta}_{-i}^*) \leq \log R_i(\overline{\theta}_i^*, \overline{\theta}_{-i}^*).$$

The first inequality holds because $\partial \log R_i(\theta_i, \theta_{-i})/\partial \theta_j \geq 0$; the second inequality holds because $(\overline{\theta}_i^*, \overline{\theta}_{-i}^*)$ is a Nash equilibrium. Because logarithm is a monotonic increasing transformation, then

$$R_i(\theta_i^*, \theta_{-i}^*) \leq R_i(\theta_i^*, \overline{\theta}_{-i}^*) \leq R_i(\overline{\theta}_i^*, \overline{\theta}_{-i}^*).$$

Therefore, the largest equilibrium $\overline{\boldsymbol{\theta}}^*$ is preferred by all the firms. $\square$

Proof of Theorem 4 (a) Suppose that there exists an asymmetric equilibrium, denoted by $(\theta_1^*, \theta_2^*, \theta_3^*, \ldots, \theta_n^*)$. Suppose that θ_1^* is the largest and θ_2^* is the smallest without loss of generality, then $\theta_1^* > \theta_2^*$. Because the game is symmetric, $(\theta_2^*, \theta_1^*, \theta_3^*, \ldots, \theta_n^*)$ is also an equilibrium. In other words, the best strategies for firm 1 are θ_1^* and θ_2^* respectively corresponding to other firms' strategies $(\theta_2^*, \theta_3^*, \ldots, \theta_n^*)$ and $(\theta_1^*, \theta_3^*, \ldots, \theta_n^*)$. Since the game is strictly supermodular and $(\theta_2^*, \theta_3^*, \ldots, \theta_n^*) < (\theta_1^*, \theta_3^*, \ldots, \theta_n^*)$, then $\theta_1^* \le \theta_2^*$, which contradicts that $\theta_1^* > \theta_2^*$.

(b) Consider the derivative of $R_i(\theta_i, \theta_{-i})$ with respect to θ_i as follows:

$$\frac{\partial R_i(\theta_i,\theta_{-i})}{\partial \theta_i} = -\gamma_i Q_i(\theta_i,\theta_{-i})\big(1-Q_i(\theta_i,\theta_{-i})\big)v_i(\theta_i)\left[\theta_i + \left(1 - \frac{1}{\gamma_i(1-Q_i(\theta_i,n))}\right) w_i(\theta_i)\right].$$

Define $Y(\theta_i)$ as follows:

$$Y(\theta_i) = \theta_i + \left(1 - \frac{1}{\gamma_i(1-Q_i(\theta_i,n))}\right) w_i(\theta_i).$$

Its derivative can be expressed by

$$\frac{\partial Y(\theta_i)}{\partial \theta_i} = \frac{1}{\gamma_i(1-Q_i(\theta_i,n))} \cdot \left(1 - w_i(\theta_i)v_i(\theta_i)\left(1-\gamma_i+\gamma_i\frac{(n-1)(Q_i(\theta_i,n))^2}{1-Q_i(\theta_i,n)}\right)\right),$$

Clearly, $Q_i(\theta_i, n)$, the probability to select nest i if all firms charge the same *adjusted markup* θ_i, is less than $1/n$, i.e., $Q_i(\theta_i, n) < \frac{1}{n}$. Then,

$$\frac{\partial Y(\theta_i)}{\partial \theta_i} > \frac{1}{\gamma_i(1-Q_i(\theta_i,n))} \cdot \left(1 - w_i(\theta_i)v_i(\theta_i)\left(1-\frac{n-1}{n}\gamma_i\right)\right).$$

Therefore,

(i) If $\gamma_i \ge \frac{n}{n-1}$, then $\frac{\partial Y(\theta_i)}{\partial \theta_i} > 0$ for any θ_i.

(ii) If $0 < \gamma_i < \frac{n}{n-1}$ and $\frac{\max_s \beta_{is}}{\min_s \beta_{is}} \le \frac{1}{1-\frac{n-1}{n}\cdot\gamma_i}$, we claim that $w_i(\theta_i)v_i(\theta_i)\left(1-\frac{n-1}{n}\gamma_i\right) < 1$. If there are more than one products with different price coefficients, $w_i(\theta_i)v_i(\theta_i) < \frac{\max_s \beta_{is}}{\min_s \beta_{is}}$; otherwise $w_i(\theta_i)v_i(\theta_i) = 1$ for any θ_i. In both cases, $w_i(\theta_i)v_i(\theta_i)\left(1-\frac{n-1}{n}\gamma_i\right) < 1$.

Thus, $\frac{\partial Y(\theta_i)}{\partial \theta_i} > 0$ for any θ_i under Condition 2. Then, $Y(\theta_i)$ is strictly increasing from negative to positive as θ_i goes from $-\infty$ to ∞. Hence, there exists a unique solution to the equation $Y(\theta_i) = 0$ and it is also the unique equilibrium to the symmetric game. $\square$

Proof of Theorem 5. Similar to the proof of Theorem 1, consider the FOC for the profit $R(\mathbf{p})$ under the Nested Attraction model:

$$\frac{\partial R(\mathbf{p})}{\partial p_{ij}} = \frac{\pi_{ij}(\mathbf{p})a'_{ij}(p_{ij})}{\beta_{ij}a_{ij}(p_{ij})} \cdot \left[(p_{ij}-c_{ij}) + \frac{a_{ij}(p_{ij})}{a'_{ij}(p_{ij})} - (1-\gamma_i)\sum_{s=1}^{m_i}(p_{is}-c_{is})q_{s|i}(\mathbf{p}_i) - \gamma_i \sum_{l=1}^{n}\sum_{s=1}^{m_i}(p_{is}-c_{is})\pi_{is}(\mathbf{p})\right] = 0.$$

The above equation is satisfied when either $\frac{\beta_{ij}\pi_{ij}(\mathbf{p})a'_{ij}(p_{ij})}{a_{ij}(p_{ij})}=0$, which requires $a'_{ij}(p_{ij})=0$, or the inner term of the square bracket is equal to zero, i.e.,

$$(p_{ij}-c_{ij})+\frac{a_{ij}(p_{ij})}{a'_{ij}(p_{ij})}=(1-\gamma_i)\sum_{s=1}^{m_i}(p_{is}-c_{is})q_{s|i}(\mathbf{p}_i)+\gamma_i\sum_{l=1}^{n}\sum_{s=1}^{m_i}(p_{is}-c_{is})\pi_{is}(\mathbf{p}). \tag{24}$$

We remark that the RHS of (24) is independent of product index j, so $(p_{ij}-c_{ij})+\frac{a_{ij}(p_{ij})}{a'_{ij}(p_{ij})}$ is invariant in products of set F_i in nest i, where F_i is the offer set of products in nest i at prices such that equation (24) is satisfied. Denote $\theta_i^{F_i}=(p_{ij}-c_{ij})+a_{ij}(p_{ij})/a'_{ij}(p_{ij})$ for each product j in set F_i.

Suppose that the prices of all other nests are fixed. Consider the price optimization problem with the market share constraint as follows:

$$\begin{aligned}R^{F_i}(\rho,\mathbf{p}_{-i}):=&\max_{p_{ij}<\infty,\forall j\in F_i}\textstyle\sum_{l\neq i}\sum_{s=1}^{m_l}(p_{ls}-c_{ls})\pi_{ls}(\mathbf{p}_i,\mathbf{p}_{-i})+\sum_{k\in F_i}(p_{ik}-c_{ik})\pi_{ik}(\mathbf{p}_i,\mathbf{p}_{-i})\\ &s.t.,\ Q_i(\mathbf{p}_i,\mathbf{p}_{-i})=\rho.\end{aligned}$$

Let $a_l=\sum_{s=1}^{m_l}a_{ls}(p_{ls})$ for each $l\neq i$ for the ease of notation. After some algebra, the total market share constraint $Q_i(\mathbf{p}_i,\mathbf{p}_{-i})=\rho$ results in

$$\sum_{s\in F_i}a_{is}(p_{is})=\left(\frac{\rho}{1-\rho}-\sum_{l\neq i}a_l^{\gamma_l}\right)^{1/\gamma_i}.$$

Then, $R^{F_i}(\rho,\mathbf{p}_{-i})$ can be rewritten as follows

$$R^{F_i}(\rho,\mathbf{p}_{-i})=\left(\rho-(1-\rho)\sum_{l\neq i}a_l^{\gamma_l}\right)\cdot\left(\theta_i^{F_i}-\big(\rho/(1-\rho)-\sum_{l\neq i}a_l^{\gamma_l}\big)^{-1/\gamma_i}\sum_{s\in F_i}\frac{(a_{is}(p_{is}))^2}{a'_{is}(p_{is})}\right)+\sum_{l\neq i}(1-\rho)a_l^{\gamma_l}\sum_{s=1}^{m_l}(p_{ls}-c_{ls})q_{s|l}(\mathbf{p}_l),$$

where p_{is} is uniquely determined by $(p_{is}-c_{is})+a_{is}(p_{is})/a'_{is}(p_{is})=\theta_i^{F_i}$ for each $s\in F_i$.

Define function $H^{F_i}(\theta_i)$ as follows

$$H^{F_i}(\theta_i)=\theta_i-\big(\rho/(1-\rho)-\sum_{l\neq i}a_l^{\gamma_l}\big)^{-1/\gamma_i}\sum_{s\in F_i}\frac{(a_{is}(p_{is}))^2}{a'_{is}(p_{is})}, \tag{25}$$

where p_{is} is uniquely determined by $(p_{is}-c_{is})+a_{is}(p_{is})/a'_{is}(p_{is})=\theta_i$ for each $s\in F_i$. Consider the first order derivative for function $H^{F_i}(\theta_i)$:

$$\begin{aligned}\frac{\partial H^{F_i}(\theta_i)}{\partial\theta_i}&=1-\big(\rho/(1-\rho)-\sum_{l\neq i}a_l^{\gamma_l}\big)^{-1/\gamma_i}\sum_{s\in F_i}\frac{\frac{\partial}{\partial p_{is}}\frac{(a_{is}(p_{is}))^2}{a'_{is}(p_{is})}}{\partial\theta_i/\partial p_{is}}\\ &=1-\big(\rho/(1-\rho)-\sum_{l\neq i}a_l^{\gamma_l}\big)^{-1/\gamma_i}\sum_{s\in F_i}\frac{\frac{2a_{is}(p_{is})(a'_{is}(p_{is}))^2-(a_{is}(p_{is}))^2a''_{is}(p_{is})}{(a'_{is}(p_{is}))^2}}{1+\frac{(a'_{is}(p_{is}))^2-a_{is}(p_{is})a''_{is}(p_{is})}{(a'_{is}(p_{is}))^2}}\\ &=1-\big(\rho/(1-\rho)-\sum_{l\neq i}a_l^{\gamma_l}\big)^{-1/\gamma_i}\sum_{s\in F_i}a_{is}(p_{is}).\end{aligned}$$

Then, it follows that

$$\left.\frac{\partial H^{F_i}(\theta_i)}{\partial\theta_i}\right|_{\theta_i=\theta_i^{F_i}}=0.$$

The equality holds because $\sum_{s\in F_i} a_{is}(p_{is}) = \left(\frac{\rho}{1-\rho} - \sum_{l\neq i} a_l^{\gamma_l}\right)^{1/\gamma_i}$. The second order derivative of $H^{F_i}(\eta_i)$ is

$$\frac{\partial^2 H^{F_i}(\theta_i)}{\partial \theta_i^2} = \sum_{s\in F_i} \frac{\frac{\partial}{\partial p_{is}}(\partial H^{F_i}(\theta_i)/\partial\theta_i)}{\frac{\partial \theta_i}{\partial p_{is}}} = -\left(\rho/(1-\rho) - \sum_{l\neq i} a_l^{\gamma_l}\right)^{-1/\gamma_i} \sum_{s\in F_i} \frac{(a'_{is}(p_{is}))^3}{2(a'_{is}(p_{is}))^2 - a_{is}(p_{is})a''_{is}(p_{is})} \geq 0.$$

The inequality holds because $a'_{is}(p_{is}) \leq 0$ and $2(a'_{is}(p_{is}))^2 - a_{is}(p_{is})a''_{is}(p_{is}) > 0$ from Condition 3. Thus, $H^{F_i}(\theta_i)$ is convex in θ_i and $R^{F_i}(\rho, \mathbf{p}_{-i}) = \min_{\theta_i} H^{F_i}(\theta_i) = H^{F_i}(\theta_i^{F_i})$.

Suppose that another product z is added to set F_i and denote $F_i^+ := F_i \cup \{z\}$. Similarly, we have $R^{F_i^+}(\rho) = \min_{\theta_i} H^{F_i^+}(\theta_i) = H^{F_i^+}(\theta_i^{F_i^+})$, where $H^{F_i^+}(\theta_i)$ is defined in function (25) and $\theta_i^{F_i^+}$ is the unique solution to equation (16) corresponding to offer set F_i^+. It is apparent that $H^{F_i^+}(\theta_i) > H^{F_i}(\theta_i)$ for any θ_i. Then,

$$R^{F_i^+}(\rho) = H^{F_i^+}(\theta_i^{F_i^+}) > H^{F_i}(\theta_i^{F_i^+}) > H^{F_i}(\theta_i^{F_i}) = R^{F_i}(\rho).$$

Therefore, it is optimal to offer all the products in nest i, i.e., $F_i = \{1, 2, \ldots, m_i\}$. Moreover, their prices satisfy that $(p_{ij} - c_{ij}) + \frac{a_{ij}(p_{ij})}{a'_{ij}(p_{ij})}$ is constant for all the products in nest i. $\square$

Appendix C: Uniqueness of the Nash Equilibrium

In this section, we will investigate the uniqueness of the Nash equilibrium in the general case. First, we state sufficient conditions.

Condition 4 *(a) Denote Ψ as the region such that*

$$-\frac{\partial Q_i(\theta_i, \theta_{-i})}{\partial \theta_i} > \sum_{j\neq i} \frac{\partial Q_i(\theta_i, \theta_{-i})}{\partial \theta_j}, \ \boldsymbol{\theta} \in \Psi, \ i = 1, 2, \ldots, n.$$

(b) Denote Ω_i as the region such that $\theta_i + w_i(\theta_i)$ is log-*concave in $\theta_i \in \Omega_i$, $i = 1, 2, \ldots, n$.*

Notice that the NL model with product-differentiated price-sensitivity parameters within a nest and homogeneous nest coefficients, satisfies Condition 4 for any $\boldsymbol{\theta}$. Condition 4(a) is a standard diagonal dominant condition (see e.g., Vives 2001) and it says that a uniform increase of the *adjusted markups* by all the firms would result in a decrease of any firm's market share. In the NL model, Condition 4(a) is equivalent to

$$\gamma_i v_i(\theta_i) > \sum_{j=1}^{n} \gamma_j Q_j(\theta_j, \theta_{-j}) v_j(\theta_j). \tag{26}$$

From Lemma 1, inequality (26) can be implied by the following condition that is stronger but easier to be verified:

$$\min_i \gamma_i \min_{l,s} \beta_{l,s} > \max_i \gamma_i \max_{l,s} \beta_{l,s} \sum_{j=1}^{n} Q_j(\theta_j, \theta_{-j}),$$

which is equivalent to

$$\sum_{l=1}^{n}\left(\sum_{s=1}^{m_l} e^{\widetilde{\alpha}_{ls}-\beta_{ls}\theta_l}\right)^{\gamma_l} < \frac{\min_i \gamma_i \min_{l,s}\beta_{ls}}{\max_i \gamma_i \max_{l,s}\beta_{ls} - \min_i \gamma_i \min_{l,s}\beta_{ls}}. \tag{27}$$

From inequalition (27), Condition 4(b) can be satisfied when the *adjusted markups* θ_i are sufficiently large for all the firms.

Apparently, $\theta_i + w_i(\theta_i) > 0$ because each firm sells all her products at a positive average margin. Then, Condition 4(b) can be implied by a stronger condition that $\theta_i + w_i(\theta_i)$ or $w_i(\theta_i)$ is concave in θ_i for each $i = 1, 2, \ldots, n$ because if $w_i(\theta_i)$ is concave, then

$$\frac{\partial^2 \log(\theta_i + w_i(\theta_i))}{\partial \theta_i^2} = \frac{w_i''(\theta_i)(\theta_i + w_i(\theta_i)) - (w_i'(\theta_i))^2}{(\theta_i + w_i(\theta_i))^2} \le 0,$$

where $w_i'(\theta_i) = \partial w_i(\theta_i)/\partial\theta_i$ and $w_i''(\theta_i) = \partial^2 w_i(\theta_i)/\partial\theta_i^2$.

When θ_i is large enough, Condition 4(b) can also be satisfied without requiring the concavity of $\theta_i + w_i(\theta_i)$ or $w_i(\theta_i)$.

Lemma 2 *There exist a threshold $\widetilde{\theta}_i$ for each firm i such that $\theta_i + w_i(\theta_i)$ is* log-*concave in θ_i for $\theta_i \ge \widetilde{\theta}_i$.*

Proof of Lemma 2. Consider the derivatives of $\log(\theta_i + w_i(\theta_i))$,

$$\begin{aligned}
\frac{\partial \log(\theta_i + w_i(\theta_i))}{\partial \theta_i} &= \frac{w_i(\theta_i)v_i(\theta_i)}{\theta_i + w_i(\theta_i)},\\
\frac{\partial^2 \log(\theta_i + w_i(\theta_i))}{\partial \theta_i^2} &= \frac{w_i(\theta_i)}{\theta_i + w_i(\theta_i)} \cdot \frac{\partial v_i(\theta_i)}{\partial \theta_i} + v_i(\theta_i) \cdot \frac{\partial \frac{w_i(\theta_i)}{\theta_i + w_i(\theta_i)}}{\partial \theta_i}\\
&= \frac{w_i(\theta_i)}{\theta_i + w_i(\theta_i)} \cdot \frac{\partial v_i(\theta_i)}{\partial \theta_i} + v_i(\theta_i) \cdot \frac{\theta_i \cdot \big(-1 + w_i(\theta_i)v_i(\theta_i)\big) - w_i(\theta_i)}{(\theta_i + w_i(\theta_i))^2}
\end{aligned}$$

The log-concavity of $\theta_i + w_i(\theta_i)$ can be guaranteed by $\theta_i \cdot \big(-1 + w_i(\theta_i)v_i(\theta_i)\big) - w_i(\theta_i) \le 0$ because $v_i(\theta_i)$ is decreasing from Lemma 1. We will next show that $\theta_i \cdot \big(-1 + w_i(\theta_i)v_i(\theta_i)\big) \to 0$ as $\theta_i \to \infty$. Denote $\underline{\beta}_i = \min_s \beta_{is}$ and let Ξ_i be the set $\Xi_i = \{s : \beta_{is} = \underline{\beta}_i\}$. Then,

$$\begin{aligned}
-1 + w_i(\theta_i)v_i(\theta_i) &= \frac{-\big(\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\big)^2 + \Big(\sum_{s=1}^{m_i} \frac{1}{\beta_{is}} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\Big) \cdot \big(\sum_{s=1}^{m_i} \beta_{is} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\big)}{\big(\sum_{s=1}^{m_i} e^{\widetilde{\alpha}_{is}-\beta_{is}\theta_i}\big)^2}\\
&= \frac{1}{\Big(\sum_{s\in\Xi_i} e^{\widetilde{\alpha}_{is}} + \sum_{s\notin\Xi_i} e^{\widetilde{\alpha}_{is}-(\beta_{is}-\underline{\beta}_i)\theta_i}\Big)^2} \cdot \Bigg(-\Big(\sum_{s\in\Xi_i} e^{\widetilde{\alpha}_{is}} + \sum_{s\notin\Xi_i} e^{\widetilde{\alpha}_{is}-(\beta_{is}-\underline{\beta}_i)\theta_i}\Big)^2\\
&\quad + \Big(\sum_{s\in\Xi_i} \frac{1}{\underline{\beta}_i} e^{\widetilde{\alpha}_{is}} + \sum_{s\notin\Xi_i} \frac{1}{\beta_{is}} e^{\widetilde{\alpha}_{is}-(\beta_{is}-\underline{\beta}_i)\theta_i}\Big) \cdot \Big(\sum_{s\in\Xi_i} \underline{\beta}_i e^{\widetilde{\alpha}_{is}} + \sum_{s\notin\Xi_i} \beta_{is} e^{\widetilde{\alpha}_{is}-(\beta_{is}-\underline{\beta}_i)\theta_i}\Big)\Bigg)\\
&\approx \frac{\sum_{s\in\Xi_i} e^{\widetilde{\alpha}_{is}} \cdot \Big(\sum_{s\notin\Xi_i} (\frac{\beta_{is}}{\underline{\beta}_i} + \frac{\underline{\beta}_i}{\beta_{is}} - 2) e^{\widetilde{\alpha}_{is}-(\beta_{is}-\underline{\beta}_i)\theta_i}\Big)}{\Big(\sum_{s\in\Xi_i} e^{\widetilde{\alpha}_{is}}\Big)^2 + 2\Big(\sum_{s\in\Xi_i} e^{\widetilde{\alpha}_{is}}\Big) \cdot \Big(\sum_{s\notin\Xi_i} e^{\widetilde{\alpha}_{is}-(\beta_{is}-\underline{\beta}_i)\theta_i}\Big)}
\end{aligned}$$

In the above approximation, the higher order terms are ignored. Because $\frac{\beta_{is}}{\underline{\beta}_i}+\frac{\underline{\beta}_i}{\beta_{is}}-2>0$ and $\beta_{is}-\underline{\beta}_i>0$, then

$$\theta_i\cdot\Big(\big(\frac{\beta_{is}}{\underline{\beta}_i}+\frac{\underline{\beta}_i}{\beta_{is}}-2\big)e^{\widetilde{\alpha}_{is}-(\beta_{is}-\underline{\beta}_i)\theta_i}\Big)=\frac{\theta_i\cdot\big(\frac{\beta_{is}}{\underline{\beta}_i}+\frac{\underline{\beta}_i}{\beta_{is}}-2\big)}{e^{-\widetilde{\alpha}_{is}+(\beta_{is}-\underline{\beta}_i)\theta_i}}\to 0,\ \text{as } \theta_i\to\infty.$$

The above convergence holds because the exponential function is increasing faster than the linear function. Since $\left(\sum_{s\in\Xi_i}e^{\widetilde{\alpha}_{is}}\right)^2+2\left(\sum_{s\in\Xi_i}e^{\widetilde{\alpha}_{is}}\right)\cdot\left(\sum_{s\notin\Xi_i}e^{\widetilde{\alpha}_{is}-(\beta_{is}-\underline{\beta}_i)\theta_i}\right)\to\left(\sum_{s\in\Xi_i}e^{\widetilde{\alpha}_{is}}\right)^2$ as $\theta_i\to\infty$, therefore, $\theta_i\cdot\big(-1+w_i(\theta_i)v_i(\theta_i)\big)\to 0$. There exists $\widetilde{\theta}_i$ such that

$$\theta_i\cdot\big(-1+w_i(\theta_i)v_i(\theta_i)\big)\le\frac{1}{\max_s\beta_{is}}\le w_i(\theta_i),\ \text{for } \theta_i\ge\widetilde{\theta}_i.$$

Thus, $\theta_i+w_i(\theta_i)$ is log-concave for $\theta_i\ge\widetilde{\theta}_i$. $\square$

Tatonnement process can reach equilibrium under some mild conditions. In the basic *tatonnement* process, firms take turns in adjusting their price decisions and each firm reacts optimally to all other firms' prices without anticipating others' response, which can be interpreted as a way of expressing bounded rationality of agents. In each iteration, firms respond myopically to the choices of other firms in the previous iteration and the dynamic process can be expressed below.

Tatonnement Process: Select a feasible vector $\boldsymbol{\theta}^{(0)}$; in the k^{th} iteration determine the optimal response for each firm i as follows:

$$\theta_i^{(k)}=\arg\max_{\theta_i\in\Omega_i\cap\Psi}R_i(\theta_i,\theta_{-i}^{(k-1)}). \tag{28}$$

Theorem 6 *Suppose $\boldsymbol{\theta}^*$ is an equilibrium under Condition 4,*

(a) $\boldsymbol{\theta}^*$ is the unique pure Nash equilibrium of **Game II** *in region $(\bigcap_{i=1}^n\Omega_i)\bigcap\Psi$.*

(b) The unique pure Nash equilibrium $\boldsymbol{\theta}^*$ can be computed by the tatonnement scheme, starting from an arbitrary price vector $\boldsymbol{\theta}^{(0)}$ in the region $(\bigcap_{i=1}^n\Omega_i)\bigcap\Psi$, i.e., $\boldsymbol{\theta}^{(k)}$ converges to $\boldsymbol{\theta}^*$.

Proof of Theorem 6. Consider the first order and second order derivatives of $\log R_i(\theta_i,\theta_{-i})$ with respect to θ_i:

$$\begin{aligned}\frac{\partial\log R_i(\theta_i,\theta_{-i})}{\partial\theta_i}&=\frac{\partial\log(Q_i(\theta_i,\theta_{-i}))}{\partial\theta_i}+\frac{\partial\log(\theta_i+w_i(\theta_i))}{\partial\theta_i}=-\gamma_i(1-Q_i(\theta_i,\theta_{-i}))v_i(\theta_i)+\frac{\partial\log(\theta_i+w_i(\theta_i))}{\partial\theta_i},\\ \frac{\partial^2\log R_i(\theta_i,\theta_{-i})}{\partial\theta_i^2}&=-\gamma_i^2Q_i(\theta_i,\theta_{-i})(1-Q_i(\theta_i,\theta_{-i}))\big(v_i(\theta_i)\big)^2+\frac{\partial^2\log(\theta_i+w_i(\theta_i))}{\partial\theta_i^2}.\end{aligned}$$

The cross-derivative of $\log R_i(\theta_i,\theta_{-i})$ is

$$\frac{\partial^2\log R_i(\theta_i,\theta_{-i})}{\partial\theta_i\partial\theta_j}=\gamma_i\gamma_jQ_i(\theta_i,\theta_{-i})Q_j(\theta_j,\theta_{-j})v_i(\theta_i)v_j(\theta_j)\ge 0,\ \forall j\ne i.$$

Then,

$$\sum_{j\neq i}\frac{\partial^2 \log R_i(\theta_i,\theta_{-i})}{\partial\theta_i\partial\theta_j}=\gamma_i Q_i(\theta_i,\theta_{-i})v_i(\theta_i)\sum_{j\neq i}\gamma_j Q_j(\theta_j,\theta_{-j})v_j(\theta_j).$$

Under Condition 4, $R_i(\theta_i,\theta_{-i})$ is log-dominant diagonal,

$$-\frac{\partial^2 \log R_i(\theta_i,\theta_{-i})}{\partial\theta_i^2}\geq\sum_{j\neq i}\frac{\partial^2 \log R_i(\theta_i,\theta_{-i})}{\partial\theta_i\partial\theta_j} \tag{29}$$

The inequality holds because

$$\gamma_i^2 Q_i(\theta_i,\theta_{-i})(1-Q_i(\theta_i,\theta_{-i}))\big(v_i(\theta_i)\big)^2\geq\gamma_i Q_i(\theta_i,\theta_{-i})v_i(\theta_i)\sum_{j\neq i}\gamma_j Q_j(\theta_j,\theta_{-j})v_j(\theta_j)$$

under Condition 4(a), and

$$\frac{\partial^2 \log(\theta_i+w_i(\theta_i))}{\partial\theta_i^2}\leq 0$$

under Condition 4(b). The inequality (29) establishes the uniqueness of the Nash equilibrium to **Game II** (see e.g., Vives 2001).

If the equilibrium of a log-supermodular game with continuous payoff is unique, it is globally stable and a tatonnement process with dynamic response (28) converges to it from any initial point in the feasible region.

□

References

Akcay, Y., H. P. Natarajan, S. H. Xu. 2010. Joint dynamic pricing of multiple perishable products under consumer choice. *Management Sci.* **56**(8) 1345 – 1361.

Allon, G., A. Federgruen, M. Pierson. 2011. Price competition under multinomial logit demand functions with random coecients. Working paper, Northwestern University.

Anderson, S. P., A. de Palma. 1992. Multiproduct firms: A nested logit approach. *The Journal of Industrial Economics,* **40**(3) 261–276.

Anderson, S. P., A. de Palma, J. Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. 1st ed. The MIT Press.

Aydin, G., E. L. Porteus. 2008. Joint inventory and pricing decisions for an assortment. *Oper. Res.* **56**(5) 1247–1255.

Aydin, G., Jennifer K. Ryan. 2000. Product line selection and pricing under the multinomial logit choice model. Working paper.

Bernstein, F., A. Federgruen. 2004. A general equilibrium model for industries with price and service competition. *Oper. Res.* **52**(6) 868–886.

Berry, S., J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **63**(4) 841–890.

Borsch-Supan, A. 1990. On the compatibility of nested logit models with utility maximization. *Journal of Econometrics* **43** 373–388.

Cachon, G. P., A. G. Kok. 2007. Category management and coordination in retail assortment planning in the presence of basket shopping consumers. *Management Sci.* **53**(6) 934 – 951.

Corless, R. M., G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, D. E. Knuth. 1996. On the Lambert W function. *Advances in Computational Mathematics* **5**(1) 329–359.

Daganzo, C. R., M. Kusnic. 1993. Two properties of the nested logit model. *Transportation Science* **27**(4) 395–400.

Debreu, G. 1952. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences* **38**(10) 886–93.

Dong, L., P. Kouvelis, Z. Tian. 2009. Dynamic pricing and inventory control of substitute products. *Manufacturing Service Oper. Management* **11**(2) 317–339.

Erdem, T., J. Swait, Louviere. 2002. The impact of brand credibility on consumer price sensitivity. *International Journal of Research in Marketing* **19**(1) 1–19.

Federgruen, A., N. Yang. 2009. Competition under generalized attraction models: Applications to quality competition under yield uncertainty. *Management Sci.* **55**(12) 2028–2043.

Gallego, G., W. T. Huh, W. Kang, R. Phillips. 2006. Price competition with the attraction demand model: Existence of unique equilibrium and its stability. *Manufacturing Service Oper. Management* **8**(4) 359–375.

Gallego, G., C. Stefanescu. 2011. Upgrades, upsells and pricing in revenue management. Working paper, Columbia University.

Greene, W. H. 2007. *Econometric Analysis*. Pearson Education.

Hanson, W., K. Martin. 1996. Optimizing multinomial logit profit functions. *Management Sci.* **42**(7) 992–1003.

Hopp, W. J., X. Xu. 2005. Product line selection and pricing with modularity in design. *Management Sci.* **7**(3) 172–187.

Kok, A. G., Y. Xu. 2011. Optimal and competitive assortments with endogenous pricing under hierarchical consumer choice models. *Management Sci.* **57**(9) 1546–1563.

Li, H., W. T. Huh. 2011. Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing Service Oper. Management* **13**(4) 549 – 563.

Liu, G. 2006. On nash equilibrium in prices in an oligopolistic market with demand characterized by a nested multinomial logit model and multiproduct firm as nest. Discussion Papers, Statistics Norway.

Luce, R. D. 1959. *Individual Choice Behavior: A Theoretical Analysis*. Wiley New York.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. P. Zarembka, ed., *Frontiers in Econometrics*. Academic Press, New York, 105–142.

McFadden, D. 1980. Econometric models for probabilistic choice among products. *The Journal of Business* **53**(3) 13–29.

Milgrom, P., J. Roberts. 1990. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica: Journal of the Econometric Society* **58**(6) 1255–1277.

Nash, J. F. 1951. Non-coorperative games. *The Annals of Mathematics* **54**(2) 286–295.

Song, J., Z. Xue. 2007. Demand management and inventory control for substitutable products. Working paper, Duke University.

Topkis, D. M. 1998. *Supermodularity and complementarity*. Princeton University Press.

Vives, X. 2001. *Oligopoly pricing: Old ideas and new tools*. The MIT Press.

Williams, H. C. W. L. 1977. On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning* **9**(A) 285–344.

Williams, H. C. W. L., J. de D. Ortuzar. 1982. Behavioural theories of dispersion and the mis-specication of travel demand models. *Transportation Research* **16**(B) 167–219.