# A New Compact Linear Programming Formulation for Choice Network Revenue Management

Summit Kunnumkal
Kalyan Talluri

December 2012

*Barcelona GSE Working Paper Series*

*Working Paper nº 677*

# A new compact linear programming formulation for choice network revenue management

Sumit Kunnumkal* Kalyan Talluri†

December 13, 2012

**Abstract**

The choice network revenue management model incorporates customer purchase behavior as a function of the offered products, and is the appropriate model for airline and hotel network revenue management, dynamic sales of bundles, and dynamic assortment optimization. The optimization problem is a stochastic dynamic program and is intractable. A certainty-equivalence relaxation of the dynamic program, called the choice deterministic linear program ($CDLP$) is usually used to generate dyamic controls. Recently, a compact linear programming formulation of this linear program was given for the multi-segment multinomial-logit (MNL) model of customer choice with non-overlapping consideration sets. Our objective is to obtain a tighter bound than this formulation while retaining the appealing properties of a compact linear programming representation. To this end, it is natural to consider the affine relaxation of the dynamic program. We first show that the affine relaxation is NP-complete even for a single-segment MNL model. Nevertheless, by analyzing the affine relaxation we derive a new compact linear program that approximates the dynamic programming value function better than $CDLP$, provably between the $CDLP$ value and the affine relaxation, and often coming close to the latter in our numerical experiments. When the segment consideration sets overlap, we show that some strong equalities called product cuts developed for the $CDLP$ remain valid for our new formulation. Finally we perform extensive numerical comparisons on the various bounds to evaluate their performance.

## 1 Introduction and literature review

Revenue management is the control of the sale of a limited quantity of a resource (hotel rooms for a night, airline seats, advertising slots etc.) to a heterogenous population with different valuations for a unit of the resource. The resource is perishable, and for simplicity sake, we assume that it perishes at a fixed point of time in the future. Sale is online, and the firm has to decide what products to offer (at a given price for each product), the tradeoff being selling too much at too low a price early and running out of capacity, or, rejecting too many low-valuation customers and ending up with excess unsold inventory.

*Indian School of Business, Hyderabad, 500032, India, email: sumit_kunnumkal@isb.edu

†ICREA and Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain, email: kalyan.talluri@upf.edu

In industries such as hotels, advertising and airlines, the products consume bundles of different resources (multi-night stays, multi-leg itineraries) and the decision to accept or reject a particular product at a certain price depends on the future demands and revenues for all the resources used by the product and indirectly, on all the resources in the network. Network revenue management (network RM) is control based on the demands for the entire network. Chapter 3 of Talluri and van Ryzin [15] contains all the necessary background on network RM.

RM incorporating more realistic models of customer behavior as customers choosing from set of offered products have recently become popular, initiated in Talluri and van Ryzin [14] for the single-resource problem. Bodea, Ferguson, and Garrow [2] for instance use choice data from a large hotel chain and empirically study the suitability of choice models.

The choice network RM problem can be formulated as a dynamic program with exponentially large state and action spaces. Since the dynamic programming formulation is computationally intractable, many approximation methods have been proposed starting with Gallego, Iyengar, Phillips, and Dubey [4] and Liu and van Ryzin [7], who formulate the choice deterministic linear program ($CDLP$). They show $CDLP$ gives an upper bound on the value function. Since $CDLP$ has an exponential number of decision variables it has to be solved using column generation. The column generation procedure turns out to be tractable for the MNL model of choice when the consideration sets of the different customer segments are disjoint ([7]). However, generating the columns is difficult (NP-complete) when the segment consideration sets overlap under the MNL model with just two segments ([3], [11]).

Given the difficulty of solving $CDLP$, Talluri [13] explores a weaker segment-based deterministic concave program ($SDCP$) formulation. The $SDCP$ formulation is further strengthened by adding equalities called product-cuts in Meissner, Strauss, and Talluri [9]. Strauss and Talluri [12] show that $SDCP$ with the product-cuts added is equivalent to $CDLP$ when the consideration set intersections have a tree structure.

Kunnumkal and Topaloglu [6] and Zhang and Adelman [17] study decomposition procedures and an affine relaxation of the dynamic program. In the same vein, Meissner and Strauss [8] look at time-sensitive bid-price controls based on a decomposition procedure. All these methods yield upper bounds on the value function that are provably tighter than the $CDLP$ upper bound. However they are not easy to solve, even for a single-segment MNL model of choice.

Recently, Gallego, Ratliff, and Shebalov [5] give a new compact formulation of $CDLP$ called the sales-based linear program ($SBLP$) for the case of MNL with non-overlapping segment consideration sets. This formulation is very appealing as it is compact—not requiring column or constraint generation—and hence scalable to industrial-size problems.

Can we obtain a tighter bound than ($SBLP$) while maintaining a compact formulation? To this end, it is natural to consider the affine relaxation of the dynamic program. Unfortunately, we show that the affine relaxation is NP-complete even for a single segment MNL model, possibly marking the limit of tractability of dynamic programming approximations. Nevertheless, by analyzing the affine relaxation we derive a new linear-programming formulation that yields an upper bound on the dynamic programming value function and is provably tighter than the $CDLP$ bound. Although theoretically weaker than the affine relaxation, we find in our numerical study, that our relaxation is often close to the affine relaxation upper bound. Moreover, for the MNL model, our formulation is compact and similar to the one discovered in [5]. Next, when the segment consideration sets overlap, we show that the strong constraints called product cuts developed for the $CDLP$ in Meissner et al. [9] remain valid for our new formulation.

The remainder of the paper is organized as follows: In §2 we describe the network choice RM model, the notation, and the basic dynamic program. In §3 we state the $CDLP$ and the affine relaxation of the dynamic program and show that the affine relaxation is NP-complete to solve even for a single-segment MNL model. Recently Vossen and Zhang [16] give an equivalent, reduced formulation of the affine relaxation using Dantzig-Wolfe decomposition ideas. We first give a simpler, alternative proof of the same result. We then propose compact linear programming formulations in §4 that fall in between $CDLP$ and the affine relaxation for the MNL model with disjoint consideration sets. §6 discusses a tightening of the formulation when segment consideration sets overlap. §7 contains our computational study using the new formulations.

# 2 Model and notation

A product is a specification of a price and the set of resources that it consumes. For example, a product could be an itinerary-fare class combination for an airline network, where an itinerary is a combination of flight legs; in a hotel network, a product would be a multi-night stay for a particular room type at a certain price point. Time is discrete and assumed to consist of $\tau$ intervals, indexed by $t$. The booking horizon begins at time $t = 1$ and ends at $t = \tau$; all the resources perish instantaneously at time $\tau + 1$. We make the standard assumption that the time intervals are fine enough so that the probability of more than one customer arriving in any single time period is negligible. The underlying network has $m$ resources (indexed by $i$) and $n$ products (indexed by $j$), and we refer to the set of all resources as $I$ and the set of all products as $J$. A product $j$ uses a subset of resources, and is identified (possibly) with a set of sale restrictions or features and a revenue of $f_j$. A resource $i$ is said to be in product $j$ ($i \in j$) if $j$ uses resource $i$.

*Throughout, we index resources by $i$, products by $j$, and time periods by $t$.* We simplify the notation significantly by making this consistent; for instance if $j$ uses resource $i$, we represent it as $i \in j$, and all $j$ that use resource $i$ by $\{j \,|\, j \ni i\}$. We use this notation instead of the somewhat more cumbersome, albeit a bit more precise, option of defining $\mathcal{I}_j \subseteq \mathcal{I}$ as the set of resources used by product $j$ and $\mathcal{J}_i \subseteq \mathcal{J}$ as the set of products that use resource $i$ and then writing $i \in \mathcal{I}_j$ and $j \in \mathcal{J}_i$.

The resources used by $j$ are represented by $a_{ij} = 1$ if $i \in j$, and $a_{ij} = 0$ if $i \notin j$, or alternately with the 0-1 incidence vector $\boldsymbol{A}^j$ of product $j$. Let $A$ denote the resource-product incidence matrix; columns of $A$ are then $\boldsymbol{A}^j$.

We use superscripts on vectors to index the vectors (for example, the resource capacity vector associated with time period $t$ would be $\boldsymbol{r}^t$) and subscripts to indicate components (for example, the capacity on resource $i$ in time period $t$ would be $r_i^t$). We use $\mathbb{1}_{[\cdot]}$ as the indicator function, 1 if true and 0 if false.

We let $\boldsymbol{r}^1 = [r_i^1]$ represent the initial capacity on the resources and $\boldsymbol{r}^t = [r_i^t]$ denote the remaining capacity on resource $i$ at beginning of time period $t$. The remaining capacity $r_i^t$ takes values in the set $\mathcal{R}_i = \{0, \ldots, r_i^1\}$ and $\mathcal{R} = \prod_i \mathcal{R}_i$ represents the state space.

## 2.1 Demand model

We assume there are $\mathcal{L} = \{1, \ldots, L\}$ customer segments, each with distinct purchase behavior. In each period a customer arrives with a rate $\lambda$ and that customer belongs to segment $l$ with probability

$\lambda_l$. So the total arrival rate $\lambda = \sum_{l=1}^{L} \lambda_l$.

Each segment $l$ has a *consideration set* $\mathcal{C}_l \subseteq J$ of products that it considers for purchase. We assume this consideration set is known to the firm (by a previous process of estimation and analysis). A segment-$l$ customer is indifferent to a product outside his consideration set; i.e., his choice probabilities are not affected by products offered not in the consideration set.

In each period the firm offers a subset $S$ of its products for sale, called the *offer set*. Given an offer set $S$, an arriving customer purchases a product $j$ in the set $S$ or decides not to purchase. The no-purchase option is indexed by 0 and is always present for the customer.

A segment-$l$ customer purchases $j \in S$ with given probability $P_j^l(S)$. This is a set-function defined on all subsets of $J$. For the moment we assume these set functions are given by an oracle; it could conceivably be given by a simple formula such as the Multinomial Logit (MNL) model. If $S_l = \mathcal{C}_l \cap S$ note that $P_j^l(S) = P_j^l(S_l)$. We define the vector $\boldsymbol{P}^l(S) = [P_1^l(S_l), \ldots, P_n^l(S_l)]$ (recall the no-purchase option is indexed by 0, so it is not included in this vector).

Given a customer arrival, and an offer set $S$, the probability that the firm sells $j \in S$ is then given by $P_j(S) = \sum_l \lambda_l P_j^l(S_l)$ and makes no sale with probability $P_0(S) = 1 - \sum_{j \in S} P_j(S)$. We define the vector $\boldsymbol{P}(S) = [P_1(S), \ldots, P_n(S)]$. Notice that $\boldsymbol{P}(S) = \sum_l \lambda_l \boldsymbol{P}^l(S)$. We define the vectors $\boldsymbol{Q}^l(S) = A\boldsymbol{P}^l(S)$ and $\boldsymbol{Q}(S) = A\boldsymbol{P}(S)$. $\boldsymbol{Q}^l(S)$ represents the expected capacity consumed on the resources when set $S$ is offered conditional on a segment-$l$ customer arrival, while $\boldsymbol{Q}(S)$ represents the expected capacity consumed over all arrivals when set $S$ is offered. The revenue functions can be written as $R^l(S) = \sum_{j \in S_l} f_j P_j^l(S_l)$ and $R(S) = \sum_{j \in S} f_j P_j(S)$.

For brevity of notation, we write $i \in S$ or $S \ni i$ whenever there is a $j \in S$ with $i \in j$.

## 2.2 Multiple segment MNL model with non-overlapping consideration sets

Some of the results apply to general discrete-choice models, but the main formulation applies to the MNL model of choice with multiple segments and non-overlapping consideration sets.[1]

In this model, the $L$ segments have considerations sets that do not overlap ($\mathcal{C}_l \cap \mathcal{C}_{l'} = \emptyset, l \neq l'$). A customer in segment $l$ (the firm does not observe segment membership), when a subset $S_l \subseteq \mathcal{C}_l$ of products are offered by the firm, chooses product $j \in S_l$ with probability

$$P_j^l(S_l) = \frac{w_j^l}{1 + \sum_{k \in S_l} w_k^l},$$

where $w_j^l$ is a *weight* associated with product $j$. This weight represents the exponential of a utility the customer derives from $j$ as a function of some attributes of the product (such as price etc.). As we fix all the attributes and our decision is on subsets to offer, we do not delve too much into how the weights are formed. We refer the reader to Ben-Akiva and Lerman [1] for background on this popular model.

In our choice model, the no-purchase option is indexed 0, and we normalize the weights so that the no-purchase weight is 1.0. So if $S_l$ is offered, a customer does not purchase any of the offered products and leaves the system with a probability $P_0^l(S_l) = \frac{1}{1+\sum_{k \in S_l} w_k^l}$.

[1] The formulation in fact extends unchanged to the slightly more general attraction model of Gallego et al. [5].

## 2.3 Choice dynamic program

The dynamic program (DP) to determine optimal controls can be written down as follows. Let $V_t(\boldsymbol{r}^t)$ denote the maximum expected revenue to go, given remaining capacity $\boldsymbol{r}^t$ at the beginning of period $t$. Then $V_t(\boldsymbol{r}^t)$ must satisfy the Bellman equation

$$V_t(\boldsymbol{r}^t) = \max_{S \subseteq \mathcal{S}(\boldsymbol{r}^t)} \left\{ \sum_{j \in S} \lambda P_j(S) \left[ f_j + V_{t+1}(\boldsymbol{r}^t - \boldsymbol{A}^j) \right] + \left[ \lambda P_0(S) + 1 - \lambda \right] V_{t+1}(\boldsymbol{r}^t) \right\}, \tag{1}$$

where

$$\mathcal{S}(\boldsymbol{r}) = \{ j | a_{ij} \le r_i \forall i \}$$

represents the set of products that can be offered given the capacity vector $\boldsymbol{r}$. The boundary conditions are $V_{\tau+1}(\boldsymbol{r}) = V_t(\boldsymbol{0}) = 0$ for all $\boldsymbol{r}$ and for all $t$, where $\boldsymbol{0}$ is a vector of all zeroes. $V^{DP} = V_1(\boldsymbol{r}^1)$ denotes the optimal expected total revenue over the booking horizon, given the initial capacity vector $\boldsymbol{r}^1$.

## 2.4 Linear programming formulation of the dynamic program

The value functions can, alternatively, be obtained by solving a linear program; Zhang and Adelman [17]. The linear programming formulation of the network choice RM $DP$ given below, has a decision variable for each state vector in each period $V_t(\boldsymbol{r})$ and is as follows:

$$\begin{aligned} V^{DP} = \quad & \min_V V_1(\boldsymbol{r}^1) \\ \text{s.t.} \quad & \\ (DP) \quad & V_t(\boldsymbol{r}) \ge \sum_j \lambda P_j(S)[f_j + V_{t+1}(\boldsymbol{r} - \boldsymbol{A}^j) - V_{t+1}(\boldsymbol{r})] + V_{t+1}(\boldsymbol{r}) \\ & \qquad \forall \boldsymbol{r} \in \mathcal{R}, S \subset \mathcal{S}(\boldsymbol{r}), t, \end{aligned}$$

with the boundary condition that $V_{\tau+1}(\cdot) = 0$. Both the dynamic program (1) and linear program $DP$ are computationally intractable, but the linear program $DP$ turns out to be useful in developing value function approximation methods. In the next section, we describe methods to approximate the value function.

# 3 Approximations and upper bounds

In the following, we outline the two approximations studied in this paper. We first describe the choice deterministic linear program and then outline the affine relaxation method.

## 3.1 Choice deterministic linear program ($CDLP$) and its compact formulation for MNL

The choice deterministic linear program ($CDLP$) proposed in Gallego et al. [4] and Liu and van Ryzin [7] is a certainty-equivalence approximation to (1) and is given by the following linear program:

$$\begin{aligned} V^{CDLP} = \quad & \max_h \quad \sum_{S \subseteq J} \lambda R(S) h_S \\ \text{s.t.} \quad & \sum_{S \subseteq J} \lambda h_S \boldsymbol{Q}(S) \leq \boldsymbol{r}^1 \\ (CDLP) \quad & \sum_{S \subseteq J} h_S = \tau \\ & 0 \leq h_S, \quad \forall\, S \subseteq J. \end{aligned}$$

The decision variables $h_S$ can be interpreted as the number of time periods each set is offered (including the empty set). Liu and van Ryzin [7] show that the optimal objective function value of $CDLP$, $V^{CDLP}$ is an upper bound on $V^{DP}$. Since $CDLP$ has $2^n$ decision variables, it has to be solved using column generation. Liu and van Ryzin [7] show that the column generation procedure can be efficiently carried out when choice is according to the MNL model and the consideration sets of the different segments do not overlap. Bront et al. [3] and Rusmevichientong et al. [11] investigate this further and show that column generation is NP-complete whenever the consideration sets for the segments overlap, for the MNL choice model with just two segments.

### Compact formulation

Gallego et al. [5] give the following equivalent formulation of $CDLP$ for MNL with non-overlapping consideration sets:

$$V^{SBLP} = \quad \max_x \quad \sum_{l=1}^{L} \sum_{j \in \mathcal{C}_l} f_j x_j^l \tag{2}$$

$$\text{s.t.} \quad \sum_{l=1}^{L} \sum_{j \in \mathcal{C}_l} \boldsymbol{A}^j x_j^l \leq \boldsymbol{r}^1 \tag{3}$$

$$\begin{aligned} (SBLP) \quad & x_0^l + \sum_{j \in \mathcal{C}_l} x_j^l = \lambda_l \tau \quad \forall l \\ & \frac{x_j^l}{w_j^l} - x_0^l \leq 0 \quad \forall l, \forall j \in \mathcal{C}_l \\ & x_j^l \geq 0. \end{aligned}$$

In the above linear program, the decision variables $x_j^l$ can be viewed as the rate of sales of product $j$ to segment $l$. Note also that we assumed homogenous arrival rates; if the $\lambda_l$ change by period $t$, we have to expand the formulation to use time-specific sales rates.

This formulation, referred to as the sales-based linear program ($SBLP$), vastly reduces the complexity of solving $CDLP$, albeit restricted to the MNL model and disjoint consideration sets.

## 3.2 Affine relaxation

The second approximation method we consider is the affine relaxation, where a functional form (in this case affine) is substituted for the value function variables in the formulation (2) as follows:

$$V_t(\boldsymbol{r}) = \theta_t + \sum_i V_{it} r_i.$$

The affine relaxation LP then becomes

$$\begin{aligned} V^{AF} = \quad & \min_{\theta, V} \quad \theta_1 + \sum_i V_{i1} r_i^1 \\ \text{s.t.} \quad & \\ (AF) \quad & \theta_t + \sum_i V_{it} r_i \geq \sum_j \lambda P_j(S)[f_j - \sum_{i \in j} V_{i,t+1}] + \theta_{t+1} + \sum_i V_{i,t+1} r_i \qquad (4) \\ & \qquad \forall \boldsymbol{r} \in \mathcal{R}, S \subset \mathcal{S}(\boldsymbol{r}), t, \\ & \theta_t \geq 0, V_{it} \geq 0, \end{aligned}$$

with the boundary conditions $\theta_{\tau+1} = 0, V_{i,\tau+1} = 0$. Zhang and Adelman [17] show that the optimal objective function value $V^{AF}$ is an upper bound on $V^{DP}$ and that there exists an optimal solution $(\hat{\theta}, \hat{V})$ of $AF$ that satisfies $\hat{V}_{it} - \hat{V}_{i,t+1} \geq 0$ for all $i$ and $t$.

The number of decision variables in $AF$ is manageable, but the number of constraints is of the order of $|\mathcal{R}|2^n\tau$. Vossen and Zhang [16] use Dantzig-Wolfe decomposition to derive a reduced formulation of $AF$, where the number of constraints is of the order of $2^n\tau$.

We give an alternative, simpler proof of the reduction. We begin by noting that constraints (4) can be written as

$$\min_{\boldsymbol{r} \in \mathcal{R}, S \subset \mathcal{S}(\boldsymbol{r})} \left\{ \theta_t - \theta_{t+1} + \sum_j \lambda P_j(S)[\sum_{i \in j} V_{i,t+1} - f_j] + \sum_i [V_{it} - V_{i,t+1}] r_i \right\} \geq 0 \qquad (5)$$

for all $t$. Now recall that an optimal solution to $AF$ satisfies $V_{it} - V_{i,t+1} \geq 0$ (shown in [17]). Therefore, the coefficient of $r_i$ in the minimization problem (5) is nonnegative, and we can assume $r_i \in \{0, 1\}$ in the minimization. Moreover, since $V_{it} - V_{i,t+1} \geq 0$, for any set $S$, we have $r_i = 0$ for $i \notin S$. On the other hand, feasibility requires we have $r_i = 1$ for $i \in S$. Therefore, (5) can be written as

$$\min_{S \subset J} \left\{ \theta_t - \theta_{t+1} + \sum_j \lambda P_j(S)[\sum_{i \in j} V_{i,t+1} - f_j] + \sum_i \mathbb{1}_{[S \in i]}[V_{it} - V_{i,t+1}] \right\} \geq 0. \qquad (6)$$

Now, just set $\beta_t = \theta_t - \theta_{t+1}$, and $\gamma_{it} = V_{it} - V_{i,t+1} \geq 0$, and observe that $\sum_{k=t+1}^{\tau} \gamma_{ik} = V_{i,t+1}$ when we use the boundary condition $V_{i,\tau+1} = 0$. With this change of variables, we can write $AF$ equivalently as

$$\begin{aligned} V^{RAF} = \min_{\beta, \gamma} \quad & \sum_t \beta_t + \sum_t \sum_i \gamma_{it} r_i^1 \\ \text{s.t.} \quad & \\ (RAF) \quad & \beta_t + \sum_i \left[ \mathbb{1}_{[S \ni i]} \gamma_{it} + (\sum_{k=t+1}^{\tau} \gamma_{ik}) \lambda Q_i(S) \right] \geq \lambda R(S) \quad \forall t, S \qquad (7) \\ & \gamma_{it} \geq 0. \end{aligned}$$

Notice that the number of constraints in the reduced formulation $RAF$ is an order of magnitude smaller than $AF$. Taking the dual of $RAF$ by associating dual variables $h_{t,S}$ with constraints (7), we get

$$\begin{aligned}
V^{dRAF} = \max_{h} \quad & \sum_{t, S \subseteq N} \lambda R(S) h_{t,S} \\
\text{s.t.} \quad & \sum_{k=1}^{t-1} \sum_{S \subseteq N} \lambda Q_i(S) h_{k,S} + \sum_{\substack{S \subseteq N \\ S \ni i}} h_{k,S} \leq r_i^1 \quad \forall i, t \\
(dRAF) \quad & \sum_{S \subseteq N} h_{t,S} = 1 \quad \forall t \\
& h_{t,S} \geq 0.
\end{aligned}$$

Vossen and Zhang [16] derive the formulation $dRAF$ using Dantzig-Wolfe decomposition. The above arguments also imply that

**Proposition 1.** $V^{AF} = V^{RAF} = V^{dRAF}$.

# 4 Tractable formulations for MNL with a single segment

In this section we restrict our attention to the MNL model with a single segment and develop our tractable approximations. We first show that the affine relaxation is NP-complete even for the MNL model with a single segment. On the other hand, for the same choice model $CDLP$ remains tractable. We compare the affine relaxation with $CDLP$, which gives crucial insight for deriving our tractable approximations. In §5 we extend the formulations to the MNL model with multiple segments, and disjoint consideration sets.

## 4.1 NP-completeness of the affine separation for single-segment MNL

Although $RAF$ has fewer constraints than $AF$, it is still exponential in the number of products. Therefore, we have to generate constraints (7) on the fly. Given a set of values $(\beta_t, \gamma_{it})$, the separation problem at time $t$ is to decide if constraint (7) is satisfied for all $S$, and if not, add the violated constraint to the LP.

Note that for the single-segment MNL, the formulas for the probabilities, expected resource consumptions and revenues are

$$P_j(S) = \frac{\mathbb{1}_{[j \in S]} w_j}{1 + \sum_{j' \in S} w_{j'}} \qquad Q_i(S) = \frac{\sum_{j \in S, j \ni i} w_j}{1 + \sum_{j \in S} w_j} \qquad R(S) = \frac{\sum_{j \in S} f_j w_j}{1 + \sum_{j \in S} w_j},$$

where we drop the segment superscript in $l$ in $w_j^l$ and write the weights as $w_j$.

Inequalities (7) for MNL then become, for a given $t$ and $S$,

$$\beta_t + \gamma_{St} + \sum_i \left[ \left( \sum_{k=t+1}^{\tau} \gamma_{ik} \right) \lambda \frac{\sum_{j \ni i, j \in S} w_j}{1 + \sum_{j \in S} w_j} \right] \geq \lambda \frac{\sum_{j \in S} f_j w_j}{1 + \sum_{j \in S} w_j}$$

where $\gamma_{St} = \sum_i \mathbb{1}_{[S \ni i]} \gamma_{it}$. Multiplying both sides by the positive quantity $1 + \sum_{j \in S} w_j$ and simplifying, constraint (7) can be written as

$$\beta_t + \sum_{j \in S} w_j \Big[ \beta_t + \lambda \big( \sum_{i \in j} ( \sum_{k=t+1}^{\tau} \gamma_{ik} ) - f_j \big) \Big] \geq -\gamma_{St} \big( 1 + \sum_{j \in S} w_j \big). \tag{8}$$

Therefore, for the single segment MNL model, constraint (7) in *RAF* can be replaced by constraint (8). The separation problem then is to find, for a given $t$ and a set of values of $(\beta_t, \gamma_{it})$, if a set $S$ violates constraint (8). For arbitrary values of $\beta_t$ and $\gamma_{it} \geq 0$ we show below that this is NP-complete. Proposition 2 below states that separation problem (5) for MNL with a single segment, as given in (8) is NP-complete.

**Proposition 2.** *The following problem is NP-complete: For any set of $w_j \geq 0$, $1 \geq \lambda \geq 0$, $f_j \geq 0$, and values $\beta_t$ and $\gamma_{it} \geq 0$, find a set $S$ that violates (8).*

This limits our ambitions of improving *CDLP* as the single-segment MNL is arguably the simplest possible choice model (after the independent-class model). Nevertheless, it is useful to compare *CDLP* with affine relaxation as we do next.

## 4.2 *CDLP* vs. *AF*

CDLP can be written equivalently, in an expanded redundant way with time-dependent variables, as follows:

$$V^{CDLP'} = \max_h \sum_{k=1}^{\tau} \sum_{S \subseteq J} \lambda R(S) h_{kS}$$

$$\text{s.t.} \quad \sum_{k=1}^{t} \sum_{S \subseteq J} \lambda h_{kS} \boldsymbol{Q}(S) \leq \boldsymbol{r}^1 \quad t = 1, \ldots, \tau \tag{9}$$

$$(CDLP') \qquad \sum_{S \subseteq J} h_{kS} = 1 \quad \forall k = 1, \ldots, \tau \tag{10}$$

$$0 \leq h_{kS}, \quad \forall\, S \subseteq J.$$

Letting $\gamma_{it}$ and $\beta_t$ be the dual variables corresponding to (9) and (10), respectively, the constraints in the dual of $CDLP'$ are $\beta_t + \sum_i \big( \sum_{k=t}^{\tau} \gamma_{ik} \big) \lambda Q_i(S) \geq \lambda R(S)$ for all $t$ and $S$. Using the single segment MNL formulas for the expected resource consumptions and expected revenues, the dual constraint can be written as

$$\beta_t + \sum_{j \in S} w_j \Big[ \beta_t + \lambda \big( \sum_{i \in j} ( \sum_{k=t}^{} \gamma_{ik} ) - f_j \big) \Big] \geq 0 \quad \forall t, S$$

which is almost identical to the left-hand-side of (8) except that the summation goes from $k = t$. To make the comparison easier, we rewrite the $CDLP'$ dual constraint as

$$\beta_t + \sum_{j \in S} w_j \Big[ \beta_t + \lambda \big( \sum_{i \in j} ( \sum_{k=t+1}^{\tau} \gamma_{ik} ) - f_j \big) \Big] \geq - \sum_{j \in S} w_j \lambda \sum_{i \in j} \gamma_{it} \quad \forall t, S. \tag{11}$$

We obtain now an easy comparison of the *CDLP* and the *AF* upper bounds.

**Proposition 3.** $V^{AF} \leq V^{CDLP}$.

Proof
By Proposition 1, $V^{AF} = V^{RAF}$ and by strong duality $CDLP'$ and its dual have the same optimal objective function value. So it suffices to compare the dual of $CDLP'$ with $RAF$. Notice that the objective function of the dual of $CDLP'$ is $\sum_t \beta_t + \sum_i \sum_t \gamma_{it} r_i^1$, which is identical to the objective function of $RAF$. As the objective functions are the same, we compare the constraints of $RAF$ with those of the dual of $CDLP'$, namely, (8) with (11). The left-hand sides of both inequalities are identical, and comparing the terms on the right hand side of both, we observe

$$\gamma_{St}(1 + \sum_{j \in S} w_j) \geq \lambda \sum_{j \in S} w_j (\sum_{i \in j} \gamma_{it})$$

as $0 \leq \lambda \leq 1$, and $\gamma_{St} \geq \sum_{i \in j} \gamma_{it} \geq 0$ for all $j \in S$. Thus the affine relaxation has a larger feasible region in this minimization linear program.

□

The relation between the $CDLP$ and $AF$ bounds is shown in [17], but our short proof for the single-segment MNL model gives crucial insight for deriving tractable relaxations. In the following, we show how to obtain relaxations that remain tractable for the single-segment MNL model and that lie in between the $CDLP$ and the $AF$ bounds, by concentrating our attention on the right-hand side of (8).

### 4.3 Weak affine relaxation 1

We motivate our first relaxation as follows: The difficult term in (8) is the $\gamma_{St}(1 + \sum_{j \in S} w_j)$, and $CDLP$ is tractable as it replaces this by $\lambda \sum_{j \in S} w_j (\sum_{i \in j} \gamma_{it})$. We instead replace the right hand side of constraint (8) with $-\gamma_{St} - \sum_{j \in S} w_j (\sum_{i \in j} \gamma_{it})$ and solve the linear program

$$\begin{aligned} V^{wAR1} = \min_{\beta, \gamma} \quad & \sum_t \beta_t + \sum_t \sum_i \gamma_{it} r_i^1 \\ \text{s.t.} \quad & \\ & \beta_t + \sum_{j \in S} w_j [\beta_t + \lambda (\sum_{i \in j} (\sum_{k=t+1}^{\tau} \gamma_{ik}) - f_j)] \geq -\gamma_{St} - \sum_{j \in S} w_j (\sum_{i \in j} \gamma_{it}) \quad \forall t, S \quad (12) \\ & \gamma_{it} \geq 0. \end{aligned}$$

Proposition 4 below shows that the above LP gives an upper bound on the value function that is weaker than the $AF$ bound but stronger than CDLP.

**Proposition 4.** $V^{AF} \leq V^{wAR1} \leq V^{CDLP}$.

Proof
The proof follows by noting that $\gamma_{St}(1+\sum_{j \in S} w_j) \geq \gamma_{St} + \sum_{j \in S} w_j (\sum_{i \in j} \gamma_{it}) \geq \lambda \sum_{j \in S} w_j (\sum_{i \in j} \gamma_{it})$ and using the same arguments as in the proof of Proposition 3.

□

In the remainder of this section, we show that weak affine relaxation 1 can be formulated as a compact linear program and is therefore tractable. We begin by observing that constraint (12) is easy to separate. Given a set of values $(\beta_t, \gamma_{it})$, the separation problem at time $t$ can be formulated as the integer program

$$\begin{aligned}
V^{S1_t} = \quad & \max_{q,u} \quad -\sum_i \gamma_{it} q_i - \sum_j w_j \Big[\beta_t + \lambda\big(\sum_{i\in j}(\sum_{k=t+1}^{\tau} \gamma_{ik}) - f_j\big) + \sum_{i \in j} \gamma_{it}\Big] u_j \\
& u_j - q_i \leq 0 \quad \forall i \in j, \forall j && (13) \\
(S1_t) \quad & q_i \leq 1 \quad \forall i && (14) \\
& u_j \geq 0, \text{integer}
\end{aligned}$$

where the decision variables $q_i$ and $u_j$, respectively, indicate if resource $i$ and product $j$ are open. The first constraint ensures that a product is open only if all the resources it consumes are open. If $V^{S1_t} \leq \beta_t$, constraint (12) is satisfied by all $S$ at time $t$; otherwise there is a set which violates the constraints and we add it to the LP.

Now, just observe that the constraint matrix of the above integer program has exactly one $+1$ and one $-1$ coefficients in each row, and hence is totally unimodular. So we can ignore the integer restriction and solve $S1_t$ exactly as a linear program. Notice that in the above formulation we used the fact that $\gamma_{it} \geq 0$ to allow $q_i$ to be unrestricted.

Since the separation problem can be solved as an LP, we can fold it in into the original LP to obtain a compact formulation as follows: First take the dual of $S1_t$ with dual variables $\pi_{ij}$ corresponding to (13), and $\psi_j$ to (14) :

$$\begin{aligned}
V^{dS1_t} = \quad & \min_{\pi,\psi} \quad \sum_i \psi_i \\
\text{s.t.} \quad & \sum_{i\in j} \pi_{ij} \geq -w_j \Big[\beta_t + \lambda\big(\sum_{i\in j}(\sum_{k=t+1}^{\tau} \gamma_{ik}) - f_j\big) + \sum_{i\in j} \gamma_{it}\Big] \quad \forall j \\
(dS1_t) \quad & -\sum_{j \ni i} \pi_{ij} + \psi_i = \gamma_{it} \quad \forall i \\
& \pi_{ij}, \psi_i \geq 0.
\end{aligned}$$

Then use the second constraint in the above LP to eliminate the variable $\psi_i$ to write the dual as

$$\begin{aligned}
V^{dS1_t} = \quad & \min_{\pi} \quad \sum_i \Big[\sum_{j\ni i} \pi_{ij} - \gamma_{it}\Big] \\
\text{s.t.} \quad & \sum_{i\in j} \pi_{ij} \geq -w_j \Big[\beta_t + \lambda\big(\sum_{i\in j}(\sum_{k=t+1}^{\tau} \gamma_{ik}) - f_j\big) + \sum_{i\in j} \gamma_{it}\Big] \quad \forall j \\
& \sum_{j\ni i} \pi_{ij} \geq \gamma_{it} \quad \forall i \\
& \pi_{ij} \geq 0.
\end{aligned}$$

Thus constraints (12) amount to a condition that $V^{dS1_t} \leq \beta_t$ for all $t$, which can be written, in lieu of (12), as (15–17) below. Putting everything together, the linear program for our first weak affine relaxation in its entirety can be written as:

$$
\begin{aligned}
V^{wAR1} = \quad & \min_{\beta,\gamma,\pi} \quad \sum_t \beta_t + \sum_t \sum_i r_i^1 \gamma_{it} \\
\text{s.t.} \quad & \beta_t \geq \sum_i \Big[ \sum_{j \ni i} \pi_{ijt} - \gamma_{it} \Big] \quad \forall t && (15) \\
(wAR1) \quad & \sum_{i \in j} \pi_{ijt} \geq -w_j \Big[ \beta_t + \lambda \big( \sum_{i \in j} ( \sum_{k=t+1}^{\tau} \gamma_{ik}) - f_j \big) + \sum_{i \in j} \gamma_{it} \Big] \quad \forall t, j && (16) \\
& \sum_{j \ni i} \pi_{ijt} \geq \gamma_{it} \quad \forall i, t && (17) \\
& \gamma_{it}, \pi_{ijt} \geq 0.
\end{aligned}
$$

The size of $wAR1$ is polynomial in the number of resources, products and the length of the booking horizon. Hence, not only is it stronger than CDLP, it is also tractable. Notice that this formulation would have been hard to derive and justify without the line of reasoning starting from $AF$.

The dual of $wAR1$ gives more insight into the formulation. By associating dual variables with constraints (15), (16), and (17), and after some simplifications, we get the dual LP as

$$
\begin{aligned}
V^{dwAR1} = \quad & \max_{x,\rho} \quad \lambda \sum_t \sum_j f_j x_{jt} \\
\text{s.t} \quad & x_{0t} + \sum_{s=1}^{t-1} \sum_{j \ni i} \lambda x_{js} + \sum_{j \ni i} x_{jt} - \rho_{it} \leq r_i^1 \quad \forall i, t \\
(dwAR1) \quad & x_{0t} + \sum_j x_{jt} = 1 \quad \forall t \\
& \frac{x_{jt}}{w_j} - x_{0t} + \rho_{it} \leq 0 \quad \forall i, j \in i, t \\
& x_{0t}, x_{jt}, \rho_{it} \geq 0.
\end{aligned}
$$

Using the same interpretation as $SBLP$, the decision variable $x_{jt}$ can be viewed as the sales rate for product $j$ and at time $t$, and the variable $x_{0t}$ as the no-purchase rate at time $t$.

## 4.4 Weak affine relaxation 2

In this section we consider a complementary relaxation that uses a different bound than weak affine relaxation 1. Our second relaxation is based on the following identity:

**Proposition 5.** $\gamma_{St}\big(\sum_{j \in S} w_j\big) \geq \sum_{i,j} \big[\mathbb{1}_{[S \ni i]} + \mathbb{1}_{[S \ni j]} - 1\big] \gamma_{it} w_j + \sum_i \big[1 - \mathbb{1}_{[S \ni i]}\big] \big(\sum_{j \ni i} \gamma_{it} w_j\big)$.

Proof
Appendix.

□

So just as we did before, we replace the right hand side of constraints (8) with $-\gamma_{St} - \sum_{i,j} \big[\mathbb{1}_{[S \ni i]} +$

$\mathbb{1}_{[S\ni j]} - 1\big]\gamma_{it}w_j - \sum_i \big[1 - \mathbb{1}_{[S\ni i]}\big]\big(\sum_{j\ni i}\gamma_{it}w_j\big)$ and solve the LP

$$
\begin{aligned}
V^{wAR2} = \min_{\beta,\gamma} \quad & \sum_t \beta_t + \sum_t \sum_i \gamma_{it} r_i^1 \\
\text{s.t.} \quad & \\
& \beta_t + \sum_{j\in S} w_j \Big[\beta_t + \lambda\big(\sum_{i\in j}(\sum_{k=t+1}^{\tau}\gamma_{ik}) - f_j\big)\Big] \geq \qquad (18)\\
& \quad -\gamma_{St} - \sum_{i,j}\big[\mathbb{1}_{[S\ni i]} + \mathbb{1}_{[j\in S]} - 1\big]\gamma_{it}w_j - \sum_i [1 - \mathbb{1}_{[S\ni i]}]\big(\sum_{j\ni i}\gamma_{it}w_j\big) \forall t, S \\
& \gamma_{it} \geq 0.
\end{aligned}
$$

The separation problem (18) can be solved as the following integer program with the same totally-unimodular constraint matrix as before, and checking if its value is $\leq \beta_t$:

$$
\begin{aligned}
\max_{q,u} \quad & -\sum_j w_j[\beta_t + \lambda(\sum_{i\in j}(\sum_{k=t+1}^{\tau}\gamma_{ik}) - f_j)]u_j - \sum_{i,j}[u_j + q_i - 1]\gamma_{it}w_j \\
& \qquad - \sum_i [1 - q_i]\gamma_{it}(\sum_{j\ni i} w_j) - \sum_i \gamma_{it} q_i \\
(S2_t) \quad & u_j - q_i \leq 0 \quad \forall i \in j, \forall j \\
& q_i \leq 1 \quad \forall i \\
& u_j \geq 0, \text{integer}.
\end{aligned}
$$

This can be folded into the original problem to obtain a compact formulation the same way as in §4.3; we omit the details.

Neither of the two relaxations is uniformly tighter than the other. In our numerical experiments in §7 weak affine relaxation 1 generally tends to produce tighter bounds, but there are instances where weak affine relaxation 2 is tighter.

# 5 Formulation for multiple segments with disjoint consideration sets

We now extend the weak affine relaxation 1 of §4.3 to incorporate multiple segments, with MNL choice model and disjoint consideration sets (the development for §4.4 is similar). We show that our formulation gives an upper bound that falls in between the $CDLP$ and $AF$ bounds, as for the single segment case.

Let $\mathcal{I}_l = \{i | \exists j \in \mathcal{C}_l \text{ and } j \ni i\}$ and $\mathcal{L}_i = \{l | i \in \mathcal{I}_l\}$ and recall that we use, for $S \subset J$, $S \ni i$ if there is some $j \in S$ with $j \ni i$.

We begin with the separation problem for $AF$ as given in (6) but expressed in terms of $\beta_t$ and $\gamma_{it}$'s: For time $t$, check for all $S$ if the following holds true

$$
\beta_t + \gamma_{St} + \sum_i \left[(\sum_{k=t+1}^{\tau}\gamma_{ik})\sum_l \lambda_l Q_i^l(S)\right] \geq \sum_l \lambda_l R^l(S), \qquad (19)
$$

where we recall that $\gamma_{St} = \sum_i \mathbb{1}_{[S\ni i]}\gamma_{it}$. Our idea is now to split this constraint into $l$ separate constraints by introducing variables $\beta_{lt}$. In addition, we relax the constraint and only require that for each $l$ and $S_l = S \cap \mathcal{C}_l$, the following holds

$$\beta_{lt} + \sum_{i\in\mathcal{I}_l} \gamma_{S_l,t} \frac{\lambda_l}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} \left( \sum_{j\in S_l} P_j^l(S_l) + P_0^l(S_l) \right) + \sum_i \left[ (\sum_{k=t+1}^{\tau} \gamma_{ik})\lambda_l Q_i^l(S_l) \right] \geq \lambda_l R^l(S_l). \quad (20)$$

Since $\sum_{j\in S_l} P_j^l(S_l) + P_0^l(S_l) = 1$, and as $\mathbb{1}_{[S_l\ni i]} = \mathbb{1}_{[S\cap\mathcal{C}_l\ni i]}$, it follows that $\sum_l \sum_{i\in\mathcal{I}_l} \gamma_{S_l,t} \frac{\lambda_l}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} = \sum_l \sum_{i\in\mathcal{I}_l} \mathbb{1}_{[S\cap\mathcal{C}_l\ni i]} \frac{\lambda_l\gamma_{it}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} = \sum_{i\in S}\gamma_{it} = \gamma_{St}$. Therefore, summing (20) over all segments $l$ we obtain (19) with $\beta_t = \sum_l \beta_{lt}$. Therefore, the segment level constraints (20) imply (19). As this is a minimization problem, we obtain a looser upper bound by separating over (20) instead of (19).

Now using the same relaxation as we did for the single-segment case to obtain (12), we obtain a segment-based weak affine relaxation for MNL that solves the following separation problem $S1_{l,t}$, which is of the same form as $S1_t$, for each segment $l$ (the full derivation is in the Appendix under proof of Proposition 6):

$$\begin{aligned}
V^{S1_{l,t}} = \quad & \max_{q,u} \quad \sum_{j\in\mathcal{C}_l} \lambda_l \Big[ f_j - \sum_{i\in j}(\sum_{k=t+1}^{\tau}\gamma_{ik} + \frac{\gamma_{it}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}}) - \frac{\beta_{lt}}{\lambda_l} \Big] w_j u_j - \sum_{i\in\mathcal{I}_l} \frac{\lambda_l\gamma_{it}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} q_i \\
\text{s.t} \quad & u_j - q_i \leq 0 \quad \forall i\in j, j\in\mathcal{C}_l \\
(S1_{l,t}) \quad & q_i \leq 1 \quad \forall i\in\mathcal{I}_l \\
& u_j \geq 0 \quad \forall j\in\mathcal{C}_l.
\end{aligned}$$

Now folding in the separation problem by taking the dual of $S1_{l,t}$, and following the same steps as for the single-segment case, we get the following linear program:

$$\begin{aligned}
V^{swAR1} = \quad & \min_{\gamma,\beta,\pi} \quad \sum_i \sum_t r_i^1 \gamma_{it} + \sum_t \sum_l \beta_{lt} \\
\text{s.t} \quad & \beta_{lt} \geq \sum_{i\in\mathcal{I}_l} [ \sum_{j\in i, j\in\mathcal{C}_l} \pi_{ijt} - \frac{\lambda_l}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}}\gamma_{it}] \quad \forall l,t \\
(swAR1) \quad & \sum_{i\in j} \pi_{ijt} \geq \lambda_{\ell_j} w_j \Big[ f_j - \sum_{i\in j}(\sum_{k=t+1}^{\tau}\gamma_{ik} + \frac{\gamma_{it}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}}) - \frac{\beta_{\ell_j,t}}{\lambda_{\ell_j}}] \quad \forall j,t \\
& \sum_{j\in i, j\in\mathcal{C}_l} \pi_{ijt} - \frac{\lambda_l}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}}\gamma_{it} \geq 0 \quad \forall i, l\in\mathcal{L}_i, t \\
& \gamma_{it}, \pi_{ijt} \geq 0,
\end{aligned}$$

where $\ell_j$ denotes the segment to which product $j$ belongs. $swAR1$ can be viewed as the extension of $wAR1$ to the MNL model with multiple segments and disjoint consideration sets. Note that $swAR1$ is again tractable as it is a compact linear program. Proposition 6 below shows that it also obtains an upper bound on the value function that is tighter than $CDLP$.

**Proposition 6.** $V^{AF} \leq V^{swAR1} \leq V^{CDLP}$.

Proof
Appendix.

□

As we show in the next section, it is possible to extend the formulation to the MNL model with multiple segments when the consideration sets overlap. The dual of $swAR1$, which we give below, turns out to be useful for this purpose.

$$\begin{aligned}
V^{dswAR1} = \quad & \max_{x,\rho} \quad \sum_{l} \lambda_l [\sum_{t} \sum_{j \in \mathcal{C}_l} f_j x^l_{jt}] \\
\text{s.t} \quad & \sum_{l} \lambda_l \left[ \frac{x^l_{0t}}{\sum_{l' \in \mathcal{L}_i} \lambda_{l'}} + \sum_{s=1}^{t-1} \sum_{j \ni i, j \in \mathcal{C}_l} x^l_{js} + \frac{\sum_{j \ni i, j \in \mathcal{C}_l} x^l_{jt}}{\sum_{l' \in \mathcal{L}_i} \lambda_{l'}} - \frac{\rho^l_{it}}{\sum_{l' \in \mathcal{L}_i} \lambda_{l'}} \right] \leq r^1_i \quad \forall i, t \\
(dswAR1) \quad & x^l_{0t} + \sum_{j \in \mathcal{C}_l} x^l_{jt} = 1 \quad \forall l, t \\
& \frac{x^l_{jt}}{w_j} - x^l_{0t} + \rho^l_{it} \leq 0 \quad \forall l, i, j \ni i, j \in \mathcal{C}_l, t \qquad (21) \\
& x^l_{0t}, x^l_{jt}, \rho^l_{it} \geq 0.
\end{aligned}$$

# 6 Overlapping consideration sets

When the segment consideration sets overlap, the $CDLP$ formulation is difficult to solve, even for MNL with just two segments. So one would imagine that it is difficult to find a tractable bound tighter than $CDLP$ in this case. One strategy, pursued in Meissner et al. [9] is to formulate the problem by segments and then add a set of consistency conditions called *product-cut* equalities (PC-equalities). These equalities apply to any general discrete-choice model and appear to be quite powerful in numerical experiments, often bringing the solution close to $CDLP$ value. Strauss and Talluri [12] subsequently show that when the consideration set structure has a certain tree structure, the cuts in fact achieve the $CDLP$ value.

In this section we show that the PC-equalities, specialized for MNL, continue to be valid for our formulation $dswAR1$, in the sense that after adding them, the value of the resulting linear program is an upper bound on $V^{DP}$.

## 6.1 PC-equalities

We first state the equalities for MNL, to be added to the formulation $SBLP$ of (2) (an explanation for their validity can be found in [13]):

$$\frac{x^l_k}{\tau w^l_k} = \lambda_l \sum_{\{S_{lm}\subseteq(\mathcal{C}_l\cap\mathcal{C}_m)|S_{lm}\ni k\}} x^{lm}_{S_{lm}}, \quad \forall k \in \mathcal{C}_l \cap \mathcal{C}_m,\ \forall l,m \tag{22}$$

$$x^{lm}_{S_{lm},k} \leq x^{lm}_{S_{lm}}, \quad \forall S_{lm} \subseteq \mathcal{C}_l \cap \mathcal{C}_m, k \in \mathcal{C}_l \setminus \mathcal{C}_m,\ \forall l,m \tag{23}$$

$$\sum_{\{S_{lm}\subseteq(\mathcal{C}_l\cap\mathcal{C}_m)|S_{lm}\supseteq\tilde{S}_{lm}\}} \left\{ \sum_{k\in\mathcal{C}_l\setminus\mathcal{C}_m} w^l_k x^{lm}_{S_{lm},k} + (1+w^l_{S_{lm}})x^{lm}_{S_{lm}} \right\} = \tag{24}$$

$$\sum_{\{S_{ml}\subseteq(\mathcal{C}_m\cap\mathcal{C}_l)|S_{ml}\supseteq\tilde{S}_{lm}\}} \left\{ \sum_{k\in\mathcal{C}_m\setminus\mathcal{C}_l} w^m_k x^{ml}_{S_{ml},k} + (1+w^m_{S_{ml}})x^{ml}_{S_{ml}} \right\}, \forall \tilde{S}_{lm} \subseteq \mathcal{C}_l \cap \mathcal{C}_m, \forall l,m$$

We refer to $SBLP$ with (22–24) as $SBLP$+; note the new variables of the form $x^{lm}_{S_{lm}}$ defined for all pairs of segments $l,m$ and for all $S_{lm} \subseteq \mathcal{C}_l \cap \mathcal{C}_m$; and $w^l_{S_{lm}} = \sum_{j\in S_{lm}} w^l_j$. While not compact, when the size of the intersections $|\mathcal{C}_l \cap \mathcal{C}_m|$ is small, this formulation is tractable.

Now we show that (22–24) can be added to $dswAR1$ and the resulting linear program gives an upper bound on $V^{DP}$.

## 6.2 Validity of PC-equalities for Weak Affine formulation

To show validity, as the feasible region of $DP$ is contained in the feasible regions of $dswAR1$ as well as that of $SBLP$+ (the latter shown in [13]), all we have to show is that the feasible region of $dswAR1$ is contained in the feasible region of $SBLP$. This fact is implied by Proposition 6, but we give a direct proof: We make the connection between $dswAR1$ and $SBLP$ variables via $x^l_j = \lambda_l \sum_{t=1}^{\tau} x^l_{jt}$ and $x^l_0 = \lambda_l \sum_{t=1}^{\tau} x^l_{0t}$. So $SBLP$ can be written in terms of the time-indexed variables $x^l_{jt}$ and we consider $dwAR1$ as an extended formulation with new variables $\rho^l_{it}$, and the projection of $dswAR1$ into the space of the variables $x^l_{jt}$ is now shown to be a subset of $SBLP$. Consider a feasible solution of $dswAR1$. The solution clearly satisfies $\sum_t \lambda_l(x^l_{0t} + \sum_{j\in\mathcal{C}_l} x^l_{jt}) = \lambda_l\tau$ and hence $x^l_0 + \sum_{j\in\mathcal{C}_l} x^l_j = \lambda_l\tau$. Likewise, $\frac{x^l_j}{w^l_j} - x^l_0 \leq -\lambda_l \sum_t \rho^l_{it} \leq 0$.

So the only remaining set of constraints to verify is (3). Consider the constraints of $dswAR1$ for period $\tau$:

$$\sum_l \lambda_l \left[ \frac{x^l_{0\tau}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} + \sum_{s=1}^{\tau-1} \sum_{j\ni i, j\in\mathcal{C}_l} x^l_{js} + \frac{\sum_{j\ni i,j\in\mathcal{C}_l} x^l_{j\tau}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} - \frac{\rho^l_{i\tau}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} \right] \leq r^1_i$$

which can be rewritten as

$$\sum_l \sum_{j\ni i,j\in\mathcal{C}_l} (x^l_j - \lambda_l x^l_{j\tau}) + \lambda_l \left[ \frac{x^l_{0\tau}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} + \frac{\sum_{j\ni i,j\in\mathcal{C}_l} x^l_{j\tau}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} - \frac{\rho^l_{i\tau}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} \right] \leq r^1_i.$$

So it is enough to show

$$\sum_l \lambda_l \left[ \frac{x^l_{0\tau}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} + \frac{\sum_{j\ni i,j\in\mathcal{C}_l} x^l_{j\tau}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} - \frac{\rho^l_{i\tau}}{\sum_{l'\in\mathcal{L}_i}\lambda_{l'}} \right] \geq \sum_l \lambda_l \sum_{j\ni i,j\in\mathcal{C}_l} x^l_{j\tau},$$

or

$$\sum_{l} \lambda_l x^l_{0\tau} + \lambda_l (1 - \sum_{l' \in \mathcal{L}_i} \lambda_{l'}) \sum_{j \ni i, j \in \mathcal{C}_l} x^l_{j\tau} - \lambda_l \rho^l_{i\tau} \geq 0,$$

which is true as (21) implies $\rho^l_{i\tau} \leq x^l_{0\tau}$ and $(1 - \sum_{l' \in \mathcal{L}_i} \lambda_{l'}) \geq 0$.

So in conclusion, when segment consideration sets overlap, we also have

**Proposition 7.** *The objective function value of $swAR1$ with (22–24) added, with $x^l_k = \lambda_l \sum_{t=1}^{\tau} x^l_{kt}$ in (22), is less than or equal to $V^{SBLP+}$.*

# 7 Computational Experiments

In this section, we compare the upper bounds obtained by the tractable formulations with those obtained by the choice deterministic linear program and the affine relaxation on different test problems. We test the performance of our benchmark solution methods on a parallel flights network and a hub and spoke network, with a single hub serving multiple spokes. In all of our test problems, we have multiple customer segments with disjoint consideration sets and choice within each segment is governed by the MNL model. We begin by describing the different benchmark solution methods and the experimental setup.

*Choice deterministic linear program ($CDLP$)* This is the solution method described in §3.1. Since all our test problems involve the MNL choice model with disjoint consideration sets, we use the compact sales-based formulation, $SBLP$, to compute the upper bound.

*Weak affine relaxation 1 ($wAR1$)* This is the version of weak affine relaxation 1 that applies to multiple segments and described in §5 ($swAR1$).

*Weak affine relaxation 2 ($wAR2$)* This is the version of weak affine relaxation 2 that applies to multiple segments. As mentioned, it is possible to extend the weak affine relaxation 2 method described in §4.4 to the setting with multiple segments by following the steps in §5.

*Affine relaxation ($AF$)* This is the solution method described in §3.2. We use the reduced formulation $RAF$ of [16] to compute the affine relaxation upper bound. Note that while the number of decision variables in $RAF$ is manageable, it has a large number of constraints. We solve $RAF$ by generating constraints on the fly (using integer programming) and stop when we are within 1% of optimality.

## 7.1 Parallel Flights

We consider $N$ parallel flights that operate between the same origin-destination pair. Note that the flight legs correspond to the resources in our network RM formulation. There is a high fare-product and a low fare-product on each flight leg so that the total number of products is $2N$. The high fare-product is 50% more expensive than the low fare-product.

We have two customer segments. The first segment is interested only in the low fare-products while the second segment is interested only in the high fare-products. So the consideration sets of the two segments are disjoint. Moreover, within each segment choice is according to the MNL model.

We measure the tightness of the leg capacities using the nominal load factor, which is defined in the following manner. Letting $\hat{S}_{lt} = \text{ argmax }_{S_l} R^l(S_l)$ denote the optimal set of products offered to segment $l$ at time period $t$ when there is ample capacity on all flight legs, we define the nominal load factor

$$\alpha = \frac{\sum_l \sum_t \sum_i \lambda_{lt} Q_i^l(\hat{S}_{lt})}{\sum_i r_i^1},$$

where $\lambda_{lt}$ denotes the arrival rate for segment $l$ at time period $t$.

We consider one set of test problems where the arrival rates remain the same throughout the booking period. The total arrival rate in each period is 0.9. We refer to these test problems as stationary arrivals. We also consider a second set of test problems where we divide the booking period into three intervals of equal length. The arrival rates remain the same within each interval, but increase from the first interval to the third. The total arrival rate in the first, second, and third intervals are 0.3, 0.6 and 0.9, respectively. We refer to the second set of test problems as non-stationary arrivals. For both stationary and non-stationary arrivals, we label our test problems by $(N, \alpha)$ where $N \in \{4, 6, 8\}$ and $\alpha \in \{0.8, 1.0, 1.2, 1.6\}$. We have $\tau = 200$ in all of our test problems.

Table 1 gives the upper bounds obtained by $CDLP$, $wAR1$, $wAR2$ and $AF$ for parallel flights with stationary arrivals. The first column in Table 1 gives the problem characteristics. The second, third, fourth and fifth columns give the upper bounds obtained by $CDLP$, $wAR1$, $wAR2$, and $AF$, respectively. The last three columns give the percentage gap between the upper bounds obtained by $CDLP$ and $wAR1$, $CDLP$ and $wAR2$, and $CDLP$ and $AF$, respectively. The upper bounds obtained by $wAR1$ are on average 0.19% tighter than $CDLP$. $wAR2$ and $AF$ obtain upper bounds that are on average 0.13% and 0.38% tighter than $CDLP$. $wAR1$ in general obtains tighter bounds than $wAR2$, although we observe instances where the upper bound obtained by $wAR2$ is slightly tighter.

Table 2 gives the upper bounds obtained by the four benchmark solution methods for parallel flights with non-stationary arrivals. The columns have the same interpretation as before. The percentage gap between $CDLP$ and the other three solution methods increases compared to the stationary arrivals case. $wAR1$, $wAR2$, and $AF$ on average produce upper bounds that are 0.7%, 0.2%, and 1.45% tighter than $CDLP$. $wAR1$ obtains upper bounds that are noticeably tighter than $wAR2$, and roughly closes 50% of the gap between the $CDLP$ and $AF$ upper bounds.

| Problem | Upper Bound | | | | % Gap with $CDLP$ | | |
|---|---|---|---|---|---|---|---|
| $(N, \alpha)$ | $CDLP$ | $wAR1$ | $wAR2$ | $AF$ | $wAR1$ | $wAR2$ | $AF$ |
| (4, 0.8) | 11,101 | 11,079 | 11,077 | 11,050 | 0.20 | 0.22 | 0.46 |
| (4, 1.0) | 9,899 | 9,864 | 9,882 | 9,848 | 0.36 | 0.17 | 0.52 |
| (4, 1.2) | 8,342 | 8,342 | 8,342 | 8,341 | 0.00 | 0.00 | 0.01 |
| (4, 1.6) | 6,217 | 6,217 | 6,217 | 6,217 | 0.00 | 0.00 | 0.00 |
| (6, 0.8) | 12,880 | 12,850 | 12,834 | 12,807 | 0.23 | 0.36 | 0.57 |
| (6, 1.0) | 11,667 | 11,599 | 11,633 | 11,548 | 0.58 | 0.29 | 1.02 |
| (6, 1.2) | 9,861 | 9,859 | 9,861 | 9,843 | 0.02 | 0.00 | 0.18 |
| (6, 1.6) | 7,460 | 7,460 | 7,460 | 7,460 | 0.00 | 0.00 | 0.00 |
| (8, 0.8) | 12,695 | 12,690 | 12,684 | 12,673 | 0.04 | 0.09 | 0.18 |
| (8, 1.0) | 11,817 | 11,720 | 11,770 | 11,647 | 0.82 | 0.40 | 1.43 |
| (8, 1.2) | 10,070 | 10,063 | 10,070 | 10,049 | 0.06 | 0.00 | 0.20 |
| (8, 1.6) | 7,524 | 7,524 | 7,524 | 7,524 | 0.00 | 0.00 | 0.01 |
| | | | | avg. | 0.19 | 0.13 | 0.38 |

Table 1: Comparison of the upper bounds for the parallel flights test problems with stationary arrival rates.

| Problem | | Upper Bound | | | % Gap with $CDLP$ | | |
|---|---|---|---|---|---|---|---|
| $(N,\alpha)$ | $CDLP$ | $wAR1$ | $wAR2$ | $AF$ | $wAR1$ | $wAR2$ | $AF$ |
| (4, 0.8) | 7,935 | 7,935 | 7,935 | 7,935 | 0.00 | 0.00 | 0.00 |
| (4, 1.0) | 7,862 | 7,838 | 7,840 | 7,806 | 0.30 | 0.27 | 0.71 |
| (4, 1.2) | 6,877 | 6,841 | 6,865 | 6,806 | 0.53 | 0.18 | 1.04 |
| (4, 1.6) | 5,504 | 5,467 | 5,492 | 5,439 | 0.67 | 0.23 | 1.19 |
| (6, 0.8) | 7,537 | 7,524 | 7,527 | 7,501 | 0.18 | 0.13 | 0.48 |
| (6, 1.0) | 6,615 | 6,569 | 6,605 | 6,514 | 0.70 | 0.16 | 1.53 |
| (6, 1.2) | 5,819 | 5,768 | 5,807 | 5,720 | 0.87 | 0.21 | 1.70 |
| (6, 1.6) | 4,704 | 4,659 | 4,694 | 4,614 | 0.97 | 0.22 | 1.92 |
| (8, 0.8) | 7,043 | 6,981 | 7,016 | 6,919 | 0.88 | 0.38 | 1.76 |
| (8, 1.0) | 5,959 | 5,901 | 5,949 | 5,829 | 0.98 | 0.18 | 2.19 |
| (8, 1.2) | 5,165 | 5,101 | 5,153 | 5,037 | 1.24 | 0.24 | 2.48 |
| (8, 1.6) | 4,225 | 4,180 | 4,217 | 4,125 | 1.06 | 0.17 | 2.37 |
| | | | | avg. | 0.70 | 0.20 | 1.45 |

Table 2: Comparison of the upper bounds for the parallel flights test problems with non-stationary arrival rates.

## 7.2 Hub and Spoke Network

We consider a hub and spoke network with a single hub that serves $N$ spokes. Half of the spokes have two flights to the hub, while the remaining half have two flights from the hub so that the total number of flights is $2N$. Figure 1 shows the structure of the network with $N = 8$. The

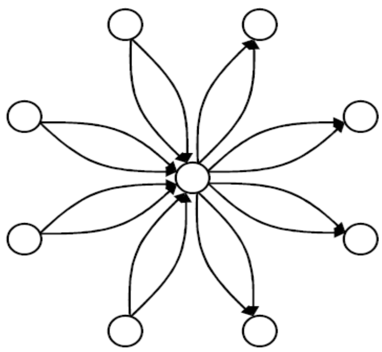

Figure 1: Structure of the airline network with a single hub and eight spokes.

total number of fare-products is $2N(N + 2)$. There are $4N$ fare products connecting spoke-to-hub and hub-to-spoke origin-destination pairs, of which half are high fare-products and the remaining half are low-fare products. The high fare-product is 50% more expensive than the corresponding low fare-product. The remaining $4N^2$ fare-products connect spoke-to-spoke origin-destination pairs. Half of the $4N^2$ fare-products are high fare-products and the rest are low fare-products, with the high fare-product being 50% more expensive than the corresponding low fare-product.

Each origin-destination pair is associated with a customer segment and each segment is only interested in the fare-products connecting that origin-destination pair. Therefore, the consideration sets are disjoint. Within each segment choice is governed by the MNL model. As in the parallel flights case, we consider two sets of test problems, one with stationary arrival rates and the second with non-stationary arrivals. We label our test problems by $(N, \alpha)$ where $N \in \{4, 6, 8\}$ and $\alpha \in \{0.8, 1.0, 1.2, 1.6\}$, which gives us 24 test problems in total. We use $\tau = 200$ in all of our test problems.

Table 3 compares the upper bounds obtained by the four solution methods for the hub and spoke network with stationary arrivals. As expected, $AF$ generates the tightest upper bound and $CDLP$ the weakest, with the $wAR1$ and $wAR2$ upper bounds sandwiched in between. $wAR1$ tends to generate tighter upper bounds than $wAR2$, although we observe once instance where the $wAR2$ bound is tighter. The average percentage gap between $wAR1$ and $CDLP$ is 1.59%, although we observe instances where the gap is as high as 2.73%. The percentage gap between $wAR1$ and $CDLP$ seems to increase with the nominal load factor and the number of spokes in the network. The average percentage gap between $wAR2$ and $CDLP$ is 0.95%, although we notice gaps as large as 1.9%. $AF$ generates bounds that are on average 2.16% tighter than the $CDLP$ bound.

Table 4 compares the upper bounds obtained by the benchmark solution methods for the hub and spoke network with non-stationary arrivals. The results display the same trends as before. The percentage gaps between $CDLP$ and the other three solution methods increases, on average, compared to the stationary arrivals case. $wAR1$, $wAR2$ and $AF$ produce upper bounds that are on average 2.75%, 1.6% and 3.43% tighter than the $CDLP$ bound. The nominal load factor and the number of the spokes in the network seem to be two factors which lead to larger gaps. Overall, $wAR1$ tends to generate tighter bounds than $wAR2$, and it closes roughly 80% of the gap between the $CDLP$ and $AF$ upper bounds.

Table 5 gives the CPU seconds required by the different solution methods for different numbers of spokes in the network and different numbers of time periods in the booking horizon. All of our computational experiments are carried out on a Pentium Core 2 Duo desktop with 3-GHz CPU and 4-GB RAM. We use CPLEX 11.2 to solve all linear programs. The running time of $CDLP$ is of the order of seconds, while those of $wAR1$, $wAR2$ and $AF$ are generally in minutes. $wAR1$ typically runs faster than $AF$ and the savings can be significant especially for relatively large networks. In light of the hardness result in Proposition 2, we only expect the savings in run times to increase with the problem size.

| Problem | Upper Bound | | | | % Gap with $CDLP$ | | |
|---|---|---|---|---|---|---|---|
| $(N, \alpha)$ | $CDLP$ | $wAR1$ | $wAR2$ | $AF$ | $wAR1$ | $wAR2$ | $AF$ |
| (4, 0.8) | 7,180 | 7,176 | 7,176 | 7,155 | 0.06 | 0.07 | 0.35 |
| (4, 1.0) | 6,462 | 6,377 | 6,391 | 6,352 | 1.31 | 1.09 | 1.70 |
| (4, 1.2) | 6,138 | 6,053 | 6,067 | 6,027 | 1.38 | 1.15 | 1.81 |
| (4, 1.6) | 5,389 | 5,304 | 5,356 | 5,277 | 1.57 | 0.62 | 2.08 |
| (6, 0.8) | 6,918 | 6,891 | 6,889 | 6,860 | 0.39 | 0.42 | 0.84 |
| (6, 1.0) | 6,357 | 6,241 | 6,268 | 6,205 | 1.83 | 1.40 | 2.39 |
| (6, 1.2) | 5,799 | 5,683 | 5,710 | 5,654 | 2.00 | 1.53 | 2.50 |
| (6, 1.6) | 4,796 | 4,704 | 4,789 | 4,672 | 1.91 | 0.13 | 2.57 |
| (8, 0.8) | 6,040 | 5,992 | 6,003 | 5,959 | 0.79 | 0.60 | 1.33 |
| (8, 1.0) | 5,460 | 5,328 | 5,365 | 5,288 | 2.43 | 1.74 | 3.15 |
| (8, 1.2) | 4,993 | 4,857 | 4,899 | 4,817 | 2.73 | 1.90 | 3.52 |
| (8, 1.6) | 4,243 | 4,129 | 4,212 | 4,089 | 2.70 | 0.74 | 3.63 |
| | | | | avg. | 1.59 | 0.95 | 2.16 |

Table 3: Comparison of the upper bounds for the hub and spoke test problems with stationary arrival rates.

| Problem | | Upper Bound | | | % Gap with $CDLP$ | | |
|---|---|---|---|---|---|---|---|
| $(N, \alpha)$ | $CDLP$ | $wAR1$ | $wAR2$ | $AF$ | $wAR1$ | $wAR2$ | $AF$ |
| (4, 0.8) | 4,400 | 4,396 | 4,395 | 4,380 | 0.09 | 0.11 | 0.45 |
| (4, 1.0) | 4,138 | 4,053 | 4,065 | 4,036 | 2.05 | 1.74 | 2.45 |
| (4, 1.2) | 3,796 | 3,711 | 3,725 | 3,689 | 2.23 | 1.85 | 2.81 |
| (4, 1.6) | 3,100 | 3,037 | 3,092 | 3,024 | 2.03 | 0.25 | 2.47 |
| (6, 0.8) | 4,311 | 4,256 | 4,260 | 4,236 | 1.28 | 1.18 | 1.72 |
| (6, 1.0) | 4,015 | 3,900 | 3,924 | 3,868 | 2.87 | 2.26 | 3.68 |
| (6, 1.2) | 3,628 | 3,508 | 3,539 | 3,481 | 3.31 | 2.46 | 4.04 |
| (6, 1.6) | 2,855 | 2,769 | 2,846 | 2,751 | 3.00 | 0.34 | 3.64 |
| (8, 0.8) | 3,802 | 3,678 | 3,704 | 3,650 | 3.25 | 2.59 | 3.99 |
| (8, 1.0) | 3,440 | 3,308 | 3,345 | 3,273 | 3.85 | 2.76 | 4.86 |
| (8, 1.2) | 3,082 | 2,940 | 2,987 | 2,909 | 4.61 | 3.08 | 5.63 |
| (8, 1.6) | 2,475 | 2,364 | 2,460 | 2,341 | 4.46 | 0.59 | 5.41 |
| | | | | avg. | 2.75 | 1.60 | 3.43 |

Table 4: Comparison of the upper bounds for the hub and spoke test problems with non-stationary arrival rates.

| No. of | CPU secs. | | | |
|---|---|---|---|---|
| spokes | $CDLP$ | $wAR1$ | $wAR2$ | $AF$ |
| 6 | 0.4 | 16 | 30 | 98 |
| 8 | 0.8 | 46 | 109 | 405 |
| 10 | 1.2 | 143 | 289 | 1,595 |
| 12 | 1.9 | 415 | 877 | 5,204 |

| No. of time | CPU secs. | | | |
|---|---|---|---|---|
| periods | $CDLP$ | $wAR1$ | $wAR2$ | $AF$ |
| 100 | 0.2 | 6 | 10 | 73 |
| 200 | 0.4 | 16 | 30 | 98 |
| 300 | 0.6 | 41 | 58 | 132 |
| 400 | 0.8 | 115 | 112 | 169 |

Table 5: CPU seconds for $CDLP$, $wAR1$, $wAR2$ and $AF$ as a function of the number of spokes in the airline network and the number of time periods in the booking horizon.

# 8 Contribution

We have proposed tractable approximation methods for the choice network RM problem, when choice is according to the MNL model. $CDLP$ and the affine relaxation are two methods in the literature that give upper bounds on the value function. While $CDLP$ is known to be tractable for the MNL model, we show that the affine relaxation is NP-hard even for the single segment MNL model. Nevertheless, our analysis helps to isolate the term in the affine relaxation which makes it hard to solve. By relaxing this difficult term, we obtain weaker, but tractable approximations. We show that our approximations yield upper bounds that are in between the $CDLP$ and affine bounds. Our relaxations retain the appeal of the formulation discovered in Gallego et al. [5] in that they involve solving compact linear programs, eliminating the need for constraint or column generation. We extend our approximations to the MNL model with multiple segments and disjoint consideration sets. We describe how the formulation can be extended to the case where the consideration sets overlap. Our computational study indicates that our approximations often produce upper bounds that are close to the affine bound and are tractable alternatives to solving the affine relaxation.

# References


[1] Ben-Akiva, M., S. Lerman. 1985. *Discrete-Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge, MA.

[2] Bodea, T., M. Ferguson, L. Garrow. 2009. Choice-based revenue management: Data from a major hotel chain. *Manufacturing and Service Operations Management* **11** 356–361.

[3] Bront, J. J. M., I. Méndez-Díaz, G. Vulcano. 2009. A column generation algorithm for choice-based network revenue management. *Operations Research* **57**(3) 769–784.

[4] Gallego, G., G. Iyengar, R. Phillips, A. Dubey. 2004. Managing flexible products on a network. Tech. Rep. TR-2004-01, Dept of Industrial Engineering, Columbia University, NY, NY.

[5] Gallego, G., R. Ratliff, S. Shebalov. 2010. A general attraction model and an efficient formulation for the network revenue management problem. Tech. rep., Department of IEOR, Columbia University, available at `http://www.columbia.edu/~gmg2/`, retrieved Jan 13, 2012.

[6] Kunnumkal, S., H. Topaloglu. 2010. A new dynamic programming decomposition method for the network revenue management problem with customer choice behavior. *Production and Operations Management* **19**(5) 575–590.

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

[8] Meissner, J., A. K. Strauss. 2012. Network revenue management with inventory-sensitive bid prices and customer choice. *European Journal of Operational Research* **216**(2) 459–468.

[9] Meissner, J., A. K. Strauss, K. T. Talluri. 2011. An enhanced concave programming method for choice network revenue management. Tech. rep., Working Paper 1259, Department of Economics, Universitat Pompeu Fabra, available at `http://www.econ.upf.edu/en/research/onepaper.php?id=1215`.

[10] Peeters, R. 2003. The maximum edge biclique problem is NP-complete. *Discrete Applied Mathematics* **131** 651–654.

[11] Rusmevichientong, P., D. Shmoys, H. Topaloglu. 2010. Assortment optimization with mixtures of logits. Tech. rep., School of IEOR, Cornell University.

[12] Strauss, A.K, K. T. Talluri. 2012. A tractable consideration set structure for network revenue management. Tech. rep., Working Paper 1303, Department of Economics, Universitat Pompeu Fabra, available at `http://www.econ.upf.edu/en/research/onepaper.php?id=1215`.

[13] Talluri, K. T. 2010. A randomized concave programming method for choice network revenue management. Tech. rep., Working Paper 1215, Department of Economics, Universitat Pompeu Fabra, available at `http://www.econ.upf.edu/en/research/onepaper.php?id=1215`.

[14] Talluri, K. T., G. J. van Ryzin. 2004. Revenue management under a general discrete choice model of consumer behavior. *Management Science* **50**(1) 15–33.

[15] Talluri, K. T., G. J. van Ryzin. 2004. *The Theory and Practice of Revenue Management*. Kluwer, New York, NY.

[16] Vossen, T. W. M., D Zhang. 2012. A dynamic disaggregation approach to approximate linear programs for network revenue management. Tech. rep., Leeds School of Business, University of Colorado at Boulder, Boulder, CO.

[17] Zhang, D., D. Adelman. 2009. An approximate dynamic programming approach to network revenue management with customer choice. *Transportation Science* **43**(3) 381–394.

# Appendix: Proofs

## Proof of Proposition 2: NP-completeness of the separation problem for single-segment MNL

We show that solving the separation problem (8) for single-segment MNL is NP-complete. Our reduction is from the following NP-complete biclique problem ([10]). We state first the definitions and notation in the problem.

The problem is defined on an undirected, bipartite graph $G = (V_1 \cup V_2, E)$, with $|V_2| = m_2$. A $(k_1, k_2)$-*biclique* is a complete bipartite subgraph of G, i.e., a subgraph consisting of a pair $(X, Y)$ of vertex subsets $X \subseteq V_1$ and $Y \subseteq V_2$, $|X| = k_1 > 1, |Y| = k_2 > 1$, such that there exists an edge $(x, y) \in E$, $\forall x \in X, y \in Y$. Note that the number of edges in the biclique is $k_1 k_2$.

**Maximum edge biclique problem (MBP)**
*Input:* A bipartite graph $G = (V_1 \cup V_2, E)$ and a positive integer $p$.
*Question:* Does $G$ contain a biclique with at least $p$ edges.

Consider the complementary bipartite graph $\bar{G}$ of $G$ defined on the same vertex set as $G$, where there is an edge $e = (u, v)$ in graph $\bar{G}$ if and only if there is no edge between $u$ and $v$ in $G$.

Define a *cover* $C_S \subseteq V_2$ of a subset $S \subseteq V_1$ in the complement graph $\bar{G}$, as $C_S = \{v \in V_2 | \exists e = (u, v) \in \bar{G}, u \in S\}$. By definition if $C_S$ is a cover of some subset $S$, it means there is no edge from *any* $u \in S$ to *any* $v \in V_2 \backslash C_S$ in the graph $\bar{G}$. Hence, as $G$ is a complement of $\bar{G}$, there is an edge from every $u \in S$ to every $v \in V \backslash C(S)$ in $G$, thus representing a biclique between $S$ and $V \backslash C(S)$ in the graph $G$.

Now we set up the reduction for the separation for (8). In equation (8), for each $u \in V_1$, we associate a product $j$ with $f_j = m_2 \frac{(p+1)}{p}$ and $w_j = m_2$. For each $v \in V_2$, we associate a resource $i$ with weights $\gamma_{it} = \frac{1}{p}$ and $\gamma_{ik} = 0, k > t$. The resource consumptions of the products $j$ are defined from the graph $\bar{G}$: $j$ contains all the $i$ such that there is an edge between the associated nodes in $\bar{G}$. We let $\lambda = 1, \beta_t = m_2$.

We now claim that $G$ has a $(k_1, k_2)$-biclique with $k_1 k_2 \geq p$ if and only if there is a set $S$ that violates the inequality (8) for this instance.

With the above values, $S \subseteq V_1$, with $|S| = k_1, |C(S)| = m_2 - k_2$ violates (8) if and only if

$$m_2 - \frac{\sum_{j \in S} \frac{(p+1)}{p} (m_2)^2}{(1 + \sum_{j \in S} m_2)} < - \sum_{i \in C(S)} \frac{1}{p}$$

or,

$$m_2 - \frac{(p+1) m_2 k_1}{p(\frac{1}{m_2} + k_1)} < - \frac{(m_2 - k_2)}{p}$$

or multiplying both sides by the positive number $p(\frac{1}{m_2} + k_1)$,

$$m_2 p(\frac{1}{m_2} + k_1) - (p+1) m_2 k_1 < -(m_2 - k_2)(\frac{1}{m_2} + k_1)$$

or,

$$p < - \frac{(m_2 - k_2)}{m_2} + k_2 k_1.$$

The term $0 < \frac{(m_2-k_2)}{m_2} < 1$ implies, if and only if

$$p \leq k_2 k_1.$$

□

## Proof of Proposition 5

The right hand side can be written as

$$\gamma_{St}(\sum_{j\in S} w_j) - \sum_{\{i,j|\mathbb{1}_{[S\ni i]}=0,\mathbb{1}_{[S\ni j]}=0\}} \gamma_{it} w_j + \sum_i [1-\mathbb{1}_{[S\ni i]}](\sum_{j\ni i} \gamma_{it} w_j).$$

So it is enough to show that

$$-\sum_{\{i,j|\mathbb{1}_{[S\ni i]}=0,\mathbb{1}_{[S\ni j]}=0\}} \gamma_{it} w_j + \sum_i [1-\mathbb{1}_{[S\ni i]}](\sum_{j\ni i} \gamma_{it} w_j) \leq 0.$$

Examining the term,

$$\begin{aligned}
-\sum_{\{i,j|\mathbb{1}_{[S\ni i]}=0,\mathbb{1}_{[S\ni j]}=0\}} \gamma_{it} w_j &+ \sum_i [1-\mathbb{1}_{[S\ni i]}](\sum_{j\ni i} \gamma_{it} w_j) \\
&= -\sum_{\{i,j|\mathbb{1}_{[S\ni i]}=0,\mathbb{1}_{[S\ni j]}=0\}} \gamma_{it} w_j + \sum_{i\notin S} \gamma_{it}(\sum_{j\ni i} w_j) \\
&= -\sum_{\{i,j|S\not\ni i,S\not\ni j,\}} \gamma_{it} w_j + \sum_{S\not\ni i} \gamma_{it}(\sum_{j\ni i} w_j) \\
&= -\sum_{\{i,j|S\not\ni i,S\not\ni j,j\not\ni i\}} \gamma_{it} w_j - \sum_{S\not\ni i} \gamma_{it}(\sum_{j\ni i} w_j) + \sum_{S\not\ni i} \gamma_{it}(\sum_{j\ni i} w_j) \leq 0,
\end{aligned}$$

where the last equality follows from the fact that for all $i \in \{i|S \not\ni i\}$ we cannot have a $j \in S$ with $i \in j$. □

## Proof of Proposition 6

We first make the following two elementary observations: (i) If $\Pi_1 = \min_x f(x)$ s.t $g(x) \geq 0$ and $\Pi_2 = \min f(x)$ s.t $h(x) \geq 0$ and if $h(x) \leq g(x)$ for all $x$, then $\Pi_1 \leq \Pi_2$. (ii) Similarly if $\Pi_1' = \min_x f(x)$ s.t $g_1(x) + g_2(x) \geq 0$ and $\Pi_2' = \min f(x)$ s.t $g_1(x) \geq 0, g_2(x) \geq 0$, then $\Pi_1' \leq \Pi_2'$.

Consider (20) and do the same relaxation as we did in (12): Replace

$$\sum_{i\in\mathcal{I}_l} \gamma_{S_l,t} \frac{\lambda_l}{\sum_{l'\in\mathcal{L}_i} \lambda_{l'}} \sum_{j\in S_l} P_j^l(S_l) = \sum_{i\in\mathcal{I}_l} \mathbb{1}_{[S_l\ni i]} \frac{\lambda_l \gamma_{it}}{\sum_{l'\in\mathcal{L}_i} \lambda_{l'}} \sum_{j\in S_l} P_j^l(S_l)$$

by a smaller quantity $\sum_{j\in S_l} P_j^l(S_l) \sum_{i\in j} \frac{\lambda_l \gamma_{it}}{\sum_{l'\in\mathcal{L}_i} \lambda_{l'}}$. We then use observation (i) above and use observation (ii) in summing over the $l$ to show we get a looser bound than $V^{AF}$. We fill in the details below.

The separation problem $S1_{l,t}$ that gives $V^{swAR1}$, in terms of $P_0^l(S_l) = \frac{1}{1+\sum_{k\in S_l} w_k^l}$ and $P_j^l(S_l) = \frac{w_j^l}{1+\sum_{k\in S_l} w_k^l}$ is

$$-\beta_{lt} + \sum_{j\in S_l} \lambda_l \Big[f_j - \sum_{i\in j}\big(\sum_{k=t+1}^{\tau} \gamma_{ik} + \frac{\gamma_{it}}{\sum_{l'\in\mathcal{L}_i} \lambda_{l'}}\big)\Big] P_j^l(S_l) - \Big[\sum_{i\in\mathcal{I}_l} \mathbb{1}_{[S_l\ni i]} \frac{\lambda_l \gamma_{it}}{\sum_{l'\in\mathcal{L}_i} \lambda_{l'}}\Big] P_0^l(S_l) \le 0. \quad (25)$$

Consider now two intermediate problems:

$$\begin{aligned} \underline{V} = \quad & \min_{\beta,\gamma} \sum_l \sum_t \beta_{lt} + \sum_i \sum_t r_i^1 \gamma_{it} \\ \text{s.t} \quad & \beta_{lt} \ge \lambda_l \big[R^l(S_l) - \sum_{i\in\mathcal{I}_l} \sum_{k=t+1}^{\tau} Q_i^l(S_l)\gamma_{ik}\big] - \sum_{i\in\mathcal{I}_l} \mathbb{1}_{[S_l\ni i]} \gamma_{it} \frac{\lambda_l}{\sum_{l'\in\mathcal{L}_i} \lambda_l'} \quad \forall t, l, S_l \subset \mathcal{C}_l \quad (26) \\ & \gamma_{it} \ge 0, \end{aligned}$$

and

$$\begin{aligned} \bar{V} = \quad & \min_{\beta,\gamma} \sum_l \sum_t \beta_{lt} + \sum_i \sum_t r_i^1 \gamma_{it} \\ \text{s.t} \quad & \beta_{lt} \ge \lambda_l \big[R^l(S_l) - \sum_{i\in\mathcal{I}_l} \sum_{k=t}^{\tau} Q_i^l(S_l)\gamma_{ik}\big] \quad \forall t, l, S_l \subset \mathcal{C}_l \quad (27) \\ & \gamma_{it} \ge 0. \end{aligned}$$

We can interpret the first problem as a segment based relaxation of $AF$, while the second problem can be viewed as a segment based relaxation of $CDLP$. Lemmas 1, 2 and 3 below show that $V^{AF} \le \underline{V} \le V^{swAR1} \le \bar{V} = V^{CDLP}$ which proves the result.

**Lemma 1.** $\underline{V} \le V^{swAR1} \le \bar{V}$.

<u>Proof</u>
Since the objective functions of all the problems are the same, we only need to compare the corresponding constraints. Since $\sum_{j\in i} P_j^l(S_l) + P_0^l(S_l) = 1$, it follows that constraint (25) implies constraint (26) and we have $\underline{V} \le V^{swAR1}$.

On the other hand, the right hand side of constraint (27) can be written as

$$\lambda_l \big[R^l(S_l) - \sum_{i\in\mathcal{I}_l} \sum_{k=t+1}^{\tau} Q_i^l(S_l)\gamma_{ik}\big] - \sum_{i\in\mathcal{I}_l} \lambda_l Q_i^l(S_l)\gamma_{it}.$$

Now note that

$$\begin{aligned} \lambda_l Q_i^l(S_l)\gamma_{it} &= \lambda_l \mathbb{1}_{[S_l\ni i]} Q_i^l(S_l)\gamma_{it} = \lambda_l \mathbb{1}_{[S_l\ni i]} \big[\sum_{j\in i} P_j^l(S_l)\big]\gamma_{it} \\ &\le \frac{\lambda_l}{\sum_{l'\in\mathcal{L}_i} \lambda_{l'}} \mathbb{1}_{[S_l\ni i]} \big[\sum_{j\in i} P_j^l(S_l)\big]\gamma_{it} \le \frac{\lambda_l}{\sum_{l'\in\mathcal{L}_i} \lambda_{l'}} \mathbb{1}_{[S_l\ni i]} \big[\sum_{j\in i} P_j^l(S_l) + P_0^l(S_l)\big]\gamma_{it} \end{aligned}$$

where the first equality holds since if $\mathbb{1}_{[S_l\ni i]} = 0$, then $Q_i^l(S_l) = 0$ and the first inequality holds since $\sum_{l'\in\mathcal{L}_i} \lambda_{l'} \le 1$. Therefore constraint (27) implies constraint (25) and we have $V^{swAR1} \le \bar{V}$.

□

**Lemma 2.** $V^{AF} \leq \underline{V}$.

Proof
Suppose that $(\hat{\beta}_{lt}, \hat{\gamma}_{it})$ satisfies constraint (26). Let $S = \cup_l S_l$. We have

$$\begin{aligned}
\sum_l \hat{\beta}_{lt} &\geq \sum_l \Big\{ \lambda_l \big[ R^l(S_l) - \sum_{i \in \mathcal{I}_l} \sum_{k=t+1}^{\tau} Q_i^l(S_l) \hat{\gamma}_{ik} \big] - \sum_{i \in \mathcal{I}_l} \mathbb{1}_{[S_l \in i]} \hat{\gamma}_{it} \frac{\lambda_l}{\sum_{l' \in \mathcal{L}_i} \lambda_l'} \Big\} \\
&= \lambda \big[ R(S) - \sum_i \sum_{k=t+1}^{\tau} Q_i(S) \hat{\gamma}_{ik} \big] - \sum_i \hat{\gamma}_{it} \sum_{l \in \mathcal{L}_i} \mathbb{1}_{[S_l \in i]} \frac{\lambda_l}{\sum_{l' \in \mathcal{L}_i} \lambda_l'} \\
&\geq \lambda \big[ R(S) - \sum_i \sum_{k=t+1}^{\tau} Q_i(S) \hat{\gamma}_{ik} \big] - \sum_i \hat{\gamma}_{it} \sum_{l \in \mathcal{L}_i} \mathbb{1}_{[S \in i]} \frac{\lambda_l}{\sum_{l' \in \mathcal{L}_i} \lambda_l'} \\
&= \lambda \big[ R(S) - \sum_i \sum_{k=t+1}^{\tau} Q_i(S) \hat{\gamma}_{ik} \big] - \sum_i \hat{\gamma}_{it} \mathbb{1}_{[S \in i]},
\end{aligned}$$

where the first equality uses the fact that $Q_i^l(S_l) = 0$ for $l \notin \mathcal{L}_i$ and hence $\lambda Q_i(S) = \sum_l \lambda_l Q_i^l(S_l) = \sum_{l \in \mathcal{L}_i} \lambda_l Q_i^l(S_l)$. The inequality holds since $\mathbb{1}_{[S_l \ni i]} \leq \mathbb{1}_{[S \ni i]}$. Hence $(\sum_l \hat{\beta}_{lt}, \hat{\gamma}_{it})$ is feasible for the affine LP and the proof follows.

□

Meissner et al. [9] prove the following that we include for completeness.

**Lemma 3.** *([9])* $\bar{V} = V^{CDLP}$.

Proof
Noting that the constraint in $CDLP$ is equivalent to

$$\begin{aligned}
\beta_t = \max_S \lambda \big[ R(S) - \sum_i \sum_{k=t}^{\tau} Q_i(S) \gamma_{ik} \big] &= \max_S \sum_l \lambda_l [R^l(S \cap \mathcal{C}_l) - \sum_{i \in \mathcal{I}_l} \sum_{k=t}^{\tau} Q_i^l(S \cap \mathcal{C}_l) \gamma_{ik}] \\
&= \sum_l \max_{S_l} \lambda_l [R^l(S_l) - \sum_{i \in \mathcal{I}_l} \sum_{k=t}^{\tau} Q_i^l(S_l) \gamma_{ik}]
\end{aligned}$$

where the last inequality uses the fact that the consideration sets are disjoint. Therefore, the $CDLP$ constraint is equivalent to the constraints $\beta_t = \sum_l \beta_{lt}$ and $\beta_{lt} = \max_{S_l} \lambda_l [R^l(S_l) - \sum_{i \in \mathcal{I}_l} \sum_{k=t}^{\tau} Q_i^l(S_l) \gamma_{ik}\big]$, which is exactly constraint (27).

□