Efficient formulations for pricing under attraction demand models

Citation: Keller, Philipp W., Retsef Levi, and Georgia Perakis. "Efficient Formulations for Pricing Under Attraction Demand Models." Math. Program. 145, no. 1–2 (March 29, 2013): 223–261.

As Published: http://dx.doi.org/10.1007/s10107-013-0646-z

Publisher: Springer Berlin Heidelberg

Persistent URL: http://hdl.handle.net/1721.1/103405

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Mathematical Programming manuscript No.
(will be inserted by the editor)

Efficient Formulations for Pricing under Attraction Demand Models

Philipp W. Keller · Retsef Levi ·
Georgia Perakis

Received: date / Accepted: date

Abstract We propose a modeling and optimization framework to cast a broad range of fundamental multi-product pricing problems as tractable convex optimization problems. We consider a retailer offering an assortment of differentiated substitutable products to a population of customers that are price-sensitive. The retailer selects prices to maximize profits, subject to constraints on sales arising from inventory and capacity availability, market share goals, bounds on allowable prices and other considerations. Consumers' response to price changes is represented by attraction demand models, which subsume the well known multinomial logit (MNL) and multiplicative competitive interaction (MCI) demand models.

Our approach transforms seemingly non-convex pricing problems (both in the objective function and constraints) into convex optimization problems that can be solved efficiently with commercial software. We establish a condition which ensures that the resulting problem is convex, prove that it can be solved in polynomial time under MNL demand, and show computationally that our new formulations reduce the solution time from days to seconds. We also propose an approximation of demand models with multiple overlapping customer segments, and show that it falls within the class of demand models we are able to solve. Such mixed demand models are highly desirable in practice, but yield a pricing problem which appears computationally challenging to solve exactly.

Keywords pricing · revenue management · attraction demand models · multinomial logit

---

P. Keller
Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139
E-mail: pkeller@mit.edu

R. Levi, G. Perakis
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139
E-mail: retsef@mit.edu, georgiap@mit.edu

1 Introduction

In this paper, we study a general modeling and optimization framework that captures fundamental multi-product pricing problems. We consider the basic setting of a retailer offering an assortment of differentiated substitutable products to a population of customers who are price sensitive. The retailer selects prices to maximize the total profit subject to constraints on sales arising from inventory levels, capacity availability, market share goals, sales targets, price bounds, joint constraints on allowable prices, and other considerations.

The profit functions and constraints we consider are captured in the following general nonlinear optimization problem. The decision variables x_i are the prices charged for each of the products indexed by $i = 1, \dots, n$.

$$\begin{aligned} \max \quad & \sum_{i=1}^{n} a_i x_i d_i(\mathbf{x}) & \\ \text{s.t.} \quad & \sum_{i=1}^{n} A_{ki} d_i(\mathbf{x}) \le u_k & k = 1, 2, \dots m \\ & \underline{x}_i \le x_i \le \overline{x}_i & i = 1, 2, \dots n \end{aligned} \tag{P}$$

The input data to this model are the profit margins $a_i > 0$ for each product $i = 1, 2, \dots, n$, the matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and the vector $\mathbf{u} \in \mathbb{R}^m$ defining m constraints on the demand, and upper and lower bounds $\underline{\mathbf{x}}, \overline{\mathbf{x}} \in \mathbb{R}^n$ on allowable prices.

The price-demand functions $d_i(\mathbf{x})$, for $i = 1, \dots, n$, are central to the range of models that can be captured, as well as to the computational tractability of the resulting optimization problem. We represent customers' purchasing decisions through *attraction demand models*, which generalize the well-known *multinomial logit* (MNL) and *multiplicative competitive interaction* (MCI) demand models (McFadden 1974; Urban 1969). This approach is common in the recent revenue management literature, as well as in marketing and economics. The function $d_i : \mathbb{R}^n \to (0, 1)$ maps the prices of *all* the products to the observed customer demand for product i. We assume that attraction demand models have the following form:

$$d_0(\mathbf{x}) = \frac{1}{1 + \sum_j f_j(x_j)} \quad \text{and} \quad d_i(\mathbf{x}) = \frac{f_i(x_i)}{1 + \sum_j f_j(x_j)}, \quad i = 1, \dots, n. \tag{1}$$

The quantity $d_0(\mathbf{x})$ denotes the fraction of consumers opting not to purchase any product, and $d_i(\mathbf{x})$ is the demand for the i^{th} product when $i > 0$. The functions $f_i(x_i)$ are called *attraction functions* and are assumed to satisfy the following assumption:

Assumption 1 The attraction function $f_i : \mathbb{R} \to \mathbb{R}_{++}$ for each product $i = 1, 2, \dots, n$, satisfies:

(i) $f_i(\cdot)$ is strictly decreasing and twice differentiable on $\mathbb{R}$, and
(ii) $\lim_{x \to -\infty} f_i(x) = \infty$, and $\lim_{x \to \infty} x f_i(x) = 0$ (i.e., $f_i(x) \in o(\frac{1}{x})$).

Unlike much of the existing work in revenue management, these models allow the demands for each of the products to be interdependent functions of all the prices. However, under attraction demand models, the profit as a function of the prices set by the retailer is in general *not quasi-concave* (see Hanson and Martin (1996) and Appendix C). Moreover, many realistic constraints involving the demand model, such as production/inventory capacity bounds, give rise to a *non-convex* region of feasible prices. Maximizing the profit thus presents a challenging optimization problem. Our experiments show that commercial software may take over a day to solve a pricing-based formulation, even for relatively small instances. Furthermore, we have no a priori guarantee that such an approach will converge to a globally optimal solution.

The contributions of this paper are multifold. First, we provide equivalent reformulations of the pricing problem that are provably tractable and can be solved efficiently by commercial software. Defining the *inverse* attraction functions as $g_i(y) = f_i^{-1}(y)$, $y > 0$, our reformulations take the following general form:

$$\begin{aligned} \max \ & \mathit{\Pi}(\boldsymbol{\theta}) = \sum_i a_i \theta_i g_i\left(\frac{\theta_i}{\theta_0}\right) & & \\ \text{s.t.} \ & \sum_{i=1}^{n} A_{ki}\theta_i \le u_k & k = 1 \dots m' & \\ & \sum_{i=0} \theta_i = 1 & & \\ & \theta_i > 0 & i = 0 \dots n & \end{aligned} \tag{COP}$$

The decision variables $\theta_1, \dots, \theta_n$ represent the fraction of customers opting to purchase each product and the set of constraints defined by $(\mathbf{A}, \mathbf{u})$ has been extended with $2n$ additional linear constraints.

We establish a general and easily verifiable condition on the attraction demand model under which the reformulation gives rise to convex optimization problems. Specifically, our approach yields maximization problems of the form (COP) with *concave* objective functions $\mathit{\Pi}(\boldsymbol{\theta})$ and *linear* constraints. This is despite the nonlinear and non-separable nature of the demand model. Moreover, we prove that the logarithmic barrier method solves the pricing problem under MNL demand in a polynomial number of iterations of Newton's method. We confirm through extensive computational experiments that our formulations can be solved in seconds instead of days, compared to the naive formulations, and that they scale well to instances with thousands of products and constraints.

We then show how to apply our approach to obtain tractable approximations to the challenging pricing problem arising under *weak market segmentation*, where the pricing decisions affect the demand in multiple overlapping customer segments. Such models are exemplified by *mixed multinomial logit* models (Boyd and Mellman 1980; Cardell and Dunbar 1980). Representing even a small number of distinct segments, such as, for example, business and leisure travelers on a flight, may significantly improve accuracy over a single-segment model. Moreover, *any* random utility maximization model can be approximated arbitrarily closely by a mixed MNL model (see Mcfadden and Train (2000)).

We approximate such mixed demand functions with valid attraction demand models that yield convex optimization problems of the form (COP), and we bound the error with respect to the true multi-segment model. The attraction demand model in question is relatively complex, and the resulting objective function in our reformulation does not have a closed form. Nevertheless, we show how the objective function, its gradient and its Hessian can be computed efficiently, allowing standard optimization algorithms to be applied.

The remainder of this section reviews related work. Section 2 describes the price optimization problem, Section 3 presents our reformulation and Section 4 presents our approach for problems with multiple overlapping customer segments. Section 5 states the dual of our reformulation. Section 6 compares the different approaches we consider computationally. Omitted proofs may be found in Appendix A. Further details about specific attraction demand models may be found in Appendices B and C. The derivation of the dual of our reformulation and an algorithm for solving it are provided in Appendix D.

1.1 Literature Review

Pricing as a tool in revenue management usually arises in the context of perishable and nonrenewable inventory such as seats on a flight, hotel rooms, rental cars, internet service and electrical power supplies (see, e.g., the survey by Bitran and Caldentey (2003)). Dynamic pricing policies are also adopted in retail and other industries where short-term supply is more flexible, and the interplay between inventory management and pricing may thus take on even greater importance. (See Elmaghraby and Keskinocak (2003) for a survey of the literature on pricing with inventory considerations.)

The modeling framework studied in this paper generalizes a variety of more specialized pricing problems considered in the operations management literature. Much of the work focuses on *dynamic* pricing, under *stochastic* customer demand. The stochastic dynamic program arising under such models is generally intractable. Solution methods common in practice often rely on periodically re-solving single-period deterministic pricing problems, and this approach is known to be asymptotically optimal in some cases (Gallego and van Ryzin 1997; Talluri and van Ryzin 1998). Thus, the single-stage deterministic problems we consider play a central role. Our modeling framework relaxes two common but restrictive assumptions imposed in most existing multi-product pricing work. First, customers' substitution behavior can be modeled since demands are functions of *all* the prices. Secondly, a broad class of practical constraints can be enforced, well beyond just capacity bounds or inventory constraints allowed in existing models.

Specialized algorithms have been developed for solving certain single-period pricing problems. Hanson and Martin (1996) devise a path following heuristic for the *unconstrained* pricing problem under mixtures of MNL models, and Gallego and Stefanescu (2009) propose a column generation algorithm for a class of constrained problems with MNL demand (discussed in Section 2.2). Neither of these heuristics is guaranteed to find a globally optimal solution in finite time, nor can they be implemented directly with commercial software. They may also be computationally expensive in practice.

A number of recent papers have proposed multi-product pricing formulations using the *inverse* demand model to yield a concave objective function in terms of sales. Examples include Aydin and Porteus (2008), Dong et al (2008) and Song and Xue (2007). In contrast to our framework, these papers focus on capturing inventory holding and replenishment costs in the objective function rather than considering explicit constraints on prices and sales. They also limit their attention to the MNL or other specific demand models. Schön (2010) proposes a formulation of the *product line design* (PLD) problem with continuous prices. The PLD problem is closely related to the pricing problems we consider, but it involves *discrete* decision variables and specific types of capacity constraints. The variable transformation which arises when inverting the demand in our pricing problem is similar to the generalized Charnes-Cooper transformation described by Schaible (1974) for concave-convex fractional programs. However, the pricing problems studied in this paper are in general *not* concave-convex fractional programs.

Recent work on quantity-based revenue management also relaxes the assumption that the demands for different products are independent. In contrast to multi-product price-based revenue management, *network revenue management* (NRM) consists of choosing which subset of products to offer customers at each period from a menu with *fixed* prices, under inventory or capacity constraints. Gallego et al (2004) and Liu and van Ryzin (2008) consider customer substitution in this setting. Miranda Bront et al (2009) additionally consider overlapping customer segments, like in the weak market segmentation setting for which we provide an approximation. Talluri and van Ryzin (2004) provide an in-depth treatment of both price- and quantity-based revenue management and their relationship.

2 Modeling Framework

We first discuss the general pricing formulation (P), which has a non-concave (nor quasi-concave) objective and non-convex constraints in general. Then in Section 3 we describe the alternative formulation (COP), which is tractable.

A key question in pricing optimization is how to model the relationship between prices and the demands for each product. Assumption 1 is natural and captures, for example, the well known multinomial logit (MNL) model, with the attraction functions

$$f_i(x_i) = v_i e^{-x_i}, \tag{2}$$

and constant parameters $v_i > 0, i = 1, \ldots, n$. The technical requirement (*ii*) is mild, and in fact any attraction function can be modified to satisfy it without changing the objective values over the feasible region of (P). The demand $d_i(\mathbf{x})$ is equal to the attraction of product i normalized by the total attraction of all the customers' alternatives, including the option of not purchasing anything from the retailer in question. Without loss of generality, the attraction of the latter *outside alternative* is represented by the term 1 in the denominator. (Notice that the model is invariant to scaling of the attraction functions.) Further discussion of attraction demand models and explanations of how they can be adapted to satisfy Assumption 1 above may be found in Appendices B and C, respectively. Due to the demand model, the objective of problem (P) is nonlinear and in general not quasi-concave. The problem appears challenging even without any constraints. Nevertheless, we will show that the broad class of *constrained* problems we consider is in fact tractable (specifically, unimodal).

Indeed, a wide variety of constraints can be represented by the formulation (P). Capacity and inventory bounds are very common and arise in revenue management problems. For example, in a flight reservation system each product may represent an itinerary with given travel restrictions, while seat availability on shared flight legs is represented by coupling resource constraints. To capture these constraints in our model, A_{ki} would be set to 1 if itinerary i uses leg k and zero otherwise. The upper bound u_k would be set to the number of seats available for leg k, divided by the total size of the population. In a retail setting, the identity matrix $\mathbf{A} = \mathbf{I}$ may be used and u_k may be set to the inventory available for item k. Minimum sales targets for a group of products may be represented with additional inequality constraints, with negative coefficients since they place a lower bound on the demands. In product-line design such operational constraints may be less important, but production capacities and minimum market-share targets take a similar form. Unfortunately, even though the constraints of (P) are linear in the demands for each of the products, they yield a *non-linear* and *non-convex* feasible region of prices in general. See Appendix C for examples.

2.1 Marginal Costs, Joint Price Constraints and Other Extensions

We have defined the objective function of (P) in terms of relative profit margins, such as when a capacity reseller earns commissions. A per-unit production cost may be incorporated by using the objective

$$\sum_{i=1}^{n} (x_i - c_i) d_i(\mathbf{x}),$$

redefining the "prices" in our general formulation as the profit margins $\hat{x}_i \triangleq (x_i - c_i), i = 1, \ldots, n$, and shifting the attraction functions accordingly. This motivates why we allow negative prices in general, since the profit margin $\hat{x}_i$ may be negative even if the price x_i paid by the consumer is positive.

Moreover, any joint constraint on prices of the form

$$x_i \geq x_j + \delta_{ij}, \qquad \delta_{ij} \in \mathbb{R} \tag{3}$$

can be re-expressed as

$$\begin{aligned} f_i(x_i) \leq f_i(x_j + \delta_{ij}) \quad &\Leftrightarrow \quad f_i(x_i) d_0(\mathbf{x}) \leq f_i(x_j + \delta_{ij}) d_0(\mathbf{x}) \\ &\Leftrightarrow \quad d_i(\mathbf{x}) \leq \frac{f_i(x_j + \delta_{ij})}{f_j(x_j)} d_j(\mathbf{x}), \end{aligned} \tag{4}$$

where we have used the monotonicity of $f_i(\cdot)$, the positivity of $d_0(\mathbf{x})$, and the fact that

$$f_i(x_i) = \frac{d_i(\mathbf{x})}{d_0(\mathbf{x})}. \tag{5}$$

Under mild assumptions on an MNL demand model, the ratio in the last inequality of (4) is a constant. The resulting linear constraint is captured by the formulation (P). A similar transformation is possible for linear attraction demand models. The details can be found in Appendix C.

We briefly mention two other straightforward extensions to our model. First, multiple customer segments may be represented by distinct, independent demand models. If it is possible to present different prices to each segment, or if disjoint subsets of products are offered to each segment, the pricing problem corresponds to multiple instances of (P) coupled only through linear constraints. Our approach in the next section carries through directly in such cases. (This is in contrast to the model discussed in Section 4 where all customers are offered all products at the same prices.) Secondly, since the equations describing inventory dynamics are *linear*, deterministic multi-stage pricing problems can be expressed in a similar form. Each period can be represented by a copy of (P). The copies are coupled only through the linear inventory dynamics constraints.

2.2 Special Cases with Convex Constraints

Proposition 1 below characterizes a class of constraints for which the feasible region of (P) *is* convex. Although the remainder of this paper considers the general formulation (P), the sub-class of problems considered in this subsection has been implicitly studied in previous work. It arises naturally in specific revenue management problems, and encompasses the problems involving customer choice considered by Gallego and Stefanescu (2009).

To our knowledge, the condition given here has never been made explicit. We provide an explanation of why versions of the pricing problem satisfying it have been found relatively easy to solve in Section 3.1.

Proposition 1 *If the attraction functions $f_1, f_2, \ldots, f_n$ are convex, and the constraints satisfy*

$$A_{ki} \geq u_k \geq 0, \qquad \text{for each } k = 1, 2, \ldots, m, \text{ and } i = 1, 2, \ldots, n, \tag{6}$$

then the feasible region of (P) *is a convex set.*

Proof Clearly the bounds on the prices define a convex set. The k^{th} inequality constraint, $1 \leq k \leq m$, is convex since it may be expressed as a positive linear combination of convex functions:

$$\sum_{i=1}^{n} \frac{A_{ki} f_i(x_i)}{1 + \sum_{j=1}^{n} f_j(x_j)} \leq u_k \Leftrightarrow \sum_{i=1}^{n} A_{ki} f_i(x_i) \leq \left(1 + \sum_{j=1}^{n} f_j(x_j)\right) u_k$$
$$\Leftrightarrow \sum_{i=1}^{n} (A_{ki} - u_k) f_i(x_i) \leq u_k. \qquad \square$$

The assumption of convex attraction functions implies that the marginal number of sales lost due to price increases is *decreasing.* This is in many cases a natural assumption. Moreover, it can be verified that the MNL, MCI and linear attraction demand models do satisfy it. Condition (6) is satisfied in certain revenue management problems where the columns of the matrix $\mathbf{A}$ represent the vectors of resources consumed in producing one unit of the respective products, and $\mathbf{u}$ is the vector of the inventories available from each resource. Suppose, for example, that each product represents a seat on the same flight but with different fare restrictions. Then, as long as there are more potential customers than seats on the flight, the condition is satisfied, because the parameters A_{ki} corresponding to the k^{th} capacity constraint are all equal to one, and $0 < u_k \leq 1$ (since the demands are normalized by the population size).

However, most problems of interest do not fall into this special class, thereby motivating our more general approach. For instance, if the customers choose between seats on different flights, some of the parameters A_{ki} will be set to 0, violating the condition. Thus, we consider the case of general input data in the remainder of this paper.

3 Market Share Reformulation

In this section, we transform problem (P) into the equivalent optimization problem (COP) over the space of market shares. This transformation eliminates the need to explicitly represent the nonlinear demand model in the constraints, while preserving the bounds on prices as linear constraints. We denote the fraction of lost sales and the market share of each product i in (1) by $\theta_0 = d_0(\mathbf{x})$ and $\theta_i = d_i(\mathbf{x})$, respectively. The attraction functions $f_1, \ldots, f_n$ are invertible since they are strictly decreasing by Assumption 1(i). Define the *inverse attraction function* [1] $g_i : \mathbb{R}_{++} \to \mathbb{R}$ as the inverse of f_i, for each product i. From (5), the prices corresponding to a given vector of market shares

[1] Although the inverse attraction functions always exist, they may not have a closed form for some complex demand models. In Section 4.2, we show that the objective function's derivatives can nevertheless be computed efficiently, allowing general purpose algorithms to be used.

$\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_n)$ can thus be expressed as,

$$x_i = f_i^{-1}\left(\frac{d_i(\mathbf{x})}{d_0(\mathbf{x})}\right) = g_i\left(\frac{\theta_i}{\theta_0}\right), \qquad \text{for } i = 1, 2, \ldots, n. \tag{7}$$

Optimization problem (P) can be rewritten as (COP). The market shares θ_i play the role of decision variables, and the the prices x_i are represented as functions of $\boldsymbol{\theta}$. In addition to the original constraints in (P) there is a simplex constraint on the market shares, and strict positivity of the market shares is enforced. These constraints are implied in (P) since the fraction of lost sales and the market share of each product i in (1) naturally satisfy

$$\sum_{i=0}^{n} d_i(\mathbf{x}) = 1, \qquad \text{with} \qquad d_i(\mathbf{x}) > 0, \qquad \text{for } i = 0, 1, \ldots, n. \tag{8}$$

To solve (COP) in practice, we may relax the strict inequalities. As any of the market shares θ_i go to zero, some of the prices go to positive or negative infinity. The price bounds in (P) thus exclude such solutions. [2] These price bounds are captured in (COP) by extending the matrix $\mathbf{A}$ and vector $\mathbf{u}$. The new number of inequality constraints on the demands is $m' = m + 2n$, and the additional coefficients and right-hand sides are given by, $\forall i, j \in \{1, 2, \ldots, n\}$,

$$\begin{aligned} A_{m+2i-1,j} &= \begin{cases} f_i(\underline{x}_i) & \text{if } i \neq j \\ 1 + f_i(\underline{x}_i) & \text{if } i = j \end{cases}, \qquad & u_{m+2i-1} &= f_i(\underline{x}_i), \\ A_{m+2i,j} &= \begin{cases} -f_i(\overline{x}_i) & \text{if } i \neq j \\ -1 - f_i(\overline{x}_i) & \text{if } i = j \end{cases}, \qquad & u_{m+2i} &= -f_i(\overline{x}_i). \end{aligned} \tag{9}$$

The following lemma shows that (P) and (COP) are equivalent, in that there is a one-to-one correspondence between feasible points of the two problems which preserves the objective function value.

Lemma 1 *Formulations* (P) *and* (COP) *are equivalent.*

Proof Consider the problem

$$\begin{aligned} \max \ & \sum_i a_i x_i d_i(\mathbf{x}) \\ \text{s.t.} \ & \sum_i A_{ki} d_i(\mathbf{x}) \leq u_k \qquad k = 1 \ldots m', \end{aligned} \tag{P1}$$

where we have replaced the price bounds in (P) with the new constraints (9). Using the monotonicity of f_i and (5), the bounds on the price of the i^{th} product

[2] Even in the absence of these price bounds, Assumption 1 ensures that an optimal solution to (COP) which is strictly positive in each component exists, when it satisfies the convexity condition of the next section. This follows from Proposition 3 characterizing its dual in Appendix D.

may by expressed in terms of the attractions as

$$\begin{aligned} \underline{x}_i \leq x_i \leq \overline{x}_i &\Leftrightarrow f_i(\underline{x}_i) \geq f_i(x_i) \geq f_i(\overline{x}_i) \\ &\Leftrightarrow f_i(\underline{x}_i)d_0(\mathbf{x}) \geq d_i(\mathbf{x}) \geq f_i(\overline{x}_i)d_0(\mathbf{x}). \end{aligned} \tag{10}$$

Using (8), the above can be written in terms of the market shares for all the products as

$$f_i(\underline{x}_i)\left(1-\sum_{j=1}^{n} d_j(\mathbf{x})\right) \geq d_i(\mathbf{x}) \geq f_i(\overline{x}_i)\left(1-\sum_{j=1}^{n} d_i(\mathbf{x})\right). \tag{11}$$

These are precisely the constraints described by the $(m{+}2i{-}1)^{\text{th}}$ and $(m{+}2i)^{\text{th}}$ rows of $\mathbf{A}$ and $\mathbf{u}$ in problem (P1). Consequently, the problems (P1) and (P) are equivalent.

We define the mapping $T : \mathcal{X} \to \Theta$ from the feasible region $\mathcal{X} \subseteq \mathbb{R}^n$ of (P) to the feasible region $\Theta \subset (0,1)^{n+1}$ of (COP) by $T(\mathbf{x}) = [d_0(\mathbf{x}), \ldots, d_n(\mathbf{x})]^\top$. The inverse mapping is given by $T^{-1}(\boldsymbol{\theta}) = \left[g_1\left(\frac{\theta_1}{\theta_0}\right), \ldots, g_n\left(\frac{\theta_n}{\theta_0}\right)\right]^\top$, as in (7). For any $\mathbf{x} \in \mathcal{X}$, $T(\mathbf{x})$ is feasible for (COP) because (i) the inequality constraints are equivalent to those of (P1) above, and (ii) the simplex and positivity constraints are satisfied by (8). Similarly, for any $\boldsymbol{\theta} \in \Theta$, $T^{-1}(\boldsymbol{\theta})$ is feasible for (P1) and (P). Finally, the objective value $\Pi(T(\mathbf{x}))$ of (COP) is equal to the objective value of (P). □

3.1 Convexity condition

Theorem 1 below provides a condition on the attraction functions $f_1, f_2, \ldots, f_n$ under which (COP) is a convex problem, and can thus be solved with general purpose convex optimization algorithms. The condition is fairly general and requires only a property of the individual attraction functions. Corollary 1 of Appendix C verifies that it holds for MNL, MCI and linear attraction demand models.

In light of Lemma 1, Theorem 1 also provides insights into the structure of the original pricing problem (P). Specifically, it implies that there are no (strict) local maxima. In particular, note that together with Proposition 1, it implies that the pricing problems considered by Gallego and Stefanescu (2009) are in fact maximizations of a unimodal profit function over a convex feasible region. More generally, the theorem specifies a condition such that under Assumption 1 problem (P) does not have any strict local maxima, even when its feasible region is not convex.

Theorem 1 *If the attraction functions are such that, in the space of market shares,*

$$2g_i'(y) + y g_i''(y) \leq 0, \qquad \forall y > 0, \qquad i = 1, 2, \ldots, n, \tag{12}$$

or equivalently in the space of prices,

$$\frac{2\,|f_i'(x)|}{f_i(x)} \geq \frac{f_i''(x)}{|f_i'(x)|}, \qquad \forall x \in \mathbb{R}, \qquad i = 1, 2, \ldots, n, \tag{13}$$

then the objective $\Pi(\theta)$ of (COP) *is concave. Furthermore,*

(i) every local maximum of (COP) *is a global maximum, and*
(ii) every local maximum of (P) *is a global maximum.*

Proof Since the objective function is $\Pi(\boldsymbol{\theta}) = \sum_{i=1}^n a_i\theta_i g_i(\frac{\theta_i}{\theta_0})$, with positive coefficients a_i, we need only show concavity of each term $\Pi_i(\theta_0, \theta_i) \triangleq \theta_i g_i(\frac{\theta_i}{\theta_0})$. The gradient and Hessian of $\Pi_i(\theta_0, \theta_i)$ are

$$\nabla \Pi_i = \begin{bmatrix} -\frac{\theta_i^2}{\theta_0^2} g_i'\left(\frac{\theta_i}{\theta_0}\right) \\ g_i\left(\frac{\theta_i}{\theta_0}\right) + \frac{\theta_i}{\theta_0} g_i'\left(\frac{\theta_i}{\theta_0}\right) \end{bmatrix} \qquad \text{and} \tag{14}$$

$$\nabla^2 \Pi_i = \left(2g_i'(\frac{\theta_i}{\theta_0}) + \frac{\theta_i}{\theta_0} g_i''(\frac{\theta_i}{\theta_0})\right) \begin{bmatrix} \frac{\theta_i^2}{\theta_0^3} & \frac{-\theta_i}{\theta_0^2} \\ \frac{-\theta_i}{\theta_0^2} & \frac{1}{\theta_0} \end{bmatrix}. \tag{15}$$

The first factor is non-positive by condition (12). Taking any vector $\mathbf{z} = [u, v]^\top \in \mathbb{R}^2$, we have that

$$\mathbf{z}^\top \begin{bmatrix} \frac{\theta_i^2}{\theta_0^3} & \frac{-\theta_i}{\theta_0^2} \\ \frac{-\theta_i}{\theta_0^2} & \frac{1}{\theta_0} \end{bmatrix} \mathbf{z} = \frac{1}{\theta_0}\left(u^2\left(\frac{\theta_i}{\theta_0}\right)^2 - 2uv\left(\frac{\theta_i}{\theta_0}\right) + v^2\right) = \frac{1}{\theta_0}\left(u\left(\frac{\theta_i}{\theta_0}\right) - v\right)^2 \geq 0,$$

so the Hessian of Π_i is negative semi-definite, and Π_i is concave.

Differentiating $x = g(y) \triangleq g_i(y)$, for some fixed i, with respect to y, we have

$$g'(y) = \left(f^{-1}(y)\right)' = \frac{1}{f'\left(f^{-1}(y)\right)} = \frac{1}{f'\left(g(y)\right)} = \frac{1}{f'\left(x\right)}, \tag{16}$$

and using the chain rule,

$$g''(y) = \frac{-1}{\left(f'\left(g(y)\right)\right)^2} f''\left(g(y)\right) g'(y) = \frac{-1}{\left(f'\left(x\right)\right)^2} f''\left(x\right) \frac{1}{f'(x)} = \frac{-f''(x)}{\left(f'\left(x\right)\right)^3}. \tag{17}$$

Substituting into (12) and multiplying by the strictly positive quantity $\left(f'(x)\right)^2$,

$$\frac{2}{f'(x)} + y\frac{-f''(x)}{\left(f'\left(x\right)\right)^3} \leq 0 \quad \Leftrightarrow \quad 2f'(x) - \frac{f(x)f''(x)}{f'(x)} \leq 0 \quad \Leftrightarrow \quad \frac{2f'(x)}{f(x)} \leq \frac{f''(x)}{f'(x)}.$$

We used the fact that $f'(x) < 0$ while $f(x) > 0$. The equivalence with (13) follows since $f'(x)$ is negative.

Then (i) follows directly from the concavity of the objective function and convexity of the feasible region in (COP). As shown in the proof of Lemma 1, there is a one-to-one, invertible, continuous mapping between feasible points of (COP) and (P), and the mapping preserves the value of the continuous objective function. Thus any local maximum of (P) would correspond to a local maximum of (COP). □

3.2 Self-Concordant Barrier Method for the MNL Demand Model

In this section, we restrict our attention to the MNL demand model, and show that problem (COP) may be solved in polynomial time using interior point methods. In particular, we show that applying the barrier method to (COP) gives rise to a *self-concordant* objective. (The latter concept is defined below.)

Note that a similar treatment could be applied to other attraction demand models, but for ease of exposition, we focus on the more common MNL demand model. Other optimization algorithms could also be applied. The barrier subproblem (18) below minimizes a twice-differentiable convex function over a simplex for *any* attraction functions satisfying the conditions of Theorem 1. Ahipasaoglu et al (2008) show that a first-order modified Frank-Wolfe algorithm exhibits linear convergence for such problems. In Section 6, we use a commercially available implementation of a primal-dual interior-point method to solve (CMNL).

Under the MNL model, the attraction functions and their inverses are defined for $i = 1, \ldots, n$ as $f_i(x_i) \triangleq v_i e^{-x_i}$, and $g_i(y_i) = -\log(y_i/v_i)$. Then problem (COP) is

$$\begin{aligned} \max \quad & \mathit{\Pi}(\boldsymbol{\theta}) = -\sum_{i=1}^{n} a_i \theta_i \log \frac{\theta_i}{v_i \theta_0} \\ \text{s.t.} \quad & \mathbf{A}\boldsymbol{\theta} \le u, \qquad \mathbf{e}^\top \boldsymbol{\theta} = 1, \qquad \boldsymbol{\theta} > 0 \end{aligned} \tag{CMNL}$$

where $\mathbf{e}$ denotes the vector of ones. We note that the objective has a form similar to the relative entropy, or Kullback-Leibler (KL) divergence,

$$\mathcal{K}(\boldsymbol{\pi}, \boldsymbol{\eta}) \triangleq \sum_{i=1}^{n} \pi_i \log \frac{\pi_i}{\eta_i}.$$

This is a measure of the distance between two probability distributions $\boldsymbol{\pi}, \boldsymbol{\eta} \in \mathbb{R}_+^n, \mathbf{e}^\top \boldsymbol{\pi} = \mathbf{e}^\top \boldsymbol{\eta} = 1$. For problems involving the KL-divergence where the denominators η_i are *constant*, the existence of a self-concordant barrier is known (see (Calafiore 2007) and (den Hertog et al 1995)). Unfortunately these results cannot be used in our setting, since the objective $\mathit{\Pi}(\boldsymbol{\theta})$ is *not* separable: each term also depends on the *decision variable* θ_0.

The barrier method solves (CMNL) by solving a series of problems parameterized by $t > 0$,

$$\begin{aligned} \min \quad & \mathit{\Psi}_t(\boldsymbol{\theta}) = -t\mathit{\Pi}(\boldsymbol{\theta}) + \mathit{\Phi}(\boldsymbol{\theta}) \\ \text{s.t.} \quad & \mathbf{e}^\top \boldsymbol{\theta} = 1 \end{aligned} \tag{18}$$

where the logarithmic barrier is defined by

$$\mathit{\Phi}(\boldsymbol{\theta}) = -\sum_{i=1}^{n} \log \theta_i - n \log \theta_0 - \sum_{k=1}^{m'} \log \left(u_k - \sum_{i=1}^{n} A_{ki} \theta_i \right). \tag{19}$$

We have changed the maximization problem to a minimization problem for consistency with the literature. Each inequality constraint in (CMNL) has been replaced by a term in the barrier which goes to infinity as the constraint becomes tight. The term for the positivity constraint on θ_0 is replicated n times for reasons that will become apparent. We let $\bar{m} = m' + 2n = m + 4n$ be the number of inequality constraints in a slightly modified version of problem (CMNL), including the $m' = m + 2n$ constraints represented by the pair $(\mathbf{A}, \mathbf{u})$, the n positivity constraints on the market share of each product, and n replications of the positivity constraint $\theta_0 > 0$, for which n logarithmic barrier terms were added to the barrier (19).

Denote the optimal solution of (18) for a given value of $t > 0$ by $\boldsymbol{\theta}^*(t)$. Given an appropriate, strictly feasible starting point $\boldsymbol{\theta}$, a positive initial value for t, a constant factor $\mu > 1$ and a tolerance $\epsilon > 0$, the barrier method consists of the following steps:

1. Solve (18) using equality-constrained Newton's method with starting point $\boldsymbol{\theta}$ to obtain $\boldsymbol{\theta}^*(t)$.
2. Update the starting point $\boldsymbol{\theta} := \boldsymbol{\theta}^*(t)$.
3. Stop if $\bar{m}/t \leq \epsilon$, otherwise update $t := \mu t$ and go to Step 1.

As the value of t becomes large, the solution $\boldsymbol{\theta}^*(t)$ tends towards the optimal solution of (CMNL). The termination condition in Step 3 guarantees that the objective value is sufficiently close to its optimal value. In practice, a phase I problem must be solved to find an appropriate initial point $\boldsymbol{\theta}$, and a redundant constraint is added to (CMNL) for technical reasons.

The computational complexity of Newton's method can be analyzed when the objective function is *self-concordant.*

Definition 1 (Self-concordance.) A convex scalar function $f : \mathcal{D} \to \mathbb{R}$ is said to be self-concordant when $|f'''(x)| \leq 2\left(f''(x)\right)^{3/2}$ for every point $x \in \mathcal{D} \subseteq \mathbb{R}$ in the domain of f. A multivariate function $f : \mathcal{F} \to \mathbb{R}$ is self-concordant if it is self-concordant along every line in its domain $\mathcal{F} \subseteq \mathbb{R}^n$.

The class of self-concordant functions is closed under addition and composition with affine functions (see, for example, (Boyd and Vandenberghe 2004)). Our proof that the objective function of the problem (18) falls withing this class relies on Theorem 2 presented here.

Theorem 2 *The function*

$$f(x,y) = tx \log \frac{x}{\beta y} - \log xy \tag{20}$$

is strictly convex and self-concordant on $\mathbb{R}^2_{++}$ for $\beta > 0$ and $t \geq 0$.

Proof We explicitly compute the derivatives of (20) and obtain

$$\nabla f(x,y) = \begin{bmatrix} t + t\log\frac{x}{y} - \frac{1}{x} - t\log\beta \\ -\frac{tx+1}{y} \end{bmatrix}, \quad \nabla^2 f(x,y) = \begin{bmatrix} \frac{tx+1}{x^2} & \frac{-t}{y} \\ \frac{-t}{y} & \frac{tx+1}{y^2} \end{bmatrix},$$

$$\text{and} \qquad \nabla^3 f(x,y) = \begin{bmatrix} -\frac{tx+2}{x^3} & 0 \\ 0 & \frac{t}{y^2} \end{bmatrix} \begin{bmatrix} 0 & \frac{t}{y^2} \\ \frac{t}{y^2} & \frac{-2tx-2}{y^3} \end{bmatrix}.$$

Then for an arbitrary direction $\mathbf{h} = [a, b]^\top \in \mathbb{R}^2$ and any $[x, y]^\top \in \mathbb{R}^2_{++}$

$$\nabla^2 f[\mathbf{h}, \mathbf{h}] = \frac{a^2(tx+1)}{x^2} - \frac{2abt}{y} + \frac{b^2(tx+1)}{y^2}$$

For ease of notation, define $s \geq 0$ and $u, v \in \mathbb{R}$ such that $s = tx \geq 0$, $u = a/x$ and $v = b/y$. Then rewrite

$$\begin{aligned}\nabla^2 f[\mathbf{h}, \mathbf{h}] &= u^2(s+1) - 2uvs + v^2(s+1) = s(u^2 - 2uv + v^2) + (u^2 + v^2) \\ &= s\underbrace{(u-v)^2}_{A} + \underbrace{(u^2+v^2)}_{B}.\end{aligned}$$

Both terms A and B are non-negative, and the second term B is positive unless $u = v = 0$, i.e. $a = b = 0$. Thus the Hessian $\nabla^2 f$ is positive definite and f is strictly convex on $\mathbb{R}_{++}$. We also expand

$$\begin{aligned}-\nabla^3 f[\mathbf{h}, \mathbf{h}, \mathbf{h}] &= \frac{a^3(tx+2)}{x^3} - \frac{3ab^2t}{y^2} + \frac{2b^3(tx+1)}{y^3} \\ &= u^3(s+2) - 3uv^2 s + 2v^3(s+1) \\ &= s(u^3 - 3uv^2 + 2v^3) + 2(u^3 + v^3) \\ &= s\underbrace{(u-v)^2(u+2v)}_{K} + \underbrace{2(u^3+v^3)}_{L}.\end{aligned}$$

We now show that f is self-concordant, that is

$$|\nabla^3 f[\mathbf{h}, \mathbf{h}, \mathbf{h}]| \leq 2(\nabla^2 f[\mathbf{h}, \mathbf{h}])^{\frac{3}{2}} \quad \Leftrightarrow \quad (\nabla^3 f[\mathbf{h}, \mathbf{h}, \mathbf{h}])^2 \leq 4(\nabla^2 f[\mathbf{h}, \mathbf{h}])^3, \tag{21}$$

by appropriately factoring the difference of the two sides of the inequality, and showing that it is non-negative. That is, we verify the non-negativity of

$$\begin{aligned}4(\nabla^2 f[\mathbf{h}, \mathbf{h}])^3 - (\nabla^3 f[\mathbf{h}, \mathbf{h}, \mathbf{h}])^2 &= 4(sA + B)^3 - (sK + L)^2 \\ &= (4s^3A^3 + 12s^2A^2B + 12sAB^2 + 4B^3) - (s^2K^2 + 2sKL + L^2) \\ &= 4s^3A^3 + s^2(12A^2B - K^2) + s(12AB^2 - 2KL) + (4B^3 - L^2).\end{aligned} \tag{22}$$

For the leading term that $A^3 = (u-v)^6 \geq 0$. Then for the second term

$$\begin{aligned}12A^2B - K^2 &= 12(u-v)^4(u^2+v^2) - (u-v)^4(u+2v)^2 \\ &= (u-v)^4\left(12u^2 + 12v^2 - (u+2v)^2\right) \\ &= (u-v)^4(11u^2 - 4uv + 8v^2) \geq 0,\end{aligned}$$

since the quadratic form can be written as $\left[\begin{smallmatrix} u, v \end{smallmatrix}\right]\left[\begin{smallmatrix} 11 & -2 \\ -2 & 8 \end{smallmatrix}\right]\left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \geq 0$. For the term in s,

$$\begin{aligned}12AB^2 - 2KL &= 12(u-v)^2(u^2+v^2)^2 - 4(u-v)^2(u+2v)(u^3+v^3) \\ &= 4(u-v)^2\left(3(u^4 + 2u^2v^2 + v^4) - (u^4 + uv^3 + 2u^3v + 2v^4)\right) \\ &= 4(u-v)^2\left(2u^2(u^2 - uv + v^2) + v^2(4u^2 - uv + v^2)\right) \geq 0.\end{aligned}$$

Finally, the last term is $4B^3 - L^2 = 4(u^2+v^2)^6 - 4(u^3+y^3)^2 = 4u^2v^2(3u^2 - 2uv + 3v^2) \geq 0$. Together, the preceding four inequalities with the fact that $s \geq 0$ show that (22) is non-negative, and that (21) holds. □

In order to prove Theorem 3, we shall use the result of Section 11.5.5 in (Boyd and Vandenberghe 2004). The result applies to minimization problems, but the objective function of the maximization problem (CMNL) can be negated to obtain an equivalent minimization problem. [3] We first define some additional notation. The constant M is an *a priori* lower bound on the optimal value of (CMNL) (and thus an upper bound for the corresponding minimization problem). It is used in the phase I feasibility problem. The constant G is an upper bound on the norm of the gradient of the constraints. Since the positivity constraints have a gradient of norm 1, and the gradients of the inequality constraints are the rows of the matrix $\mathbf{A}$, which we denote here by $\mathbf{A}_{k,\cdot}$, we set

$$G = \max\left\{1, \max_{1\leq k\leq m'} \|\mathbf{A}_{k,\cdot}\|\right\}.$$

We define R to be the radius of a ball centered at the origin containing the feasible set. Since any feasible vector $\boldsymbol{\theta}$ lies in the unit simplex, we may set $R = 1$. We define two constants depending on the parameters of the backtracking line search algorithm used in Newton's method. We let $\gamma = \frac{\alpha\beta(1-2\alpha)^2}{20-8\alpha}$ and $c = \log_2 \log_2 \frac{1}{\epsilon}$. Typical values are $\alpha \in [0.01, 0.3]$ and $\beta \in [0.1, 0.8]$. The constant c can reasonably be approximated by $c = 6$. Finally, we let $p^* > M$ be the optimal value of (CMNL), and we define $\bar{p}^*$ to be the optimal value of the phase I feasibility problem used to find a suitable starting point (see Section 11.5.4 of (Boyd and Vandenberghe 2004)). The latter value is close to zero when a problem is nearly infeasible or nearly feasible, and is far from zero if the problem is clearly feasible or infeasible.

The bound of Theorem 3 depends on the number of constraints and the number of products through $\bar{m} = m + 4n$, and on two terms that depend on the problem data, $C_1 = \log_2 \frac{G}{|\bar{p}^*|}$ and $C_2 = \log_2 \frac{p^*-M}{\epsilon}$. The constant C_1 should be interpreted as measuring the difficulty of the phase I feasibility problem, while C_2 can be interpreted as measuring the difficulty of solving the phase II problem.

Theorem 3 *Problem* (CMNL) *may be solved to within a tolerance* $\epsilon > 0$ *in a polynomial number* $N = N_I + N_{II}$ *iterations of Newton's method, where*

$$N_I = \left\lceil \sqrt{\bar{m}+2}\log_2\left(\frac{(\bar{m}+1)(\bar{m}+2)GR}{|\bar{p}^*|}\right)\right\rceil \left(\frac{1}{2\gamma} + c\right)$$

[3] Because of this negation, the values of M and p^* defined below are also the negation of the corresponding values in (Boyd and Vandenberghe 2004). Therefore the phase I minimization problem is unchanged, but the objective function of the phase II minimization problem is the negation of the objective of (CMNL).

is the number of iterations required to solve the phase I problem, and

$$N_{II} = \left\lceil \sqrt{\bar{m}+1}\log_2\left(\frac{(\bar{m}+1)(p^*-M)}{\epsilon}\right)\right\rceil\left(\frac{1}{2\gamma}+c\right)$$

is the number of iterations required to solve the phase II problem.

Proof The objective of (18) is

$$\begin{aligned}\Psi_t(\boldsymbol{\theta}) &= -t\Pi(\theta) + \Phi(\theta)\\ &= t\sum_{i=1}^{n} a_i\theta_i\log(\theta_i/(\theta_0 v_i)) - \sum_{i=1}^{n}\log\theta_i - n\log\theta_0 - \sum_{k=1}^{m'}\log(u_k - \sum_{i=1}^{n}A_{ki}\theta_i)\\ &= \sum_{i=1}^{n}(a_i t\theta_i\log(\theta_i/(\theta_0 v_i)) - \log\theta_i\theta_0) - \sum_{k=1}^{m'}\log(u_k - \sum_{i=1}^{n}A_{ki}\theta_i). \qquad (23)\end{aligned}$$

We show that Φ is a self-concordant barrier for (CMNL), that is, the function $\Psi_t(\boldsymbol{\theta})$ is self-concordant, convex and closed on the domain $\{\boldsymbol{\theta}\in\mathbb{R}^{n+1} : \boldsymbol{\theta} > 0,\ \mathbf{e}^\top\boldsymbol{\theta} = 1, \mathbf{A}\boldsymbol{\theta} < \mathbf{u}\}$. The terms of the first summation in (23) are self-concordant and convex by Theorem 2, since $a_i > 0, \forall i$. The function $-\log x$ is self-concordant and convex, and so are the terms of the second summation since both properties are preserved by composition with an affine function. Finally, Ψ_t is self-concordant and convex since both properties are preserved through addition. The function is closed on its domain since the barrier terms become infinite at the boundary, and all terms are bounded from below.

We observe that the level sets of (CMNL) are bounded, since the feasible set is bounded. We can now apply the result of Section 11.5.5 in (Boyd and Vandenberghe 2004) to show that no more than N_{I} Newton steps are required to solve a phase I problem yielding an initial strictly feasible point on the central path of an appropriate auxiliary phase II problem. The phase II problem may in turn be solved to within tolerance ϵ in at most N_2 Newton steps. The total number of Newton steps required to solve (CMNL) is thus at most $N_{\mathrm{I}} + N_{\mathrm{II}}$. □

4 Multiple Overlapping Customer Segments

In many cases, it is desirable to represent a number of customer segments, such as when business and leisure travelers are buying the same airline tickets. Suppose that different attraction demand models are available for each segment of the population. Regardless of her segment, a customer may purchase any of the products offered, but her choice probabilities depend on her particular segment. That is, we would like to be able to optimize over a demand model of the form

$$d_i^{\mathrm{MIX}}(\mathbf{x}) = \sum_{\ell=1}^{L}\Gamma_\ell d_i^\ell(\mathbf{x}) = \sum_{\ell=1}^{L}\Gamma_\ell\frac{f_i^\ell(x_i)}{1+\sum_{j=1}^{n} f_j^\ell(x_j)}, \qquad (24)$$

where the mixture coefficient Γ_ℓ represents the relative size of the ℓ^{th} market segment, whose demand is itself modeled by an attraction demand model. To continue representing demand as a fraction of the population, we assume that $\sum_{\ell=1}^{L} \Gamma_\ell = 1$, and $\Gamma_\ell > 0, \forall \ell$. We define the notation $d_0^\ell(\mathbf{x}), d_1^\ell(\mathbf{x}), \ldots, d_n^\ell(\mathbf{x})$ for the lost sales and demand functions of the ℓ^{th} segment, as in (1).

We point out that this model implicitly assumes that consumers from each segment may purchase *any* product. It is more general than the models from network revenue management that assume consumers only purchase products specific to their segment. Similarly, the work of Schön (2010) assumes that the retailer can set different prices for each of the segments. Both of these situations are better represented by standard attraction demand models, as discussed in Section 2.1.

On the other hand, the mixture of attraction demand models defined in (24) is *not* itself an attraction model. How to efficiently solve the pricing problem with multiple segments to optimality remains an open problem, and is beyond the scope of this paper. Hanson and Martin (1996) have shown that the pricing objective may have multiple local maxima, and solution methods from network revenue management give rise to NP-hard sub-problems (see, e.g., Miranda Bront et al (2009), who solve them heuristically). Instead, we propose an approximation to the multi-segment demand functions $d_i^{\text{MIX}}(\mathbf{x})$ by a valid attraction demand model.

4.1 Approximation by an Attraction Demand Model

Aydin and Porteus (2008) suggest (for the specific case of MNL models) the following approximation, based on valid attraction functions

$$\bar{f}_i(x_i) = \sum_{\ell=1}^{L} \gamma_\ell f_i^\ell(x_i), \qquad i = 1, \ldots, n, \tag{25}$$

where the coefficients $\gamma_1, \ldots, \gamma_L \in \mathbb{R}_+$ are set equal to the segment sizes Γ_ℓ of (24). We also introduce the notation $\bar{d}_1(\mathbf{x}), \ldots, \bar{d}_n(\mathbf{x})$ for the approximate demands when using the attraction functions (25). We define

$$\Pi^{\text{MIX}}(\mathbf{x}) = \sum_{i=1}^{n} a_i x_i d_i^{\text{MIX}}(\mathbf{x}) \qquad \text{and} \qquad \bar{\Pi}(\mathbf{x}) = \sum_{i=1}^{n} a_i x_i \bar{d}_i(\mathbf{x})$$

as the exact and approximated profit functions, respectively.

In Theorem 4, proved in Appendix A, we show that setting coefficients γ_ℓ as in (27) instead yields a local approximation to the desired multi-segment model (24) for prices near some reference point $\mathbf{x^0} \in \mathbb{R}^n$. In particular, our approximation is exact at the reference price $\mathbf{x} = \mathbf{x^0}$.

Theorem 4 *If the sets of attraction functions* $\left\{f_i^\ell, i = 1, \ldots, n\right\}$ *satisfy Assumption 1 for* $\ell = 1, \ldots, L$*, then so do the attraction functions* $\bar{f}_1, \ldots, \bar{f}_n$ *defined in* (25).

Furthermore, suppose that for some constant $B > 0$ and reference prices $\mathbf{x^0} \in \mathbb{R}^n$ the attraction functions satisfy the local Lipschitz conditions

$$|f_i^\ell(x_i) - f_i^\ell(x_i^0)| \leq B|x_i - x_i^0|, \quad \forall \mathbf{x} \in \mathcal{X}, \quad i = 1, \dots, n, \quad \ell = 1, \dots, L, \tag{26}$$

where $\mathcal{X} \subseteq \{\mathbf{x} : \|\mathbf{x} - \mathbf{x^0}\|_1 < 1/B\} \subset \mathbb{R}^n$ is a set around the reference prices $\mathbf{x^0}$. Let the coefficients of the approximation be

$$\gamma_\ell = \frac{\Gamma_\ell d_0^\ell(\mathbf{x^0})}{\sum_{\ell=1}^L \Gamma_\ell d_0^\ell(\mathbf{x^0})}, \qquad \ell = 1, \dots, L. \tag{27}$$

Then the approximate demand functions $\bar{d}_1, \dots, \bar{d}_n$ satisfy,

$$(1 - \epsilon_x) d_i^{MIX}(\mathbf{x}) \quad \leq \quad \bar{d}_i(\mathbf{x}) \quad \leq \quad (1 + \epsilon_x) d_i^{MIX}(\mathbf{x}), \qquad \forall x \in \mathcal{X},$$

where $\epsilon_x = \frac{2B\|\mathbf{x}-\mathbf{x^0}\|_1}{1-B\|\mathbf{x}-\mathbf{x^0}\|_1}$. Moreover, if the feasible prices are positive, i.e., $\mathcal{X} \subset \mathbb{R}_+^n$, the approximate profit function $\bar{\Pi}(\mathbf{x})$ satisfies

$$(1 - \epsilon_x) \Pi^{MIX}(\mathbf{x}) \quad \leq \quad \bar{\Pi}(\mathbf{x}) \quad \leq \quad (1 + \epsilon_x) \Pi^{MIX}(\mathbf{x}), \qquad \forall x \in \mathcal{X}.$$

An appropriate set $\mathcal{X}$ can be obtained by taking any set for which the Lipschitz condition (26) holds and restricting it to $\{\mathbf{x} : \|\mathbf{x} - \mathbf{x^0}\|_1 < 1/B\}$. Clearly, the accuracy of the approximation is highly dependent on the smoothness of the attraction functions. Such limitations are to be expected since the profit function $\Pi^{\text{MIX}}(\mathbf{x})$ may have multiple local maxima, while our approximation $\bar{\Pi}(\mathbf{x})$ cannot by Theorem 1 (for common demand models).

4.2 Solving the Approximated Pricing Problem

Although the approximation presented in the preceding section uses an attraction demand model, the corresponding (COP) formulations cannot be solved directly in practice even if Assumption 1 is satisfied. This is because, unlike for simpler attraction demand models, the attraction functions (25) do not have closed-form inverses. The same issue arises under other complex attraction demand models, such as the semi-parametric attraction models proposed by Hruschka (2002).

To solve (COP) using standard nonlinear optimization algorithms, we need to evaluate the gradient and the Hessian matrix of the objective function $\Pi(\boldsymbol{\theta})$ for any market shares $\boldsymbol{\theta}$. First, we note that the prices $\mathbf{x}$ corresponding to market shares $\boldsymbol{\theta}$ can be obtained efficiently. The equations (7) are equivalent to $f_i(x_i) = \theta_i/\theta_0, i = 1, \dots, n$. Since the attraction functions f_i are decreasing by Assumption 1, they have a unique solution and may be solved by n one-dimensional line searches. The following proposition, proved in Appendix A, then shows how the partial derivatives of the objective may be recovered from the prices $x_1, \dots, x_n$ corresponding to the market shares $\boldsymbol{\theta}$, and from the derivatives of the original attraction functions $f_1, \dots, f_n$. In particular, the derivatives of each $\bar{f}_i$ in (25) are readily obtained from those of $f_i^1, \dots, f_i^L$.

Proposition 2 *Let $x_1 \dots x_n$ be the unique prices solving the equations* (7) *for a given market share vector $\boldsymbol{\theta} \in \mathbb{R}_+^{n+1}$. Then the elements of the gradient $\nabla \Pi(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ of the objective function $\Pi(\boldsymbol{\theta})$ are*

$$\frac{\partial \Pi}{\partial \theta_0} = -\sum_{i=1}^{n} a_i \frac{\left(f_i(x_i)\right)^2}{f_i'(x_i)}, \qquad \text{and} \qquad \frac{\partial \Pi}{\partial \theta_i} = a_i \left(x_i + \frac{f_i(x_i)}{f_i'(x_i)} \right), \quad i = 1, \dots n.$$

The elements of the Hessian $\nabla^2 \Pi(\boldsymbol{\theta})$ are $\partial^2 \Pi / \partial \theta_i \partial \theta_j = 0$, for $1 \le i < j \le n$,

$$\frac{\partial^2 \Pi}{\partial \theta_0^2} = \sum_{i=1}^{n} \frac{a_i}{\theta_0} \left(\frac{2(f_i(x_i))^2}{f_i'(x_i)} - \frac{\left(f_i(x_i)\right)^3 f_i''(x_i)}{\left(f_i'(x_i)\right)^3} \right),$$

$$\frac{\partial^2 \Pi}{\partial \theta_i^2} = \frac{a_i}{\theta_0} \left(\frac{1}{f_i'(x_i)} - \frac{f_i(x_i) f_i''(x_i)}{\left(f_i'(x_i)\right)^3} \right), \qquad \text{for } i = 1, \dots, n,$$

and

$$\frac{\partial^2 \Pi}{\partial \theta_i \partial \theta_0} = -\frac{a_i}{\theta_0} \left(\frac{f_i(x_i)}{f_i'(x_i)} - \frac{(f_i(x_i))^2 f_i''(x_i)}{\left(f_i'(x_i)\right)^3} \right), \qquad \text{for } i = 1, \dots, n.$$

5 The Dual Problem

The structure of the pricing problem (P) goes beyond concavity of the transformed objective. Notice that the reformulation (COP) is not *separable* over the market shares $\theta_1, \dots, \theta_n$ only because of the occurrence of θ_0 in each term of the objective. Nevertheless, as is often the case with separable problems, its Lagrangian dual yields a tractable decomposition. The dual of (COP) is

$$\begin{aligned} \min \quad & \mu + \sum_{k=1}^{m'} \lambda_k u_k \\ \text{s.t.} \quad & \mu = \sum_{i=1}^{n} \max_{y_i > 0} \phi_i(y_i, \boldsymbol{\lambda}, \mu) \\ & \boldsymbol{\lambda} \ge 0 \end{aligned} \tag{DCOP}$$

where we define for $i = 1, 2, \dots, n$

$$\phi_i(y, \boldsymbol{\lambda}, \mu) \triangleq y \left(a_i g_i(y) - \sum_{k=1}^{m'} \lambda_k A_{ki} - \mu \right), \qquad y > 0. \tag{28}$$

The dual (DCOP) is expressed in terms of one-dimensional maximization problems for each product. These subproblems are coupled through a *single* linear constraint.

From a practical point of view, the dual problem does not require working with the inverse attractions, or with their derivatives, directly. A column generation algorithm for solving the dual is provided in Appendix D, along with the derivation of the dual itself. The algorithm requires the solution of a linear program and n one-dimensional maximization problems involving the original attraction functions f_i at each iteration (as opposed to their inverses g_i). It can be used even when the convexity condition (12) is not satisfied, or when the derivatives of the attraction functions described in Proposition 2 are not readily available. Moreover, Proposition 5 of Appendix D provides an alternate condition on the attraction functions which guarantees that a unique primal solution corresponds to each dual solution, without requiring the convexity condition (12).

For the special case of the MNL demand model, the inner maximization problems can in fact be solved in closed form, yielding the following dual problem.

$$\begin{aligned} \min \quad & \mu + \sum_{k=1}^{m'} \lambda_k u_k \\ \text{s.t.} \quad & \mu \geq \sum_{i=1}^{n} a_i v_i \exp\left\{ -1 - \frac{\sum_{k=1}^{m'} \lambda_k A_{ki} + \mu}{a_i} \right\} \\ & \boldsymbol{\lambda} \geq 0 \end{aligned} \tag{DMNL}$$

The minimization (DMNL) is a convex optimization problem. As a result, it can also be solved with general-purpose algorithms. However, our experiments in the next section suggest this is less efficient than solving the primal (COP) directly.

6 Computational Experiments

We have proposed three formulations of the pricing problem in (P), (COP) and (DCOP). First, we have shown that under certain conditions, any local maximum of the non-convex problem (P) in terms of the prices is also a global maximum. Second, under the same conditions, the equivalent problem (COP) is a convex optimization problem with linear constraints. Third, we can recover a solution from the dual problem (DCOP).

To compare the efficiency of the three formulations, we evaluate the solution times of instances with an MNL demand model using the commercial LOQO solver (see Vanderbei and Shanno (1999); Shanno and Vanderbei (2000)). This solver uses a primal-dual interior point algorithm for sequential quadratic programming. It was chosen because it is commercially available and is intended for both convex and non-convex problems. However, LOQO does *not* employ the barrier method analyzed in Section 3.2. The AMPL (see Fourer et al (1990)) modeling language provides automatic differentiation for all problems. All the experiments were run on computer with dual 2.83GHz Intel Xeon CPUs and 32GB of RAM.

The MNL demand model was chosen since it allows (DCOP) to be solved directly. The demand model parameters are sampled as described in Appendix B.1 to ensure that the aggregate demand is near 0 and 1 as the prices approach

Products	Constraints	Price Formulation (P)		Market Share Form. (COP)		Dual Formulation (DCOP)	
(n)	(m)	Iterations	Time	Iterations	Time	Iterations	Time
2	256	7	0.02	15	0.00	19	0.28
4	256	7	0.03	15	0.01	20	0.34
8	256	10	0.12	17	0.01	22	0.50
16	256	16	0.83	24	0.03	28	0.99
32	256	21	4.22	26	0.07	25	1.68
64	256	24	17.48	25	0.16	26	4.28
128	256	27	75.69	29	0.64	25	13.30
256	256	24	265.10	29	1.90	24	55.77
512	256	25	1,307.78	34	4.38	33	451.61
1,024	256	27	5,181.45	34	9.03	27	2,346.63
2,048	256	29	22,818.30	36	20.66	33	19,560.10
4,096	256	38	123,123.50	38	49.09	-	-
256	2	19	1.63	26	0.04	28	28.72
256	4	16	2.13	26	0.04	30	30.42
256	8	19	4.38	28	0.05	28	28.80
256	16	23	14.95	27	0.07	29	30.84
256	32	21	31.09	30	0.11	27	31.10
256	64	21	59.52	29	0.21	38	49.54
256	128	20	114.26	29	0.56	26	41.40
256	256	24	265.10	29	1.90	24	55.77
256	512	36	964.53	36	5.26	28	122.72
256	1,024	50	2,411.04	36	8.83	31	338.77
256	2,048	60	5,549.19	45	22.95	35	1,226.33
256	4,096	69	13,003.43	53	59.36	43	5,918.61

Table 1 Number of iterations and solution time in seconds as a function of the number of products and constraints for the three problem formulations. (Averages over 10 randomly generated instances.)

the bounds $\overline{x}_i$ and $\underline{x}_i = 0$, respectively, for each product i. Constraints are sampled uniformly from the tangents to the sphere of radius $\frac{1}{2} \cdot \frac{1}{n+1}$ centered at the uniform distribution $\theta_0 = \theta_1 = \ldots = \theta_n = \frac{1}{n+1}$. Specifically, the k^{th} constraint is defined by

$$\mathbf{z}_k^\top \left(\boldsymbol{\theta} - \frac{1}{n+1} \mathbf{e} \right) \leq \frac{1}{2} \cdot \frac{1}{n+1} \qquad \Leftrightarrow \qquad \mathbf{z}_k^\top \boldsymbol{\theta} \leq \frac{1}{n+1} \left(\frac{1}{2} + \mathbf{z}_k^\top \mathbf{e} \right), \quad (29)$$

where $\mathbf{z}_k$ is sampled uniformly from the unit sphere centered at the origin, and $\mathbf{e}$ is the vector of all ones. This choice ensures that we do not generate any redundant constraints, and that a number of constraints are likely to be active at optimality.

Table 1 shows the average number of iterations and the average solution times over 10 randomly generated instances of various size when solving each of the three formulations. We note that for the market share formulation (COP) and the dual formulation (DCOP), the price bounds on x_i are converted to linear constraints by replacing $d_i(\mathbf{x})$ with θ_i in equation (10) of Lemma 1, yielding a total of $m' = m + 2n$ constraints. However, the additional $2n$ constraints are sparse and we do not expect them to be active at optimality. In contrast, using (11) would result in dense constraints for (COP). Default parameters are used for the LOQO solver except that the tolerance is reduced from 8 to 6 significant digits of agreement between the primal and dual solution, and a parameter governing the criteria used to declare problems infeasible is relaxed to prevent premature termination (we set `inftol2=100`). These adjustments are necessary since convergence is sometimes very slow in the first few iterations, and again after five or six digits or accuracy have been achieved.

As is generally the case with interior point methods, LOQO terminates in a moderate number of iterations for all the instances, but there is significant variability in the time per iteration. Examining first the results for the price formulation (P), we observe that the total solution time increases rapidly with both the number of products and the number of constraints. With 4,096 products and 256 constraints, approximately 34 hours of computation time are needed. The solution time scales somewhat better with the number of constraints, but 3.6 hours are still needed with 4,096 constraints and only 256 products. We cannot theoretically guarantee convergence when solving (P) in general. Nevertheless, the optimal solution is eventually found in all the instances we considered, with the chosen parameters.

In contrast, the market share formulation we introduced, even for the largest instances we considered, is solved in about one minute, and is about 2,500 times faster than the price formulation in the most extreme case. We believe this difference may be due to the sparsity of the Hessian matrix of the objective function and the linearity of the constraint in (COP). In the price formulation (P), the Hessian of the Lagrangian is dense since each term of both the constraints and the objective depends nonlinearly on *all* the variables. The non-convexity of the constraints and objective in (P) may also be the reason for the slow performance, though we would expect this effect to increase the number of iterations rather than the time per iteration, which does not appear to be the case.

Finally, the dual formulation (DCOP) is observed to be much slower for a large number of products than for a large number of constraints. In fact, the specific dual formulation (DMNL) has a *single* constraint involving a summation over all n products with nonlinear terms. The number of terms in the sum increases with the number of products, but adding constraints in the primal problem (i.e., increasing m) only increases the number of dual variables λ_i, which are zero unless the constraint in question is active. We note that we were not able to obtain a solution for the (DCOP) instances with $n = 4,096$ and $m = 256$ because AMPL required an excessive amount of memory. This may be due to the fact that the actual number of primal constraints (corresponding to the number of dual variables) is not $m = 256$, but $m' \triangleq m + 2n = 8,448$ when including the converted price bounds. Then the Hessian of the dual constraint is a dense matrix with $(m')^2$ entries. It seems likely that a more efficient solution approach is possible for the dual problem, since most of the dual variables are zero at optimality (i.e., most of the constraints are inactive). Indeed, an efficient algorithm tailored to the special case of Proposition 1 is possible, but we see limited interest in pursuing this approach for the MNL demand model since the primal problem (COP) can be solved efficiently.

6.1 Approximation to Multiple Segment Demand Models

To illustrate the performance of the algorithm for solving the multiple-segment approximation, Table 2 shows the number of iterations and the running time needed to solve instances of varying size. For each evaluation of the gradient and Hessian, the equations in (7) are solved with Brent's method (see Brent (1973)) and the derivatives in Proposition 2 are computed.

Segments	Products	Constraints	(P)		(COP) with demand model approximation			
(k)	(n)	(m)	Iterations	Time (sec.)	Iterations	Time (sec.)	Error	Infeasibility
2	256	256	23	601.94	58	4.73	0.57%	1.2E-03
	1024	256	26	10,812.00	46	11.94	2.22%	2.3E-03
	4096	256	19	165,090.00	35	46.11	0.56%	3.4E-04
	256	256	23	601.94	58	4.73	0.57%	1.2E-03
	256	1024	39	4,066.70	67	16.46	0.05%	2.4E-04
	256	4096	45	18,741.00	50	65.31	0.03%	1.3E-04
4	256	256	27	1,400.30	48	3.96	0.51%	1.0E-03
	1024	256	27	25,549.00	73	18.51	2.27%	1.9E-03
	4096	256	-	-	33	44.67	-	3.5E-04
	256	256	27	1,400.30	48	3.96	0.51%	1.0E-03
	256	1024	47	9,705.20	52	13.05	0.02%	1.0E-04
	256	4096	53	43,809.00	59	74.96	0.03%	1.5E-04
8	256	256	36	3,787.10	37	3.12	0.55%	1.4E-03
	1024	256	37	77,454.00	68	17.53	2.02%	2.4E-03
	4096	256	-	-	38	49.79	-	2.6E-04
	256	256	36	3,787.10	37	3.12	0.55%	1.4E-03
	256	1024	58	24,259.00	56	13.98	0.01%	6.5E-05
	256	4096	-	-	57	72.80	-	5.5E-05
16	256	256	48	11,818.00	37	3.21	0.55%	1.4E-03
	1024	256	68	298,770.00	59	15.77	1.86%	1.9E-03
	4096	256	-	-	42	53.79	-	2.9E-04
	256	256	48	11,818.00	37	3.21	0.55%	1.4E-03
	256	1024	82	80,675.00	55	13.90	0.02%	5.9E-05
	256	4096	-	-	86	103.45	-	2.4E-05

Table 2 Number of iterations and solution time in seconds to solve the exact, non-convex multi-segment pricing problem in terms of prices, as well as the convex multi-segment approximation in terms of market shares.

We observe that the solution times for (COP) are comparable to those for the single-segment instances, since the computational cost is dominated by the optimization algorithm rather than by the function evaluations. Indeed, the number of distinct segments only impact the time needed to evaluate the objective, and we observe that increasing the number of segments does not increase the solution time significantly.

In contrast, solving the price formulation (P) with the exact demands $d_i^{\mathrm{MIX}}(\mathbf{x})$ of (24) takes significantly longer with multiple segments. For a given problem size, doubling the number of segments more than doubles the solution time. The largest successfully solved instance took over three days (298,770 seconds) to solve, compared to only 15.77 seconds for the (COP) formulation with the approximate demand model.

For our experiments, the reference price $\mathbf{x^0}$ and the coefficients γ_ℓ in (27) were chosen such that the demand model approximation was exact at the uniform demand distribution, $\theta_i = \bar{d}_i(\mathbf{x^0}) = d_i^{\text{MIX}}(\mathbf{x^0}) = 1/(n+1)$, for $i = 0, \ldots, n$. Thus the demand model approximation is inexact at the computed optimum. The second-last column in Table 2 shows the error in the objective value relative to the solution obtained by solving (P) with the exact demand model. The solution to (COP) overestimates the true maximum by at most 2.27% in our experiments, and generally by much less. It is somewhat unsurprising that the maximum of the approximation (COP) exceeds the true maximum, because this occurs whenever the approximation exceeds the true maximum of (P) at *any* feasible point. The rightmost column of Table 2 shows the maximum absolute constraint violation for each instance. The constraint violations are small despite the approximation, since the right hand side of the constraints (29) is on the order of $1/(n+1)$. For reference, the tolerance used for the solver is 10^{-6}. We remark that a different choice of the reference point $\mathbf{x^0}$ may improve the accuracy at the optimum. Although it is unclear how best to select this parameter of the approximation *a priori*, the accuracy of the approximation and the constraint violations at the computed optimal value can easily be checked empirically once the solution has been obtained.

7 Conclusions

We have developed an optimization framework for solving a large class of constrained pricing problems under the important class of attraction demand models. Our formulations incorporate a variety of constraints which naturally occur in numerous problems studied in the literature. They provide increased representation power for typical revenue management settings such as airline, hotel and other booking systems. Moreover, they capture problems such as product line pricing, or joint inventory and pricing problems where capacity constraints alone may not be sufficient.

We provided a condition on the demand model guaranteeing that our formulations are convex optimization problems with linear constraints. It is satisfied by MNL, MCI and linear attraction demand models, in particular. We further proved that the pricing problem can be solved in polynomial time under MNL demand models using interior point methods. Our computational experiments show that our new formulations can be solved orders of magnitude faster than naive formulations, using commercially available software. The efficiency of the solutions suggests our models may be effectively adapted for use in multi-period stochastic pricing problems, where they promise to increase modeling power at reasonable computational cost.

Furthermore, we proposed an approximation to the demand encountered when there are multiple overlapping market segments. Such scenarios are an active research topic in the closely related area of network revenue management. We provided a bound on the approximation error, and showed that the resulting pricing problems can be solved using standard nonlinear optimization algorithms despite the lack of a closed-form objective function. Our approximation represents a new way to approach pricing in the presence of multiple overlapping market segments, and provides an efficient way to solve certain instances.

Acknowledgements The authors would like to thank the Associate Editor and two anonymous referees for their valuable comments. They helped improve both the content and exposition of this work. Preparation of this paper was partially supported, for the second author by grants CMMI-0846554 (CAREER Award) and DMS-0732175 from the National Science Foundation, AFOSR awards FA9550-08-1-0369 and FA9550-08-1-0369, an SMA grant and the Buschbaum Research Fund from MIT and for the third author by the CMMI-0758061 Award, the EFRI-0735905 Award and the CMMI-0824674 Award from the National Science Foundation and an SMA grant.

References

Ahipasaoglu SD, Sun P, Todd MJ (2008) Linear convergence of a modified Frank-Wolfe algorithm for computing minimum-volume enclosing ellipsoids. Optimization Methods Software 23(1):5–19

Aydin G, Porteus EL (2008) Joint inventory and pricing decisions for an assortment. Operations Research 56(5):1247–1255

Bertsekas DP (1999) Nonlinear Programming, 2nd edn. Athena Scientific

Bitran G, Caldentey R (2003) An overview of pricing models for revenue management. Manufacturing Service Operations Management 5(3):203–229

Boyd JH, Mellman RE (1980) The effect of fuel economy standards on the U.S. automotive market: An hedonic demand analysis. Transportation Research Part A: General 14(5-6):367 – 378

Boyd S, Vandenberghe L (2004) Convex Optimization. Cambridge University Press

Brent R (1973) Algorithms for Minimization without Derivatives. Prentice-Hall, Englewood Cliffs, NJ

Calafiore GC (2007) Ambiguous risk measures and optimal robust portfolios. SIAM J on Optimization 18(3):853–877

Cardell NS, Dunbar FC (1980) Measuring the societal impacts of automobile downsizing. Transportation Research Part A: General 14(5-6):423 – 434

Dong L, Kouvelis P, Tian Z (2008) Dynamic pricing and inventory control of substitute products. Manufacturing and Service Operations Management

Elmaghraby W, Keskinocak P (2003) Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. Management Science 49(10):1287–1309

Fourer R, Gay DM, Kernighan BW (1990) A modeling language for mathematical programming. Management Science 36(5):519–554

Gallego G, van Ryzin G (1997) A multiproduct dynamic pricing problem and its applications to network yield management. Operations Research 45(1):24–41

Gallego G, Stefanescu C (2009) Upgrades, upsells and pricing in revenue management Under review.

Gallego G, Iyengar G, Phillips R, Dubey A (2004) Managing flexible products on a network. Tech. rep., CORS Technical Report TR-2004-01

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

den Hertog D, Jarre F, Roos C, Terlaky T (1995) A sufficient condition for self-concordance, with application to some classes of structured convex programming problems. Mathematical Programming 69(1):75–88

Hruschka H (2002) Market share analysis using semi-parametric attraction models. European Journal of Operational Research 138(1):212 – 225

Liu Q, van Ryzin G (2008) On the choice-based linear programming model for network revenue management. Manufacturing Service Operations Management 10(2):288–310

McFadden D (1974) Conditional logit analysis of qualitative choice behavior. Frontiers in Econometrics pp 105–142., Zarembka, P. (Ed.)

Mcfadden D, Train K (2000) Mixed MNL models for discrete response. Journal of Applied Econometrics 15:447–470

Miranda Bront JJ, Mendez-Diaz I, Vulcano G (2009) A column generation algorithm for choice-based network revenue management. Operations Research 57(3):769–784

Nakanishi M, Cooper LG (1982) Simplified estimation procedures for mci models. Marketing Science 1(3):314–322

Schaible S (1974) Parameter-free convex equivalent and dual programs of fractional programming problems. Mathematical Methods of Operations Research 18:187–196

Schön C (2010) On the product line selection problem under attraction choice models of consumer behavior. European Journal of Operational Research 206(1):260 – 264

Shanno DF, Vanderbei RJ (2000) Interior-point methods for nonconvex nonlinear programming: orderings and higher-order methods. Mathematical Programming 87(2):303–316

Song JS, Xue Z (2007) Demand management and inventory control for substitutable products Working paper.

Talluri KT, van Ryzin GJ (1998) An analysis of bid-price controls for network revenue management. Management Science 44(11-Part-1):1577–1593

Talluri KT, van Ryzin GJ (2004) The Theory and Practice of Revenue Management. Springer

Urban GL (1969) A mathematical modeling approach to product line decisions. Journal of Marketing Research 6(1):40–47

Vanderbei RJ, Shanno DF (1999) An interior-point algorithm for nonconvex nonlinear programming. Computational Optimization and Applications 13(1):231–252

A Approximation for Multiple Customer Segments

Proof of Theorem 4. Assumption 1 holds since (i) the sum of decreasing functions is decreasing and the sum of differentiable functions is differentiable, and (ii) the limit of a finite sum is the sum of the limits.

Since from the choice of coefficients $\sum_{\ell=1}^{\ell} \gamma_\ell = 1$, we rewrite

$$\bar{d}_i(\mathbf{x}) = \frac{\bar{f}_i(x_i)}{1+\sum_{j=1}^n \bar{f}_j(x_j)} = \frac{\sum_{\ell=1}^L \gamma_\ell f_i^\ell(x_i)}{1+\sum_{j=1}^n \sum_{\ell=1}^L \gamma_\ell f_j^\ell(x_j)} = \frac{\sum_{\ell=1}^L \gamma_\ell f_i^\ell(x_i)}{\sum_{\ell=1}^L \gamma_\ell \left(1+\sum_{j=1}^n f_j^\ell(x_j)\right)}$$

$$= \frac{\sum_{\ell=1}^L \Gamma_\ell d_0^\ell(\mathbf{x^0}) \left(d_i^\ell(\mathbf{x})/d_0^\ell(\mathbf{x})\right)}{\sum_{\ell=1}^L \Gamma_\ell d_0^\ell(\mathbf{x^0}) \left(1/d_0^\ell(\mathbf{x})\right)} = \sum_{\ell=1}^L \frac{\left(d_0^\ell(\mathbf{x^0})/d_0^\ell(\mathbf{x})\right)}{\sum_{\ell=1}^L \Gamma_\ell \left(d_0^\ell(\mathbf{x^0})/d_0^\ell(\mathbf{x})\right)} \Gamma_\ell d_i^\ell(\mathbf{x}) \tag{30}$$

where we use fact (5) in the fourth equality. The ratios appearing in the last expression can be expressed as

$$\frac{d_0^\ell(\mathbf{x^0})}{d_0^\ell(\mathbf{x})} = \frac{1+\sum_{i=1}^n f_i^\ell(x_i)}{1+\sum_{i=1}^n f_i^\ell(\overline{x}_i)} = 1 + d_0^\ell(\mathbf{x^0}) \sum_{i=1}^n \left(f_i^\ell(x_i) - f_i^\ell(\overline{x}_i)\right)$$

where $d_0^\ell(\mathbf{x^0}) < 1$. Then, using assumption (26), we obtain

$$1 - B\|\mathbf{x} - \mathbf{x^0}\|_1 \leq \frac{d_0^\ell(\mathbf{x^0})}{d_0^\ell(\mathbf{x})} \leq 1 + B\|\mathbf{x} - \mathbf{x^0}\|_1.$$

Note that the lower bound is non-negative by the definition of $\mathcal{X}$. Since $\sum_{\ell=1}^{\ell} \Gamma_\ell = 1$, we obtain from (30)

$$\bar{d}_i(\mathbf{x}) \leq \frac{1 + B\|\mathbf{x} - \mathbf{x^0}\|_1}{1 - B\|\mathbf{x} - \mathbf{x^0}\|_1} \sum_{\ell=1}^{\ell} \Gamma_\ell d_i^\ell(\mathbf{x}) = \frac{1 + B\|\mathbf{x} - \mathbf{x^0}\|_1}{1 - B\|\mathbf{x} - \mathbf{x^0}\|_1} d_i^{\mathrm{MIX}}(\mathbf{x})$$

$$\bar{d}_i(\mathbf{x}) \geq \frac{1 - B\|\mathbf{x} - \mathbf{x^0}\|_1}{1 + B\|\mathbf{x} - \mathbf{x^0}\|_1} \sum_{\ell=1}^{\ell} \Gamma_\ell d_i^\ell(\mathbf{x}) = \frac{1 - B\|\mathbf{x} - \mathbf{x^0}\|_1}{1 + B\|\mathbf{x} - \mathbf{x^0}\|_1} d_i^{\mathrm{MIX}}(\mathbf{x})$$

Defining shorthand $a = B\|\mathbf{x} - \mathbf{x^0}\|_1 < 1$, notice that

$$\frac{1+a}{1-a} = 1 + \frac{2a}{1-a} = 1 + \epsilon_x, \qquad \text{and} \qquad \frac{1-a}{1+a} = 1 - \frac{2a}{1+a} \geq 1 - \frac{2a}{1-a} = 1 - \epsilon_x.$$

The statement regarding the profit function follows immediately if the prices are also positive, by bounding each term of $\bar{\Pi}(\mathbf{x})$ individually. □

Proof of Proposition 2. The quantities in the statement are obtained by summing

$$\nabla \Pi(\boldsymbol{\theta}) = \sum_{i=1}^n a_i \nabla \Pi_i(\boldsymbol{\theta}) \qquad \text{and} \qquad \nabla^2 \Pi(\boldsymbol{\theta}) = \sum_{i=1}^n a_i \nabla^2 \Pi_i(\boldsymbol{\theta}),$$

with the non-zero elements of the terms $\nabla \Pi_i(\boldsymbol{\theta})$ and $\nabla^2 \Pi_i(\boldsymbol{\theta})$ given in (14) and (15). (By a slight abuse of notation, we now consider the terms $\Pi_i(\boldsymbol{\theta})$ to be functions of the entire market share vector $\boldsymbol{\theta}$ instead of only the variables θ_0 and θ_i on which they each depend.) Then, we substitute in $f_i(x_i)$ and its first and second derivatives using (16) and (17). □

B Background on Attraction Demand Models

The class of attraction demand models subsumes a number of important customer choice models by retaining only their fundamental properties. Namely, the form of the demands (1) ensures that they are positive and sum to one. A related feature is the well known *independence from irrelevant alternatives* (IIA) property which implies that the demand lost from increasing the price of one product is distributed to other alternatives proportionally to their initial demands.

The attraction functions $f_i(\cdot), i = 1, \ldots, n$ may depend on a number of product attributes in general, but we limit our attention to the effect of price. The requirements of Assumption 1 are mild. The positivity assumption and (i) imply that demand for a product is smoothly decreasing in its price but always positive. The requirement (ii) implies that the demand grows to 1 if the price is sufficiently negative, and ensures that increasing the price eventually becomes unprofitable for a seller. As we demonstrate for specific instances below, if the latter two assumptions are not satisfied, the attraction functions can be suitably modified.

Though the class of attraction demand models is very general, certain instances are well studied and admit straightforward estimation methods to calibrate their parameters. This is the case for the MNL and MCI demand models (McFadden 1974; Nakanishi and Cooper 1982). On the other hand, if assumptions have been made on customers responses to price changes, appropriate attraction functions can be defined to model the desired behavior. Examples of this approach include the linear attraction demand model, and the mixtures of attraction functions discussed in Section 4.

B.1 The multinomial logit (MNL) demand model

The MNL demand model is a discrete choice model founded on utility theory, where $d_i(\mathbf{x}_i)$ is interpreted as the *probability* that a utility-maximizing consumer will elect to purchase product i. The utility a customer derives from buying product i is $U_i = V_i + \epsilon_i$ whereas making no purchase is has utility $U_0 = \epsilon_0$. The V_i terms are deterministic quantities depending on the product characteristics (including price) and the random variables ϵ_i are independent with a standard Gumbel distribution. It can be shown that the probability of product i having the highest realized utility is then in fact given by $d_i(\mathbf{x})$ in (1), with $f_i(x_i)$ replaced by e^{V_i}. To model the impact of pricing, we let, for each alternative $i = 1, \ldots, n$,

$$V_i \triangleq V_i(\hat{x_i}) = \beta_{0,i} - \beta_{1,i}\hat{x_i}, \tag{31}$$

where $\beta_{0,i} > 0$ represents the quality of product i and $\beta_{1,i} > 0$ determines how sensitive a customer is to its price, denoted here by $\hat{x}_i$. When there is a population of consumers with independent utilities, the fractions $d_i(\mathbf{x})$ represent the portion of the population opting for each product in expectation. For ease of notation, we re-scale the true price $\hat{x}_i$ by $\beta_{1,i}$ to obtain the single-parameter attraction functions (2), with $v_i = e^{\beta_{0,i}}$ and $x_i = \beta_{1,i}\hat{x_i}$, rather than using the form of the exponents (31) directly. These functions clearly satisfy Assumption 1.

Parameters for the demand model used in the experiments of Section 6 are generated by sampling the mean linear utilities $V_i(\hat{x}_i)$ in equation (31) for each product i. Specifically, $V_i(0)$ and $V_i(x_{max})$ are chosen uniformly over $[2\sigma, 4\sigma]$ and $[-4\sigma, -2\sigma]$ respectively, where $\sigma = \pi/\sqrt{6}$ is the standard deviation of the random Gumbel-distributed customer utility terms ϵ_i. The parameters $\beta_{0,i}$ and $\beta_{1,i}$ are set accordingly. Recall that the mean utility of the outside alternative is fixed at $V_0 = 0$. The choice of parameters thus ensures that purchasing each product is preferred with large probability when its price is set to 0, and that no purchase is made with large probability when the (unscaled) prices $\hat{x}_i$ are near x_{max}.

B.2 The multiplicative competitive interaction (MCI) demand model

Another common choice of attraction functions is Cobb-Douglas attraction functions $\hat{f}_i(x_i) = \alpha_i x_i^{-\beta_i}$, with parameters $\alpha_i > 0$ and $\beta_i > 1$. It yields the *multiplicative competitive interaction* (MCI) model. Since the attraction is not defined for negative prices, we use its linear extension below a small price ϵ. Let

$$f_i(x_i) = \begin{cases} \alpha_i \epsilon^{-\beta_i} - (x_i - \epsilon)\alpha_i \beta_i \epsilon^{-\beta_i - 1} & \text{if } x_i < \epsilon, \\ \alpha_i x_i^{-\beta_i} & \text{otherwise.} \end{cases} \tag{32}$$

This is a mathematical convenience, since one would expect problems involving MCI demand to enforce positivity of the prices. The approximation can be made arbitrarily precise by reducing ϵ.

B.3 The linear attraction demand model

This demand model approximates a linear relationship between prices and demands, while ensuring that the demands remain positive and sum to less than one. The attraction function for the i^{th} product is $\hat{f}_i(x_i) = \alpha_i - \beta_i x_i$, with parameters $\alpha_i, \beta_i > 0$. An appropriate extension is needed to ensure positivity. For instance, by choosing the upper bound $\bar{x_i} = \alpha_i/\beta_i - \epsilon$, the following attraction function satisfies our assumptions:

$$f_i(x_i) = \begin{cases} \alpha_i - \beta_i x_i & \text{if } x_i \leq \overline{x}_i, \\ \beta_i e^{-(x_i - \overline{x}_i)/\epsilon} & \text{otherwise.} \end{cases} \tag{33}$$

C Pricing under Attraction Demand Models

C.1 Non-Convexity of the Naive Pricing Problem under MNL Demand

This sections illustrates why the pricing problem (P) is difficult to solve directly in terms of prices, as claimed in Section 2. Figure 1 shows the profit in terms of the prices under an MNL demand model when the number of products is $n = 1$ and $n = 2$. The dashed line in the first plot shows the demand as a function of the price. The profit function is not concave even for a single product. With multiple products, the level sets of the objective are not convex, i.e., the objective is not even quasi-concave.

Furthermore, combining nonlinear constraints with a non-quasi-concave objective function introduces additional complications. First, it is easy to see that, because the objective is not quasi-concave, even a linear inequality constraint in terms of prices could exclude the global maximum in the right panel of Figure 1, and thus give rise to a local maximum on each of the ridges leading to the peak. Secondly, the feasible region of (P) is in general not convex. Figure 2 illustrates the constraints of problem (P) with data

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ 1 & -\frac{1}{2} \end{bmatrix}, \qquad \mathbf{u} = \begin{bmatrix} 0.6 \\ -0.4 \\ 0.4 \end{bmatrix}.$$

The left panel shows the feasible region in terms of the prices, and the right panel shows the polyhedral feasible region in terms of the demands. Observe that the last two constraints are clearly non-convex in the space of prices. On the other hand, the first constraint happens to belong to the class of convex constraints characterized by Proposition 1.

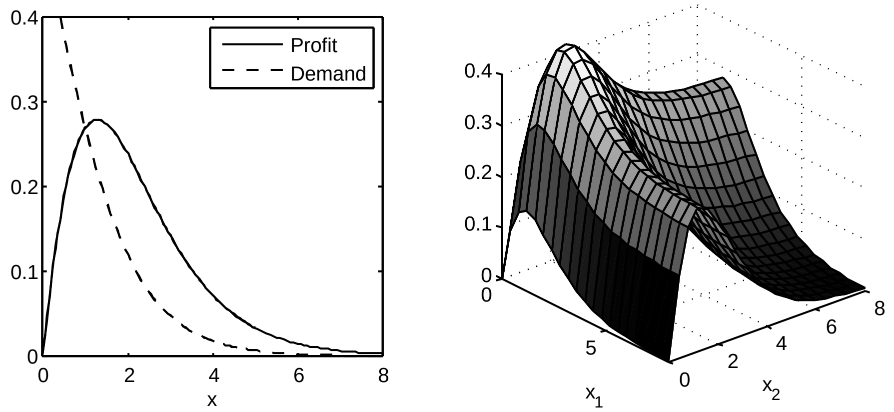

Fig. 1 The objective functions of (P) with a single product (left) and two-products (right).

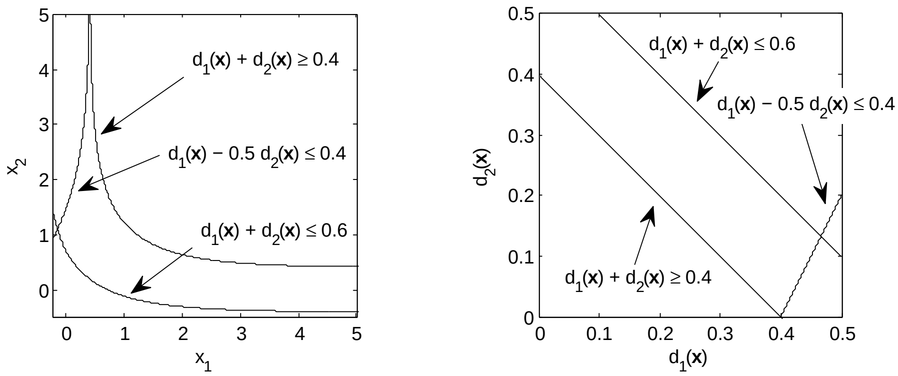

Fig. 2 The feasible region of (P) with two products and three constraints on the demands.

C.2 Representing Joint Price Constraints

This section shows how to incorporate certain joint price constraints into the formulations we have proposed. Under MNL demand models, it is natural to assume that the consumer's utility (31) is equally sensitive to the price regardless of the alternative she considers. That is, $\beta_{1,i} = \beta_{1,j}$. Then the constraint (4) can be expressed as

$$d_i(\mathbf{x}) \leq \frac{f_i(x_j + \delta_{ij})}{f_j(x_j)} d_j(\mathbf{x}) = \frac{v_i f_j(x_j + \delta_{ij})}{v_j f_j(x_j)} d_j(\mathbf{x}) = \left(\frac{v_i}{v_j} e^{-\delta_{ij}} \right) d_j(\mathbf{x}).$$

The assumption regarding the sensitivity to price is required so that the same scaling described in Appendix B.1 is used to relate x_i and x_j with $\hat{x}_i$ and $\hat{x}_j$ of equation (31), respectively. This allows f_i to be replaced with f_j in the preceding equation. The resulting constraint is evidently linear in terms of the demands and is captured by the formulation (P). This transformation depends on the relationship between the attraction functions for different products and is thus specific to the MNL model. A similar transformation is possible for the linear attraction demand model with the analogous uniform price sensitivity assumption, $\beta_i = \beta_j$ in (33). From (4), we then have

$$f_i(x_i)d_0(\mathbf{x}) \le f_i(x_j + \delta_{ij})d_0(\mathbf{x}) \quad \Leftrightarrow$$
$$d_i(\mathbf{x}) \le (\alpha_i - \beta_i(x_j + \delta_{ij}))d_0(\mathbf{x}) = (\alpha_i - \alpha_j - \beta_j\delta_{ij})d_0(\mathbf{x}) + d_j(\mathbf{x}),$$

where the $d_0(\mathbf{x})$ terms can be substituted out using the simplex constraint (8).

C.3 Convexity of (COP) Under Common Attraction Demand Models

Corollary 1 *Under the linear, MNL and MCI attraction demand models, the objective of* (COP) *is a concave function and any local maximum of either* (COP) *or* (P) *is also a global maximum.*

Proof For each model, we verify the condition (12). For the MNL model (2), we have

$$g_i(y) = -\log\frac{y}{v_i},\; g_i'(y) = \frac{-1}{y},\; g_i''(y) = \frac{1}{y^2},\text{ and } 2g_i'(y) + yg_i''(y) = \frac{-2}{y} + \frac{y}{y^2} = \frac{-1}{y} < 0.$$

Now consider the attractions (33) for the linear model. For $x > \overline{x}_i$ (i.e., $y < f_i(\overline{x}_i)$) we have the MNL attraction function so the condition (12) holds as shown above. Elsewhere, when $x \le \overline{x}_i$,

$$g_i(y) = \frac{\alpha_i - y}{\beta_i}, \quad g_i'(y) = \frac{-1}{\beta_i}, \quad g_i''(y) = 0, \quad \text{and} \quad 2g_i'(y) + yg_i''(y) = \frac{-2}{\beta_i} + 0 = \frac{-2}{\beta_i} < 0.$$

as desired. For the MCI attraction functions (32), we have the linear attraction function for $x_i < \epsilon$, otherwise

$$g_i(y) = \left(\frac{y}{\alpha_i}\right)^{\frac{-1}{\beta_i}}, \quad g_i'(y) = \frac{-1}{\alpha_i\beta_i}\left(\frac{y}{\alpha_i}\right)^{\frac{-1}{\beta_i}-1}, \quad g_i''(y) = \frac{-1}{\alpha_i^2\beta_i}\left(\frac{-1}{\beta_i} - 1\right)\left(\frac{y}{\alpha_i}\right)^{\frac{-1}{\beta_i}-2},$$

$$\begin{aligned} 2g_i'(y) + yg_i''(y) &= \frac{-2}{\alpha_i\beta_i}\left(\frac{y}{\alpha_i}\right)^{\frac{-1}{\beta_i}-1} + y\frac{-1}{\alpha_i^2\beta_i}\left(\frac{-1}{\beta_i} - 1\right)\left(\frac{y}{\alpha_i}\right)^{\frac{-1}{\beta_i}-2} \\ &= \left(-2 + \frac{1}{\beta_i} + 1\right)\frac{1}{\alpha_i\beta_i}\left(\frac{y}{\alpha_i}\right)^{\frac{-1}{\beta_i}-1} = \left(\frac{1}{\beta_i} - 1\right)\frac{1}{\alpha_i\beta_i}\left(\frac{y}{\alpha_i}\right)^{\frac{-1}{\beta_i}-1} < 0 \end{aligned}$$

where the inequality uses that $\beta_i > 1$. So the condition (12) is also satisfied. □

D The Dual Market Share Problem

Proposition 3 *The dual of* (COP) *is given by* (DCOP). *For any $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\mu \in \mathbb{R}$, there exist optimal $y_i^* > 0$, for $i = 1, \ldots, n$, so that $\phi_i(y_i^*, \boldsymbol{\lambda}, \mu) > 0$ in each of the inner maximization problems that appear in the equality constraint of* (DCOP). *Furthermore, when condition* (12) *(or equivalently, condition* (13)*) is satisfied and $(\boldsymbol{\lambda}, \mu)$ is an optimal solution of* (DCOP), *a primal optimal solution $\boldsymbol{\theta}^*$ of* (COP) *is given by*

$$\theta_0^* = \frac{1}{1 + \sum_{i=1}^n y_i^*}, \qquad \textit{and} \qquad \theta_i^* = \frac{y_i^*}{1 + \sum_{i=1}^n y_i^*}, \quad i = 1, \ldots, n. \tag{34}$$

Proof For each $i = 1, \dots, m'$, let λ_i be the Lagrange multiplier associated with the i^{th} constraint in (COP). Let μ be the multiplier associated with the equality constraint. The Lagrangian is

$$\begin{aligned} L(\boldsymbol{\theta}; \boldsymbol{\lambda}, \mu) &= \sum_{i=1}^{n} a_i \theta_i g_i \left(\frac{\theta_i}{\theta_0}\right) - \sum_{k=1}^{m'} \lambda_k \left(\sum_{i=1}^{n} A_{ki}\theta_i - u_k\right) - \mu \left(\sum_{i=0}^{n} \theta_i - 1\right) \\ &= \sum_{i=1}^{n} \theta_i \left(a_i g_i \left(\frac{\theta_i}{\theta_0}\right) - \sum_{k=1}^{m'} \lambda_k A_{ki} - \mu\right) - \mu\theta_0 + \mu + \sum_{k=1}^{m'} \lambda_k u_k. \end{aligned}$$

Taking the supremum successively over the different variables, we obtain the dual function

$$\begin{aligned} L^*(\boldsymbol{\lambda}, \mu) &\triangleq \sup_{\boldsymbol{\theta}>0} L(\boldsymbol{\theta}; \boldsymbol{\lambda}, \mu) = \sup_{\theta_0>0} \left\{ \sup_{\theta_1 \dots \theta_n > 0} L(\boldsymbol{\theta}; \boldsymbol{\lambda}, \mu) \right\} \\ &= \mu + \sum_{k=1}^{m'} \lambda_k u_k + \sup_{\theta_0>0} \left\{ -\mu\theta_0 + \sup_{\theta_1 \dots \theta_n > 0} \sum_{i=1}^{n} \theta_i \left(a_i g_i \left(\frac{\theta_i}{\theta_0}\right) - \sum_{k=1}^{m'} \lambda_k A_{ki} - \mu \right) \right\} \\ &= \mu + \sum_{k=1}^{m'} \lambda_k u_k + \sup_{\theta_0>0} \left(-\mu\theta_0 + \sup_{\theta_1 \dots \theta_n > 0} \sum_{i=1}^{n} \theta_0 \phi_i (\frac{\theta_i}{\theta_0}, \boldsymbol{\lambda}, \mu) \right) \\ &= \mu + \sum_{k=1}^{m'} \lambda_k u_k + \sup_{\theta_0>0} \theta_0 \left(-\mu + \sum_{i=1}^{n} \sup_{\theta_i>0} \phi_i (\frac{\theta_i}{\theta_0}, \boldsymbol{\lambda}, \mu) \right), \end{aligned} \tag{35}$$

where $\phi_i(y, \boldsymbol{\lambda}, \mu)$ is defined as in (28). The value of θ_0 has no impact on the value of the inner supremums in (35) since the optimization is over the ratio $\frac{\theta_i}{\theta_0}$ with the numerator free to take any positive value. Thus we may write the dual problem as

$$\inf_{\boldsymbol{\lambda}\geq 0, \mu} L^*(\boldsymbol{\lambda}, \mu) = \inf_{\boldsymbol{\lambda}\geq 0, \mu} \left\{ \mu + \sum_{k=1}^{m'} \lambda_k u_k + \sup_{\theta_0>0} \theta_0 \left(-\mu + \sum_{i=1}^{n} \sup_{y_i>0} \phi_i(y_i, \boldsymbol{\lambda}, \mu) \right) \right\}.$$

At optimality, the quantity in the inner parentheses must be non-positive, so we may write

$$\begin{aligned} \inf \quad & \mu + \sum_{k=1}^{m'} \lambda_k u_k \\ \text{s.t.} \quad & \mu \geq \sum_{i=1}^{n} \sup_{y_i>0} \phi_i(y_i, \boldsymbol{\lambda}, \mu) \\ & \boldsymbol{\lambda} \geq 0. \end{aligned}$$

The inequality constraint is tight at optimality, because $\phi_i(y_i, \boldsymbol{\lambda}, \mu)$ are strictly decreasing in μ.

We now show that $\phi_i(y_i, \boldsymbol{\lambda}, \mu)$ achieves a maximum at some $y_i = y_i^* > 0$, for any fixed $\boldsymbol{\lambda}$ and μ. For ease of notation, we fix i and drop the subscript. Let

$$\phi(y) \triangleq \phi_i(y, \boldsymbol{\lambda}, \mu) = y\left(g(y) - \nu\right), \tag{36}$$

where we define $g(y) \triangleq a_i g_i(y)$ and $\nu \triangleq \sum_{k=1}^{m'} \lambda_k A_{ki} + \mu$. By Assumption 1, there exists a value $\hat{y} > 0$ for which $g(\hat{y}) = \nu$, since the attraction $f_i(\cdot)$ is defined everywhere on $\mathbb{R}$ and $g(\cdot)$ is its inverse. Moreover, $\phi(\cdot)$ is strictly positive on the interval $(0, \hat{y})$ and strictly negative on $(\hat{y}, \infty)$ since $g(\cdot)$ is strictly decreasing. Also by Assumption 1 (ii) ,

$$\lim_{y\downarrow 0} \phi(y) \triangleq \lim_{y\downarrow 0} (yg(y) - y\nu) = \lim_{y\downarrow 0} a_i y g_i(y) = \lim_{x\to\infty} a_i x f_i(x) = 0.$$

We consider the continuous extension of ϕ, with $\phi(0) = 0$, without loss of generality. Then, the continuous function $\phi(\cdot)$ achieves a maximum y_i^* on the closed interval $[0, \hat{y}]$ by Weierstrass' Theorem. We have $0 < y_i^* < \hat{y}$ and $\phi(y_i^*) > 0$, since $\phi(0) = \phi(\hat{y}) = 0$ and ϕ is strictly positive on the interval.

Suppose now that condition (12) (equivalently, condition (13)) holds. Then (COP) has a concave objective, a bounded polyhedral feasible set and a finite maximum (because the feasible set is bounded). Then the dual (DCOP) has an optimal solution and there is no duality gap. Consider now an optimal dual solution $(\boldsymbol{\lambda}^*, \mu^*)$ and corresponding maximizers $y_1^*, \ldots, y_n^*$. Then (34) is a primal optimal solution, since it maximizes the Lagrangian $L(\boldsymbol{\theta}; \boldsymbol{\lambda}, \mu)$ by definition of the dual: we have only made the change of variable $y_i = \frac{\theta_i}{\theta_0}$. □

D.1 The Dual Problem under MNL Demand Models

Proposition 4 *The dual problem* (DCOP) *for the special case of MNL attraction functions* (2) *is given by* (DMNL).

Proof The inverse attraction functions for the MNL model (2) and their derivatives are $g_i(y) = -\log \frac{y}{v_i}$, and $g_i'(y) = \frac{-1}{y}$, respectively. Then the first order necessary optimality condition for the i^{th} inner maximization in (DCOP) is

$$\frac{\partial \phi_i}{\partial y} = \left(a_i g_i(y) - \sum_{k=1}^{m'} \lambda_k A_{ki} - \mu \right) + a_i y g_i'(y) = 0 \quad \Leftrightarrow$$

$$y = v_i \exp\left\{ -1 - \frac{\sum_{k=1}^{m'} \lambda_k A_{ki} + \mu}{a_i} \right\}.$$

The preceding line gives the unique maximizer since one exists by Proposition 3. Substituting the optimal value of y back into (28) yields that

$$\phi_i(y_i^*, \boldsymbol{\lambda}, \mu) = a_i v_i \exp\left\{ -1 - \frac{\sum_{k=1}^{m'} \lambda_k A_{ki} + \mu}{a_i} \right\},$$

which can in turn be substituted into (DCOP) to obtain (DMNL). The constraint may be relaxed to an inequality which is tight at optimality, since the right hand side is decreasing in μ. □

D.2 Solving the Dual Problem in General

More generally, there may not exist a closed form solution for the values $\phi_i(y_i^*, \boldsymbol{\lambda}, \mu)$. Then the dual problem may not reduce to a tractable optimization problem. If there is no closed form inverse for the attraction functions, not even the primal market share problem (COP) can be solved directly, even if it has a concave objective function. This is notably the case for the demand models discussed in Section 4 (although we have shown that the primal objective function's gradient and Hessian can nevertheless be computed efficiently).

In this section, we present a column generation algorithm to solve the dual which avoids both of these difficulties. It is more general than solving either of the formulations (COP) and (DCOP) directly, since it does not require the convexity of the primal objective function assumed in Theorem 1, and it does not require a closed form solution for the inner maximizations of the dual problem.

In the dual (DCOP), fixing the variables $\boldsymbol{\lambda}$ uniquely determines the value of the remaining variable μ, because of the equality constraint. Notice that the right hand side of the constraint is decreasing in μ, because all functions $\phi_i(y, \boldsymbol{\lambda}, \mu)$ are decreasing in μ for any value of y. Furthermore, any feasible μ is positive since the maxima $\phi_i(y_i^*, \boldsymbol{\lambda}, \mu)$ are positive by Proposition 3. We define $\mu(\boldsymbol{\lambda})$ as the unique root of equation

$$F_{\boldsymbol{\lambda}}(\mu) = \mu - \sum_{i=1}^{n} \max_{y_i > 0} \phi_i(y_i, \boldsymbol{\lambda}, \mu) = 0. \tag{37}$$

Its value may be computed by a line search which computes the maximizers y_i^* at each evaluation. When these one-dimensional maximizations are tractable, it is possible to evaluate the dual objective efficiently, and the Dantzig-Wolfe column generation scheme can be applied to solve (COP). (See, for instance, (Bertsekas 1999) for details.) Specifically, we propose the following algorithm:

1. **Initialization:** Set lower and upper bounds $LB = -\infty$ and $UB = \infty$.
2. **Master Problem:** Given market share vectors $\boldsymbol{\theta}^0, \boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^{L-1}$, solve the following linear program over the variables $\xi_0, \xi_1, \ldots, \xi_{L-1}$:

$$\begin{aligned} \gamma^L = \max \quad & \sum_{\ell=0}^{L-1} \xi^\ell \Pi(\boldsymbol{\theta}^\ell) \\ \text{s.t.} \quad & \sum_{i=1}^{n} A_{ki} \left(\sum_{\ell=0}^{L-1} \xi^\ell \theta_i^\ell \right) \le u_k \qquad k = 1 \ldots m' \\ & \sum_{\ell=0}^{L-1} \xi^\ell = 1, \qquad \xi^\ell \ge 0, \quad \ell = 0, \ldots, L-1. \end{aligned} \tag{LP}$$

   Let $\boldsymbol{\lambda}^L$ be the vector of optimal dual variables associated with the inequality constraints. The master problem solves (COP) with the feasible region restricted to the convex hull of the demand vectors $\boldsymbol{\theta}^0, \boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^{L-1}$. If the optimal value γ^L of (LP) exceeds the lower bound LB, update $LB := \gamma^L$.
3. **Dual Function Evaluation:** Compute the root $\mu(\boldsymbol{\lambda}^L)$ of the dual equality constraint $F_{\boldsymbol{\lambda}^L}(\mu)$ shown in (37), and let $\boldsymbol{\theta}^L$ be the primal solution (34) corresponding to the maximizers $\{y_i^*, i = 1, \ldots, n\}$. If the dual objective value $L(\boldsymbol{\theta}^L; \boldsymbol{\lambda}^L, \mu(\boldsymbol{\lambda}^L))$ is less than the upper bound, set $UB := L(\boldsymbol{\theta}^L; \boldsymbol{\lambda}^L, \mu(\boldsymbol{\lambda}^L))$.
4. **Termination:** If $(UB - LB)$ is below a pre-specified tolerance, stop. Otherwise, let $L := L + 1$ and go to Step 2.

This algorithm requires at least one initial feasible solution $\boldsymbol{\theta}^0$, which can be found by solving any linear program with the constraints of (COP). It does not require (COP) to have a concave objective, since it computes an optimal solution to its dual, which is always a convex minimization problem. Moreover, it can be used even if there is no closed form for the inverse attraction functions $g_i(\cdot)$. Indeed, we can equivalently represent the functions $\phi_i(y, \boldsymbol{\lambda}, \mu)$ in terms of the original attraction functions $f_i(x_i)$, as

$$\psi_i(x_i, \boldsymbol{\lambda}, \mu) \triangleq f_i(x_i) \left(a_i x_i - \sum_{k=1}^{m'} \lambda_k A_{ki} - \mu \right). \tag{38}$$

Then the maximization can be performed over the price x_i, and the optimal price for given dual variables $(\boldsymbol{\lambda}, \mu)$ is

$$x_i^* \triangleq \arg\max_{x_i} \psi_i(x_i, \boldsymbol{\lambda}, \mu) = g_i(y_i^*).$$

The maximum is guaranteed to exist since y_i^* exists by Proposition 3. It can be computed via a line search if it is the unique local maximum. The unimodality of ϕ_i (and equivalently, of ψ_i) is guaranteed, for instance, by the assumption of Theorem 1, or more generally, by the assumption of Proposition 5 below. In the column generation algorithm, the objective of (LP) depends on the prices $x_i = g_i\left(\theta_i^0/\theta_0^0\right)$ corresponding to the initial feasible point. Because they must satisfy $f_i(x_i) = \theta_i^0/\theta_0^0$ and $f_i(x_i)$ is monotone, they can also be found using line search procedures in practice. For each new point $\boldsymbol{\theta}^L$, corresponding prices x_i^* are computed in the maximizations of ψ_i over x_i.

Finally, we remark that it is not necessary to dualize the price bounds represented by the constraints $k = (m+1), \ldots, m'$ defined in (9). These constraints may be omitted if the price bounds $\underline{x}_i \leq x_i \leq \overline{x}_i$ are instead enforced when computing the maximizers x_i^* (or, equivalently, the bounds $f_i(\underline{x}_i) \geq y_i \geq f_i(\overline{x}_i)$ are enforced when computing y_i^*). This modification reduces the number of constraints from $m' = (m+2n)$ to m.

The algorithm just described may also be viewed as a generalization of the procedure presented by Gallego and Stefanescu (2009) to general attraction demand models and arbitrary linear inequality constraints. (Although they arrive at their method by taking the dual of the price-based formulation (P) for the special case of MNL demand.) Because convergence of column generation algorithms is often slow near the optimum, we expect that directly solving (COP) or (DCOP) will be more efficient when it is possible. This, for example, is the case with the MNL demand models considered by Gallego and Stefanescu (2009). However, the column generation algorithm applies to demand models where it is not possible to solve the other formulations. It can provide an upper bound on the optimal profit when the objective function of (COP) is not concave, and can often compute an approximate solution quickly (accurate within a few percent in relatively few iterations, as shown in our experiments).

We end this section with the following proposition providing a sufficient condition on the inverse attractions guaranteeing unique maximizers y_i^*. It requires that the inverse attraction functions are "sufficiently concave" (though not necessarily concave) up until some $\bar{y}$, and then "sufficiently convex" afterward. Omitting the ratio $\frac{x}{y}$, conditions (39) and (40) below correspond to strict concavity and strict convexity, respectively. However, the first requirement is weaker, and the second is stronger, because this ratio is less than one. (Recall that $g_i'(x) < 0, \forall x$ since f_i and g_i are decreasing.) We note that the proposition allows $\bar{y} = 0$ or $\bar{y} = \infty$, in which case one of the assumptions holds trivially.

Proposition 5 *If for each $i = 1, 2, \ldots, n$, there exists a point $\bar{y}_i \in [0, \infty]$ such that*

$$g_i(y) < g_i(x) + \frac{x}{y}(y-x)g_i'(x), \qquad \forall x, y \in (0, \bar{y}_i], x < y, \qquad \textit{and} \tag{39}$$

$$g_i(y) > g_i(x) + \frac{x}{y}(y-x)g_i'(x), \qquad \forall x, y \in [\bar{y}_i, \infty), x < y, \tag{40}$$

then the maximizers $\left\{y_i^*, i = 1, \ldots, n\right\}$ are unique for any values of $\boldsymbol{\lambda}$ and μ.

Proof We fix i and use the simplified notation defined in (36). From Proposition 3, the maximizer $y_i^* > 0$ exists, and it must be a stationary point of ϕ. We will show that the rightmost stationary point to the left of $\bar{y}_i$ maximizes $\phi(y)$ over $(0, \bar{y}_i]$, and that the leftmost stationary point to the right of $\bar{y}_i$ maximizes $\phi(y)$ over $[\bar{y}_i, \infty)$, if they exist. At least one of them must exist since we know a maximum is attained. If both exist, we deduce that there is an additional stationary point between them by applying the mean value theorem. This contradicts the fact that they are the rightmost and leftmost stationary points on their respective intervals, proving uniqueness of the maximizer.

Suppose $y \in (0, \bar{y}_i]$ is a stationary point of $\phi(\cdot)$, i.e.

$$\phi'(y) = g(y) - \nu + yg'(y) = 0 \qquad \Rightarrow \qquad \nu = g(y) + yg'(y). \tag{41}$$

We will show that for any other point $x \in (0, y)$, whether or not it is a stationary point,

$$\begin{aligned}
&\phi(x) < \phi(y) \quad \Leftrightarrow \quad x(g(x) - \nu) < y(g(y) - \nu) \quad \Leftrightarrow \\
&x\left(g(x) - g(y) - yg'(y)\right) < -y^2 g'(y) \quad \Leftrightarrow \quad x(g(x) - g(y)) < (x-y)\, yg'(y) \quad \Leftrightarrow \\
&\qquad g(x) - g(y) < (x-y)\frac{y}{x}g'(y) \quad \Leftrightarrow \quad \frac{x}{y}\left(\frac{g(x) - g(y)}{x-y}\right) > g'(y),
\end{aligned}$$

where we used (41). Having fixed y, we denote the left hand side as a function of x by

$$h(x) = \frac{x}{y}\left(\frac{g(y) - g(x)}{y-x}\right)$$

and note that $\lim_{x\uparrow y} h(x) = g'(y)$. Thus, to prove the inequality, it is sufficient to show that the continuous function $h(x)$ is decreasing in x on the interval $(0, y)$. We consider the derivative with respect to x

$$\begin{aligned} h'(x) &= \frac{g(x)-g(y)}{y(x-y)} + \frac{x}{y}\left(\frac{g'(x)}{x-y} - \frac{g(x)-g(y)}{(x-y)^2}\right) \\ &= \frac{1}{(x-y)^2}\left(\frac{x-y}{y}(g(x)-g(y)) + \frac{x}{y}(x-y)g'(x) - \frac{x}{y}(g(x)-g(y))\right) \\ &= \frac{1}{(x-y)^2}\left(\frac{x}{y}(x-y)g'(x) - g(x) + g(y)\right) \\ &= a_i \frac{1}{(x-y)^2}\left(-\frac{x}{y}(y-x)g_i'(x) - g_i(x) + g_i(y)\right) < 0. \end{aligned} \tag{42}$$

The assumption (39) implies that the above derivative is negative, where we have substituted $g_i(\cdot)$ back in, and thus $h(x)$ is decreasing.

A similar argument shows the analogous result for stationary points to the right of $\bar{y}_i$. Take instead $x \in (\bar{y}_i, \infty)$ to be the leftmost stationary point in the half-line, and let $y \in (x, \infty)$ be some other stationary point. We still have that $x < y$, but now $\phi(x) > \phi(y) \Leftrightarrow h(x) < g'(y)$, because $h(x)$ is *increasing* in x. This is implied by the assumption (40), which shows that the derivative in (42) is now positive. □

D.3 Performance of the Column Generation Algorithm

Table 3 shows the accuracy achieved and the running time in seconds after a fixed number of iterations of the column generation algorithm, when applied to randomly generated problem instances with four overlapping customer segments, using the approximation of Section 4. Only the most recently active 512 columns are retained in the master problem (LP). We have no closed form for the inner maximizers y_i^* and instead use a numerical minimization algorithm based on Brent's method to compute them. Brent's method (see Brent (1973)) is also used to solve (37) numerically. The algorithm was halted if six significant digits of accuracy were achieved.

Products	Constraints	100 Iteration		500 Iterations		2000 Iterations	
(n)	(m)	Duality Gap	Time	Duality Gap	Time	Duality Gap	Time
16	256	0.03%	0.41	(< 1e-6)	0.86	(< 1e-6)	0.86
64	256	19.63%	1.10	0.68%	11.83	(< 1e-6)	90.05
256	256	31.28%	2.49	4.18%	25.99	0.19%	198.75
512	256	25.77%	4.05	4.75%	38.30	0.92%	263.76
1,024	256	14.39%	8.74	4.12%	65.91	2.71%	393.85
2,048	256	4.45%	18.31	1.25%	117.06	1.15%	605.92
4,096	256	1.35%	35.50	0.52%	212.99	0.52%	1,150.80
256	16	0.03%	1.49	0.01%	9.54	0.01%	55.87
256	64	6.72%	1.36	0.10%	12.74	0.07%	68.91
256	256	31.28%	2.49	4.18%	25.99	0.19%	198.75
256	512	237.23%	4.02	12.77%	75.98	1.79%	638.32
256	1,024	215.69%	9.98	24.42%	250.18	5.17%	1,882.00
256	2,048	161.09%	20.73	26.25%	583.72	6.80%	4,636.20
256	4,096	178.72%	64.33	28.96%	1,561.70	7.76%	10,944.00

Table 3 Duality gap as a percentage of LB and running time in seconds for the column generation algorithm.

As is often the case for column generation algorithms, we observe fast convergence early on. After 500 iterations, most of the instances are solved to within 10 percent of the optimal objective value. Quadrupling the number of iterations further reduces the duality gap to a few percentage points in all but the largest instances. The solution times compare favorably with the price formulation (P) and the dual formulation (DCOP) for the single-segment case, but are significantly slower than for the market-share formulation (COP). Of course, the latter formulation requires the custom objective evaluation code described in the preceding section when multiple segments are being approximated. We conclude that the column generation method offers a viable alternative when the other formulations cannot be applied easily, and only limited accuracy is needed.