New compact linear programming formulations for choice network revenue management

Sumit Kunnumkal*       Kalyan Talluri†

November 21, 2014

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 dynamic controls. Recently a compact linear programming formulation of CDLP for the multinomial-logit (MNL) model of customer choice with non-overlapping consideration sets has been proposed. Our objective is to obtain a tighter bound than CDLP while retaining the appealing properties of a 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 new linear programs that approximate 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. We give extensions to the case with multiple customer segments and nested-logit model of choice. 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 heterogeneous 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. 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.

In industries such as hotels, advertising and airlines, the products consume bundles of different resources (multi-night stays, bundles of ad slots, 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

---

*Indian School of Business, Hyderabad, 500032, India, email: sumit.kunnumkal@ish.edu
†ICREA and Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain, email: kalyan.talluri@upf.edu
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 [19] 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 [18] for the single-resource problem. Bodea, Ferguson, and Garrow [3] for instance use choice data from a large hotel chain and empirically study the suitability of choice models. Vulcano, van Ryzin, and Chaar [22] and Newman, Ferguson, Garrow, and Jacobs [12] and Talluri [16] study estimation of choice models from sales data.

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 [6] and Liu and van Ryzin [9], who formulate the choice deterministic linear program (CDLP). They show that 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 multinomial-logit (MNL) model of choice when the consideration sets of the different customer segments are disjoint ([9]). However, generating the columns is difficult (NP-complete) when the segment consideration sets overlap under the MNL model with just two segments ([4], [14]).

Given the difficulty of solving CDLP, Talluri [17] 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 [11]. Strauss and Talluri [15] show that SDCP with the product-cuts added is equivalent to CDLP when the consideration set intersections have a tree structure.

Kunnumkal and Topaloglu [8] and Zhang and Adelman [23] study decomposition procedures and an affine relaxation of the dynamic program. In the same vein, Meissner and Strauss [10] 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 [7] 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 new linear-programming formulations that yields an upper bound on the dynamic programming value function and are provably tighter than the CDLP bound. Moreover, for the MNL model, our formulations are compact and similar to the one discovered in [7]. Although theoretically weaker than the affine relaxation, we find in our numerical study, that our relaxations are often close to the affine relaxation upper bound.

To summarize, our contributions in this paper are as follows:

1. We show that the affine relaxation is NP-hard even for a single-segment MNL. Previously there
was some hope that CDLP could be improved at least for this simple and widely used choice model.

2. We propose new relaxations based on the affine relaxation. This shows that there is some value and insight possible from formulating and analyzing dynamic programming relaxations—even if the actual relaxation is difficult we obtain new insights for tractable relaxations. For the MNL model, not only are our relaxations provably tighter than CDLP and tractable, but they also can be formulated as compact linear programs. Compact representations are appealing from an implementation perspective since it eliminates the need for customized coding in the form of constraint or column generation techniques.

3. We give extensions to multiple segments, nested logit, study alternate relaxations theoretically and numerically and analyze the applicability of the product cuts developed in Meissner et al. [11].

4. We develop policies based on the relaxations and perform numerical experiments to study the effectiveness of the new formulations.

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. Next, in §4 we show that the affine relaxation is NP-hard even for single segment MNL and give various relaxations which are tractable and fall in between CDLP and the affine relaxation. §5 discusses extensions to multiple segments and nested logit. §6 contains our computational study using the new formulations.

2 Model and notation

Our model and notation is to a large part based on Liu and van Ryzin [9]. 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 (usually indexed by \( i \)) and \( n \) products (usually 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 \( I_j \subseteq I \), and its sale brings in revenue \( f_j \). A resource \( i \) is used by a subset \( J_i \subseteq J \) of products. A resource \( i \) is said to be in product \( j \) (\( i \in I_j \)) if \( j \) uses resource \( i \), and conversely we write \( j \in J_i \).

The resources used by product \( j \) are represented by the 0-1 incidence vector \( \mathbb{1}_{I_j} \), which has a 1 in the \( i \)th position if \( i \in I_j \) and a 0 otherwise.

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

We let \( r_1 = [r^1_i] \) represent the initial capacity on the resources and \( r^t = [r^t_i] \) denote the remaining capacity on resource \( i \) at beginning of time period \( t \). The remaining capacity \( r^t_i \) takes values in the
set $R_i = \{0, \ldots, r_i^j\}$ and $R = \prod_i R_i$ represents the state space.

2.1 Demand model

We assume there are $L = \{1, \ldots, L\}$ customer segments, each with distinct purchase behavior. In each period a customer from segment $l$ arrives with probability $\lambda_l$ so that $\lambda = \sum_l \lambda_l$ is the total arrival rate. Note that conditioned on a customer arrival, $\lambda_l/\lambda$ is the probability that the customer belongs to segment $l$.

Each segment $l$ has a consideration set $C_l \subseteq J$ 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). The choice probabilities of a segment-$l$ customer are not affected by products not in its 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 $S_l = C_l \cap S$ note that $P^j_l(S) = P^j_l(S_l)$.

Given a customer arrival, and an offer set $S$, the probability that the firm sells $j \in S$ is then given by $P_j(S) = \sum_l \frac{\lambda_l}{\lambda} P^j_l(S_l)$ and makes no sale with probability $P_0(S) = 1 - \sum_j P_j(S)$. The expected sales for product $j$ is therefore $\lambda P_j(S) = \sum_l \lambda_l P^j_l(S_l)$, while $1 - \lambda + \lambda P_0(S) = 1 - \sum_j \lambda P_j(S)$ is the probability of no sales in a time period. Given an offer set $S$, $Q^j_l(S) = \sum_{j \in J} P^j_l(S)$ denotes the expected capacity consumed on resource $i$ conditional on a segment-$l$ customer arrival and $Q_l(S) = \sum_j \frac{\lambda_j}{\lambda} Q^j_l(S)$ denotes the expected capacity consumed on resource $i$ conditional on a customer arrival. Note that $\lambda Q_j(S) = \sum_l \lambda_l Q^j_l(S)$ gives the expected capacity consumed on resource $i$ in a time period. The revenue functions can be written as $R^i(S) = \sum_{j \in S_l} f_j P^j_l(S_l)$ and $R(S) = \sum_{j \in S} f_j P_j(S)$.

We write $i \in I_S$ whenever there is a $j \in S$ with $i \in I_j$. We also assume that the arrival rates and choice probabilities are stationary. This is again for brevity of notation and all of our results go through with nonstationary arrival rates and choice probabilities.

2.2 MNL model

Some of the results of this paper apply to general discrete-choice models, but the main new relaxations apply to the MNL model of choice.\(^1\)

In the MNL model of choice probabilities, when a subset $S_l \subseteq C_l$ of products are offered by the firm, a customer in segment $l$ chooses product $j \in S_l$ with probability

$$P^j_l(S_l) = \frac{w_{lj}}{1 + \sum_{k \in S_l} w_{lk}},$$

where $w_{lj}$ is a weight associated with product $j$. The no-purchase option is indexed 0, and we

\(^1\)Our results in fact extend unchanged to the slightly more general attraction model of [7].
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 $P_0(S_l) = \frac{1}{1 + \sum_{k \in S_l} w_k}$.

The weight $w_l$ 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 [2] for background on this popular model. The firm does not observe segment membership.

The case of non-overlapping consideration sets for the segments is one of the few known cases where the CDLP formulation is solvable. This model is often taken as an approximation (for tractability). Under this assumption, the $L$ segments have considerations sets that do not overlap ($C_l \cap C_{l'} = \emptyset; l \neq l'$).

\subsection{Choice dynamic program}

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

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

where

$$S(r) = \{ j : \| j \| \leq r; \forall i \}$$

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

\subsection{Linear programming formulation of the dynamic program}

The value functions can, alternatively, be obtained by solving a linear program; see Zhang and Adelman [23]. The linear programming formulation of the network choice RM DP given below, has a decision variable for each state vector in each period $V_t(r)$ and is as follows:

$$V^{DPLP} = \min_{V_t(r)} V_t(r^1) \quad \text{s.t.} \quad (DPLP) \quad V_t(r) \geq \sum_j \lambda P_j(S) \left[ f_j + V_{t+1} \left( r - 1_{[j \in S]} \right) - V_{t+1} \left( r \right) \right] + V_{t+1} \left( r \right) \quad \forall r \in \mathcal{R}, S \subseteq S(r), t,$$

with the boundary condition that $V_{t+1}(\cdot) = 0$. Both the dynamic program (1) and linear program DPLP are computationally intractable, but the linear program DPLP 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)

The choice deterministic linear program CDLP proposed in Gallego et al. [6] and Liu and van Ryzin [9] is a certainty-equivalence approximation to (1). We write CDLP in an expanded, redundant way as the following linear program (LP):

\[
V_{\text{CDLP}} = \max \sum_t \sum_S \lambda R(S) h_{S,t} \\
\text{s.t. } \sum_{k=1}^t \sum_S \lambda Q_i(S) h_{S,k} \leq r_i^1 \quad \forall i, t \\
(CDLP) \quad \sum_S h_{S,t} = 1 \quad \forall t \\
\quad h_{S,t} \geq 0.
\]

The decision variable \( h_{S,t} \) can be interpreted as the frequency with which set \( S \) (including the empty set) is offered at time period \( t \). The first set of constraints ensure that the total expected capacity consumed on resource \( i \) up until time period \( t \) does not exceed the available capacity. Note that since \( h_{S,t} \geq 0 \), constraints (3) are redundant except for the last time period. Still, this expanded formulation is useful when we compare CDLP with other approximation methods. The second set of constraints states that the sum of the frequencies adds up to 1.

Associating dual variables \( \gamma = \{ \gamma_{i,t} | \forall i, t \} \) with constraints (3) and \( \beta = \{ \beta_t | \forall t \} \) with constraints (4), the dual of CDLP is

\[
V_{\text{dCDLP}} = \min_{\beta, \gamma} \sum_t \beta_t + \sum_t \sum_i \gamma_{i,t} r_i^1 \\
\text{s.t. } \quad (dCDLP) \quad \beta_t + \sum_i \left( \sum_{k=t}^\tau \gamma_{i,k} \right) \lambda Q_i(S) \geq \lambda R(S) \quad \forall t, S \\
\quad \gamma_{i,t} \geq 0.
\]

Liu and van Ryzin [9] show that the optimal objective function value of CDLP, \( V_{\text{CDLP}} \), is an upper bound on \( V_{\text{DPLP}} \). Besides giving an upper bound on the value function, CDLP can also be used to construct different heuristic control policies. One idea that is pursued in [23] is to use the optimal values of the dual variables associated with constraints (3) to come up with a value function approximation, and use this approximation in place of the value function in optimality equation (1) to decide on the offer set. Letting \( \tilde{\gamma} = \{ \tilde{\gamma}_{i,t} | \forall i, t \} \) denote the optimal values of the dual variables associated with constraints (3), we interpret \( \tilde{\gamma}_{i,t} \) as giving the value of an additional unit of capacity on resource \( i \) from time period \( t \) to \( t + 1 \). With this interpretation, \( \sum_{k=t}^\tau \tilde{\gamma}_{i,k} \) gives the marginal value of capacity on resource \( i \) at time period \( t \) and we approximate the value function as

\[
\hat{v}_t(r) = \sum_i \left( \sum_{k=t}^\tau \tilde{\gamma}_{i,k} \right) r_i.
\]
If $r^t$ is the vector of remaining resource capacities at time $t$, we solve the problem

$$
\max_{S \subseteq S(r^t)} \left\{ \sum_{j \in S} \lambda P_j(S) \left[ f_j + \hat{V}_{t+1}(r - 1 \mathbb{1}_{\mathcal{Z}_j}) \right] + [\lambda P_0(S) + 1 - \lambda] \hat{V}_{t+1}(r^t) \right\},
$$

(7)

and offer the set which achieves the maximum in the above optimization problem.

Since CDLP has $O(2^n)$ decision variables, it has to be solved using column generation. Liu and van Ryzin [9] 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. [4] and Rusmevichientong et al. [14] 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.

3.2 Affine relaxation

The second approximation method we consider is the affine relaxation, where the value function is approximated as $V_t(r) = \theta_t + \sum_i V_i(t, r_i)$. Note that $V_i(t, r)$ can be interpreted as the marginal value of capacity on resource $i$ at time $t$. Substituting this value function approximation into the formulation DPLP we get the affine relaxation LP

$$
V^{AF} = \min_{\theta, V} \theta + \sum_i V_i 1_r^i
$$

s.t.

$$(AF) \quad \theta + \sum_i V_i(t, r_i) \geq \sum_j \lambda P_j(S) \left[ f_j - \sum_{i \in \mathcal{I}_j} V_i(t+1) \right] + \theta_{t+1} + \sum_i V_{i, t+1}(t, r_i)$$

with the boundary conditions $\theta_{t+1} = 0, V_{i, t+1} = 0$. Zhang and Adelman [23] show that the optimal objective function value $V^{AF}$ is an upper bound on $V^{DPLP}$ and that there exists an optimal solution $(\hat{\theta}, \hat{V})$ of $AF$ that satisfies $\hat{V}_{t, t} - \hat{V}_{t+1, t} \geq 0$ 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 [21] 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 making a change of variables. Letting $\beta_t = \theta_t - \theta_{t+1}$, and $\gamma_{i, t} = V_i(t, r) - V_{i, t+1}(t, r)$, we can write $AF$ equivalently as

$$
\min_{\beta, \gamma} \sum_t \beta_t + \sum_t \sum_i \gamma_{i, t} 1_r^i
$$

s.t. $\beta_t + \sum_i \gamma_{i, t} + \sum_j \lambda P_j(S) \left[ \sum_{k=t+1}^{t} \sum_{i \in \mathcal{I}_j} \gamma_{i, k} \right] - f_j \geq 0 \quad \forall r \in \mathcal{R}, S \subseteq S(r), t$

(8)

where we use the fact that $V_i(t, r) = \sum_{k=t}^{t} \gamma_{i, k}$ and so $\sum_{k=t}^{t} \gamma_{i, k}$ can be interpreted as the marginal value of capacity on resource $i$ at time $t$. Note that the nonnegativity constraint on $\gamma_{i, t}$ is without
loss of generality, since there exists an optimal solution to $AF$ that satisfies $V_{i,t} - V_{i,t+1} \geq 0$. Now, constraints (8) can be written as

$$
\min_{r \in \mathcal{R}, S \subseteq S(r)} \left\{ \beta_t + \sum_i \gamma_{i,t} r_i + \sum_j \lambda P_j(S) \left[ \left( \sum_{i \in \mathcal{I}_j} \sum_{k=t+1}^{\tau} \gamma_{i,k} \right) - f_j \right] \right\} \geq 0 \tag{9}
$$

for all $t$. Since $\gamma_{i,t} \geq 0$, the coefficient of $r_i$ in the minimization problem (9) is nonnegative, and we can assume $r_i \in \{0,1\}$ in the minimization (as larger values of $r_i$ would be redundant in $S \subseteq S(r)$ and would only increase the objective value). Moreover, since $\gamma_{i,t} \geq 0$, for any set $S$, we have $r_i = 0$ for $i \not\in \mathcal{I}_S$. On the other hand, feasibility requires we have $r_i = 1$ for $i \in \mathcal{I}_S$. Therefore, (9) can be written as

$$
\min_S \left\{ \beta_t + \sum_i \mathbb{I}_{[i \in \mathcal{I}_S]} \gamma_{i,t} + \sum_j \lambda P_j(S) \left[ \left( \sum_{i \in \mathcal{I}_j} \sum_{k=t+1}^{\tau} \gamma_{i,k} \right) - f_j \right] \right\} \geq 0. \tag{10}
$$

and we can write $AF$ equivalently as

$$
V^{RAF} = \min_{\beta, \gamma} \sum_t \beta_t + \sum_t \sum_i \gamma_{i,t} r_i^1 \tag{11}
$$

(RAF) s.t. $\beta_t + \sum_i \mathbb{I}_{[i \in \mathcal{I}_S]} \gamma_{i,t} + \sum_i \left( \sum_{k=t+1}^{\tau} \gamma_{i,k} \right) \lambda Q_i(S) \geq \lambda R(S) \ \forall t, S$

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 $h_{S,t}$ with constraints (11), we get

$$
V^{dRAF} = \max_h \sum_t \sum_S \lambda R(S) h_{S,t} \tag{12}
$$

s.t. $\sum_S \sum_{k=1}^{t-1} \lambda Q_i(S) h_{S,k} + \mathbb{I}_{[i \in \mathcal{I}_S]} h_{S,t} \leq r_i^1 \ \forall i, t$ (dRAF) $\sum_S h_{S,t} = 1 \ \forall t$

$h_{S,t} \geq 0$.

The above arguments imply that

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

In addition to giving an upper bound on the optimal expected total revenue, the affine relaxation can also be used to construct heuristic control policies. Letting $(\hat{\beta}, \hat{\gamma})$, with $\hat{\beta} = \{\hat{\beta}_t | \forall t\}$ and $\hat{\gamma} = \{\hat{\gamma}_{i,t} | \forall i, t\}$, denote an optimal solution to $RAF$, we use $\sum_{k=t}^{\tau} \hat{\gamma}_{i,k}$ to approximate the marginal value of capacity on resource $i$ at time $t$. We approximate $V_t(r)$ using (6) and solve problem (7) using this value function approximation to decide on the set of products to be offered at time period $t$.

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 constraint (8).
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. After some initial preliminary results, we show that the affine relaxation is NP-complete even for the MNL model with a single segment; in contrast CDLP is tractable. We then give a series of tractable relaxations that fall in between CDLP and the affine relaxation. In §5 we extend the formulations to generalizations of the single-segment MNL model.

4.1 Preliminaries

In this section we state some known facts and some preliminary results.

Since we restrict attention to the single-segment MNL model, we drop the segment superscript \( l \) and write the weights as \( w_j \), and the choice probabilities, expected resource consumptions and expected revenues as:

\[
P_j(S) = \frac{1}{1 + \sum_{j' \in S} w_{j'}} \quad Q_i(S) = \frac{\sum_{j \in S \setminus \{i\}} w_j}{1 + \sum_{j \in S} w_j} \quad R(S) = \frac{\sum_{j \in S} f_j w_j}{1 + \sum_{j \in S} w_j}.
\]

4.1.1 Compact formulation of CDLP

Gallego et al. [7] give the following equivalent formulation of CDLP for the MNL choice model:

\[
V^{SBLP} = \max \sum_t \sum_j \lambda f_j x_{j,t} \quad \text{s.t.} \quad \sum_t \sum_j \lambda x_{j,t} \leq r_i^1 \quad \forall t \\
(SBLP) \quad x_{0,t} + \sum_j x_{j,t} = 1 \quad \forall t \\
\frac{x_{j,t}}{w_j} - x_{0,t} \leq 0 \quad \forall j, \forall t \\
x_{j,t} \geq 0.
\]

In the above LP, the decision variables \( x_{j,t} \) can be viewed as the rate of sales of product \( j \) at time period \( t \). The first constraint ensures that the total capacity consumed by the products on each resource does not exceed the available capacity. The second and third constraints ensure that the sales rates are consistent with the MNL choice probabilities. This formulation, referred to as the sales-based linear program (SBLP), vastly reduces the complexity of solving CDLP, albeit restricted to the MNL model.

4.1.2 Affine relaxation for single-segment MNL

Substituting the MNL choice probabilities, expected resource consumptions and expected revenues into constraint (11), we obtain

\[
\beta_t + \gamma_{S,t} + \sum_{k=t+1}^{T} \left( \sum_{k=t+1}^{T} \gamma_{k,h} \right) \frac{\sum_{j \in S \setminus \{i\}} w_j}{1 + \sum_{j \in S} w_j} \geq \lambda \frac{\sum_{j \in S} f_j w_j}{1 + \sum_{j \in S} w_j}
\]

9
where

$$\gamma_{s,t} = \sum_i \mathbf{1}_{[i \in I_s]} \gamma_{i,t}.$$  

Multiplying both sides by the positive quantity $$1 + \sum_{j \in S} w_j$$ and simplifying, constraint (11) of RAF can be equivalently written as

$$\beta_t \geq -\gamma_{s,t} \left( 1 + \sum_{j \in S} w_j \right) - \sum_{j \in S} w_j \left[ \beta_t + \lambda \left( \sum_{i \in I_j} \sum_{k=t+1}^\tau \gamma_{i,k} \right) - f_j \right]. \quad (14)$$

Since the constraint has to be satisfied for every $$S$$ and $$t$$, we have $$\beta_t \geq \Pi_t^{AF}(\beta, \gamma)$$ for all $$t$$, where

$$\Pi_t^{AF}(\beta, \gamma) = \max_S \left\{ -\gamma_{s,t} \left( 1 + \sum_{j \in S} w_j \right) - \sum_{j \in S} w_j \left[ \beta_t + \lambda \left( \sum_{i \in I_j} \sum_{k=t+1}^\tau \gamma_{i,k} \right) - f_j \right] \right\} \quad (15)$$

and the affine relaxation constraint (11) can be written as

$$\beta_t \geq \Pi_t^{AF}(\beta, \gamma) \quad \forall t. \quad (16)$$

4.2 NP-completeness of the affine separation for single-segment MNL

Since RAF has an exponential number of constraints, we have to generate constraints (14) on the fly. Given a set of values $$(\beta, \gamma)$$, the separation problem at time $$t$$ is to decide if constraint (14) is satisfied for all $$S$$, and if not, add the violated constraint to the LP. In other words, the separation problem at time $$t$$ involves solving optimization problem (15) and determining if $$\beta_t \geq \Pi_t^{AF}(\beta, \gamma)$$. If $$\beta_t \geq \Pi_t^{AF}(\beta, \gamma)$$, then constraint (14) is satisfied for all $$S$$ at time $$t$$. Otherwise, the set $$\hat{S}$$ which attains the maximum in problem (15) violates the constraint, and we add the constraint for set $$\hat{S}$$ to the LP. Therefore, solving problem (15) in an efficient manner is key to separating constraints (14) efficiently. Proposition 2 below states that the affine relaxation separation problem for MNL with a single segment, as given in (14) is NP-complete.

**Proposition 2.** The following problem is NP-complete:

Input: $$w_j \geq 0$$, $$1 \geq \lambda \geq 0$$, $$f_j \geq 0$$, and values $$\beta_t$$ and $$\gamma_{i,t} \geq 0$$.

Question: Is there a set $$S$$ that violates (14)?

**Proof.** Our reduction is from the NP-complete maximum edge biclique problem ([13]). 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 complement 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 $$G$$. 

10
Define a cover $C_S \subseteq V_2$ of a subset $S \subseteq V_1$ in the complement graph $\bar{G}$, as $C_S = \{v \in V_2 \mid \exists e = (u, v) \in \bar{G}, u \in S\}$. 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 $v \in V_2 \setminus C_S$ 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 $v \in V \setminus C(S)$ in $G$, thus representing a biclique between $S$ and $V \setminus C(S)$ in the graph $G$.

Now we set up the reduction for the separation for (14). In equation (14), 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 $\gamma_{i,t} = \frac{1}{p}$ and $\gamma_{i,k} = 0, k > t$. 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 $k_1 k_2 > p$ if and only if there is a set $S$ that violates the inequality (14) for this instance.

With the above values, $S \subseteq V_1$, with $|S| = k_1, |C(S)| = m_2 - k_2$ violates (14) 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,

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

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

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

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

$$p < k_2 k_1.$$  

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 the affine relaxation as we do next.

### 4.2.1 Comparing relaxations

All of our approximation methods involve solving an optimization problem of the form

$$\min_{\beta, \gamma} \sum_t \beta_t + \sum_t \sum_i \gamma_{t,i} r_{i,t}^1$$

11
subject to the constraints $\beta_t \geq \Pi_t^I(\beta, \gamma)$, where $\Pi_t^I(\cdot, \cdot)$ is a scalar function of $\beta = \{\beta_t | \forall t\}$ and $\gamma = \{\gamma_{i,t} | \forall i, t\}$. The following result is useful in comparing the upper bounds obtained by the different approximation methods.

**Lemma 1.** Let

$$V^I = \min_{\beta, \gamma} \sum_t \beta_t + \sum_i \gamma_{i,t} t_i$$

(I) s.t. $\beta_t \geq \Pi_t^I(\beta, \gamma) \; \forall t$

$\gamma_{i,t} \geq 0$, and let

$$V^{II} = \min_{\beta, \gamma} \sum_t \beta_t + \sum_i \gamma_{i,t} t_i$$

(II) s.t. $\beta_t \geq \Pi_t^{II}(\beta, \gamma) \; \forall t$

$\gamma_{i,t} \geq 0$.

If $\Pi_t^I(\beta, \gamma) \leq \Pi_t^{II}(\beta, \gamma)$ for all $t$, then $V^I \leq V^{II}$.

**Proof.** The proof follows by noting that a feasible solution to problem (II) is also feasible to problem (I) and both optimization problems have the same objective function.

## 4.2.2 CDLP vs. AF for single-segment MNL

The CDLP vs. AF comparison for single-segment MNL brings out some intuition into the difficulty of the affine relaxation. We therefore derive the comparison specializing to this model.

Using the single-segment MNL formulas for the expected resource consumptions and expected revenues, the CDLP dual constraint (5) can be written as

$$\beta_t \geq -\sum_{j \in S} w_j \left[ \beta_t + \lambda \left( \sum_{i \in I_j} \sum_{k=t}^{\tau} \gamma_{i,k} - f_j \right) \right] \; \forall t, S$$

which is almost identical to the right-hand-side of (14) except that the summation goes from $k = t$.

To make the comparison easier, we rewrite the above constraint as

$$\beta_t \geq \Pi_t^{CDLP}(\beta, \gamma) \; \forall t \tag{17}$$

where

$$\Pi_t^{CDLP}(\beta, \gamma) = \max_S \left\{ -\sum_{j \in S} w_j \lambda \sum_{i \in I_j} \gamma_{i,t} - \sum_{j \in S} w_j \left[ \beta_t + \lambda \left( \sum_{i \in I_j} \sum_{k=t}^{\tau} \gamma_{i,k} - f_j \right) \right] \right\}. \tag{18}$$

Since $0 \leq \lambda \leq 1$, and $\gamma_{S,t} = \sum_{i \in I_j} \Pi_{i \in I_j} | \gamma_{i,t} \geq \sum_{i \in I_j} \gamma_{i,t} \geq 0$ for all $j \in S$, we have

$$\gamma_{S,t} \left( 1 + \sum_{j \in S} w_j \right) \geq \lambda \sum_{j \in S} w_j \left( \sum_{i \in I_j} \gamma_{i,t} \right).$$
Therefore $\Pi_t^{AF}(\beta, \gamma) \leq \Pi_t^{CDLP}(\beta, \gamma)$ and by Lemma 1, $V_t^{AF} \leq V_t^{CDLP}$, which gives an alternative proof of the $AF$ bound being tighter than the $CDLP$ bound. More importantly, the comparison hints at how we can obtain tractable relaxations that are tighter than $CDLP$.

### 4.3 Weak affine relaxation

We motivate our first relaxation, which we refer to as weak affine relaxation $wAR$, as follows: The difficult term in (15) is the $\gamma_S,t(1 + \sum_{j \in S} w_j)$, and $CDLP$ is tractable as it replaces this by $\lambda \sum_{j \in S} w_j (\sum_{i \in I_j} \gamma_i,t)$. We instead replace the $\gamma_S,t(1 + \sum_{j \in S} w_j)$ term in (15) with $\gamma_S,t + \sum_{j \in S} w_j (\sum_{i \in I_j} \gamma_i,t)$ and solve the linear program

$$V_t^{wAR} = \min_{\beta, \gamma} \sum_{t} \beta_t + \sum_{t} \sum_{i} \gamma_{i,t} r_i^{1}$$

s.t. $\beta_t \geq \Pi_t^{wAR}(\beta, \gamma)$ \quad \forall t$

(19)

where

$$\Pi_t^{wAR} = \max_{S} \left\{ -\gamma_{S,t} - \sum_{j \in S} w_j \left( \sum_{i \in I_j} \gamma_i,t \right) - \sum_{j \in S} w_j \left[ \beta_t + \lambda \left( \sum_{i \in I_j} \sum_{k=t+1}^{\tau} \gamma_{i,k} \right) - f_j \right] \right\}. \quad (20)$$

Proposition 3 below shows that $wAR$ obtains an upper bound on the value function that is weaker than $AF$ but stronger than $CDLP$. In Appendix A, we show that it also gives a tighter upper bound than by working with a continuous relaxation of $\Pi_t^{AF}(\beta, \gamma)$.

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

**Proof**

The proof follows by noting that

$$\gamma_{S,t} \left( 1 + \sum_{j \in S} w_j \right) \geq \gamma_{S,t} + \sum_{j \in S} w_j \left( \sum_{i \in I_j} \gamma_i,t \right) \geq \lambda \sum_{j \in S} w_j \left( \sum_{i \in I_j} \gamma_i,t \right).$$

Therefore $\Pi_t^{AF}(\beta, \gamma) \leq \Pi_t^{wAR}(\beta, \gamma) \leq \Pi_t^{CDLP}(\beta, \gamma)$ and the result now follows from Lemma 1.

\[\square\]

In the remainder of this section, we show that the weak affine relaxation upper bound, $V_t^{wAR}$, can be obtained in a tractable manner; moreover we show that the weak affine relaxation LP can, in fact, be reformulated as a compact linear program where the number of variables and constraints is polynomial in the number of products and resources. This is appealing since it eliminates the need for constraint generation to solve the problem.

We begin by observing that solving problem (20) in an efficient manner is key to separating the weak affine relaxation constraints efficiently. Therefore, we focus on solving the optimization problem (20). Introducing decision variables $q_i$ and $u_j$, respectively, to indicate if resource $i$ and
product $j$ are open, problem (20) can be formulated as the integer program

$$
\Pi^\text{wAR}_t(\beta, \gamma) = \max_{q,u} - \sum_i \gamma_i t q_i - \sum_j w_j \left[ \beta_t + \lambda \left( \sum_{i \in \mathcal{I}_j} \sum_{k=t+1}^{\tau} \gamma_i k \right) - f_j \right] + \sum_i \gamma_i t u_j
$$

s.t. $u_j - q_i \leq 0 \quad \forall i \in \mathcal{I}_j, \forall j$  
$q_i \leq 1 \quad \forall i$  
$u_j \geq 0$ integer.  

Note that the first constraint ensures that a product is open only if all the resources it consumes are open.

Now, just observe that the constraint matrix of the above integer program has exactly one +1 and one −1 coefficient in each row, and hence is totally unimodular. So we can ignore the integer restriction and solve (21)–(24) exactly as a linear program. In fact, problem (21)–(24) can also be solved combinatorially as a flow problem: the dual of the LP can be transformed to be a flow problem on a bipartite graph with one set of nodes representing products and the other side resources and edges representing product-resource incidence, and flow from a source to a sink node, each connected to the product and resource nodes respectively; fast algorithms of Ahuja, Orlin, Stein, and Tarjan [1] can then be used to solve the problem in time $O(m|E| + \min(m^2, m^2 \sqrt{|E|})$ where $|E|$ is the number of edges in this graph. Therefore, problem (21)–(24) can be solved efficiently and so, separating the $wAR$ constraints is tractable.

We next show that $wAR$ can be formulated as a compact LP eliminating the need for generating constraints on the fly. Since the separation problem can be solved as an LP where all the fixed values $(\beta, \gamma)$ appear in the objective function only, we can fold it in into the original LP as follows: First take the dual of (21)–(24) with dual variables $\pi_{i,j}$ corresponding to (22), and $\psi_i$ to (23):

$$
\Pi^\text{wAR}_t(\beta, \gamma) = \min_{\pi,\psi} \sum_i \psi_i
$$

s.t. $\sum_{i \in \mathcal{I}_j} \pi_{i,j} \geq -w_j \left[ \beta_t + \lambda \left( \sum_{i \in \mathcal{I}_j} \sum_{k=t+1}^{\tau} \gamma_i k \right) - f_j \right] + \sum_i \gamma_i t \pi_{i,j} \forall j$  
$-\sum_{j \in \mathcal{J}_i} \pi_{i,j} + \psi_i = -\gamma_{i,t} \forall i$  
$\pi_{i,j}, \psi_i \geq 0.$

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

$$
\Pi^\text{wAR}_t(\beta, \gamma) = \min_{\pi,\gamma} \sum_i \left[ \sum_{j \in \mathcal{J}_i} \pi_{i,j} - \gamma_{i,t} \right]
$$

s.t. $\sum_{i \in \mathcal{I}_j} \pi_{i,j} \geq -w_j \left[ \beta_t + \lambda \left( \sum_{i \in \mathcal{I}_j} \sum_{k=t+1}^{\tau} \gamma_i k \right) - f_j \right] + \sum_i \gamma_i t \pi_{i,j} \forall j$  
$\sum_{j \in \mathcal{J}_i} \pi_{i,j} \geq \gamma_{i,t} \forall i$  
$\pi_{i,j} \geq 0.$

Constraints (19) amount to a condition that $\beta_t \geq \Pi^\text{wAR}_t(\beta, \gamma)$ for all $t$, which can be written, in lieu of (19), as (26–28) below. Putting everything together, the linear program for our first tractable
approximation in its entirety can be written as:

\[ V^{wAR} = \min_{\beta, \gamma, \pi} \sum_t \beta_t + \sum_i r_i^1 \gamma_{i,t} \]  
\[ \text{s.t.} \quad \beta_t \geq \sum_i \left( \sum_{j \in J_i} \pi_{i,j,t} - \gamma_{i,t} \right) \quad \forall t \]  
\[ \sum_{i \in I_i} \pi_{i,j,t} \geq -w_j \left( \beta_t + \lambda \left( \sum_{i \in I_i} \sum_{k=t+1}^r \gamma_{i,k} \right) - f_j \right) + \sum_i \gamma_{i,t} \quad \forall t, j \]  
\[ \sum_{j \in J_i} \pi_{i,j,t} \geq \gamma_{i,t} \quad \forall i, t \]  
\[ \gamma_{i,t}, \pi_{i,j,t} \geq 0. \]  

The size of the above LP is polynomial in the number of resources, products and the length of the booking horizon. Hence, not only is \( wAR \) stronger than \( CDLP \), it is also tractable and has a compact formulation. Notice that this formulation would have been hard to derive and justify without the line of reasoning starting from \( AF \).

The dual of problem (25)–(29) gives more insight into the formulation. By associating dual variables with constraints (26), (27), and (28), and after some simplifications, we get the dual LP as

\[ V^{wAR} = \max_{x, \rho} \sum_t \sum_j \lambda_j x_{j,t} \]  
\[ \text{s.t.} \quad x_{0,t} + \sum_{s=1}^{t-1} \sum_{j \in J_i} \lambda_j x_{j,s} + \sum_{j \in J_i} x_{j,t} - \rho_{i,t} \leq r_i^1 \quad \forall i, t \]  
\[ (dwAR) \quad x_{0,t} + \sum_j x_{j,t} = 1 \quad \forall t \]  
\[ \frac{x_{j,t}}{w_j} - x_{0,t} + \rho_{i,t} \leq 0 \quad \forall i, j \in J_i, t \]  
\[ x_{0,t}, x_{j,t}, \rho_{i,t} \geq 0. \]  

If we interpret \( x_{j,t} \) as the sales rate for product \( j \) at time \( t \) and \( x_{0,t} - \rho_{i,t} \) as the resource level no-purchase rate at time \( t \), then we can view \( wAR \) as a refinement of \( SBLP \) where the sales rates at each time period are modulated by the expected remaining resource capacities. Comparing \( dwAR \) with \( SBLP \), it is clear that a feasible solution to \( dwAR \) is also feasible to \( SBLP \), which gives an alternative proof of Proposition 3. In the following sections, we describe tractable approximation methods that further tighten the \( wAR \) bound.

### 4.4 A tighter relaxation

In this section, we describe a simple way to tighten the \( wAR \) bound. Associating decision variables \( q_i \) and \( u_{j,t} \), respectively, to indicate if resource \( i \) and product \( j \) are open, the \( AF \) separation problem (15) can be written as

\[ \Pi_t^{AF}(\beta, \gamma) = \max_{q, u} - \sum_i \gamma_{i,t} q_i \left( 1 + \sum_j w_j u_{j,t} \right) - \sum_j w_j \left[ \beta_t + \lambda \left( \sum_{i \in I_i} \sum_{k=t+1}^r \gamma_{i,k} \right) - f_j \right] u_{j,t} \]  
\[ \text{s.t.} \quad (22), (23), (24). \]
Now \( wAR \) replaces the product term \( q_i u_j \) for \( i \neq j \) in the first summation with 0 and since \( q_i u_j \geq 0 \), we have \( \Pi_t^{AF}(\beta, \gamma) \leq \Pi_t^{wAR}(\beta, \gamma) \). Noting that \( q_i u_j \geq q_i + u_j - 1 \), we propose replacing the right hand side of constraints (16) with

\[
\Pi_t^{wAR^+}(\beta, \gamma) = \max_{q, u} - \sum_i \gamma_{i,t} q_i - \sum_i \sum_{j \neq i} \gamma_{i,t} w_{i,j} \zeta_{i,j}
- \sum_j w_j \left[ \beta_t + \lambda \left( \sum_{j \in I} \sum_{k=t+1}^T \gamma_{i,k} \right) - f_j \right] + \sum_i \gamma_{i,t} u_j
\]

s.t (22), (23)

\[ \zeta_{i,j} \geq q_i + u_j - 1 \quad \forall i, j \neq i \]
\[ u_j, \zeta_{i,j} \geq 0. \]

The following lemma is immediate.

**Lemma 2.** \( \Pi_t^{AF}(\beta, \gamma) \leq \Pi_t^{wAR^+}(\beta, \gamma) \leq \Pi_t^{wAR}(\beta, \gamma) \).

Therefore, we replace the right hand side of constraints (16) with \( \Pi_t^{wAR^+}(\beta, \gamma) \) and solve the LP

\[
V^{wAR^+} = \min_{\beta, \gamma} \sum_t \beta_t + \sum_t \sum_i \gamma_{i,t} r_i^t
\]

\[(wAR^+)\text{ s.t.} \quad \beta_t \geq \Pi_t^{wAR^+}(\beta, \gamma) \quad \forall t
\]

\[ \gamma_{i,t} \geq 0. \]

We refer to this method as \( wAR^+ \). Lemma 2 together with Lemma 1 implies that \( V^{AF} \leq V^{wAR^+} \leq V^{wAR} \). The \( wAR^+ \) separation problem can be solved as a linear program and hence \( wAR^+ \) is tractable as well. Moreover, it is possible to obtain a compact formulation of \( wAR^+ \) by following the steps in §4.3; we omit the details.

### 4.5 A hierarchical family of relaxations

In this section we show how to construct a hierarchical family of relaxations that at the highest level (level-\( n \)), the number of products) gives us the affine relaxation. Naturally, because of the NP-hardness of solving the affine relaxation, we cannot expect tractability, so we concentrate on small levels. The level-1 relaxation already turns out to be a tighter relaxation than \( wAR \). While the level-1 relaxation separation problem can be solved in a tractable manner, a potential drawback is that it cannot be folded into the original problem to yield a compact formulation.

For simplicity we describe the level-1 formulation and remark on how it extends to a hierarchy of relaxations. In the level-1 relaxation, which we refer to as hierarchical affine relaxation (\( hAR \)), we replace the \( \gamma_{S,t}(1 + \sum_{j \in S} w_j) \) term in (15) with \( \gamma_{S,t} + (\sum_{j \in S} w_j)(\max_{j' \in S} \sum_i \gamma_{i,t}) \) and solve the LP

\[
V^{hAR} = \min_{\beta, \gamma} \sum_t \beta_t + \sum_t \sum_i \gamma_{i,t} r_i^t
\]

\[(hAR)\text{ s.t.} \quad \beta_t \geq \Pi_t^{hAR}(\beta, \gamma) \quad \forall t
\]

\[ \gamma_{i,t} \geq 0, \]

\[ \gamma_{i,t} \geq 0, \]
where
\[
\Pi_t^{h_{AR}} = \max_S \left\{ -\gamma_{S,t} - \left( \sum_{j \in S} w_j \right) \left( \max_{j' \in S} \sum_{i \in I_{j'}} \gamma_{i,t} \right) - \sum_{j \in S} w_j \left[ \beta_t + \lambda \left( \sum_{i \in I_j} \sum_{k=t+1}^\tau \gamma_{i,k} \right) - f_j \right] \right\}.
\] (32)

We have the following lemma.

**Lemma 3.** \(\Pi_t^{AF}(\beta, \gamma) \leq \Pi_t^{h_{AR}}(\beta, \gamma) \leq \Pi_t^{w_{AR}}(\beta, \gamma)\).

**Proof**
The proof follows by noting that \(\gamma_{S,t} \left( 1 + \sum_{j \in S} w_j \right) \geq \gamma_{S,t} + \left( \sum_{j \in S} w_j \right) \left( \max_{j' \in S} \sum_{i \in I_{j'}} \gamma_{i,t} \right) \geq \gamma_{S,t} + \sum_{j \in S} w_j \left( \sum_{i \in I_j} \gamma_{i,t} \right)\).

Lemma 3 together with Lemma 1 implies that \(V^{AF} \leq V^{h_{AR}} \leq V^{w_{AR}}\). Next, we show that the separation problem (32) can be solved in a tractable manner. Associating binary decision variables \(q_i\) and \(u_j\), respectively, to indicate if resource \(i\) and product \(j\) are open, problem (32) can be written as
\[
\Pi_t^{h_{AR}}(\beta, \gamma) = \max_{q,u} \left\{ -\gamma_{S,t} - \sum_i \gamma_{i,t} q_i - \left( \sum_j w_j u_j \right) \left( \max_{j' \in S} \sum_{i \in I_{j'}} \gamma_{i,t} u_j ' \right) - \sum_j w_j \left[ \beta_t + \lambda \left( \sum_{i \in I_j} \sum_{k=t+1}^\tau \gamma_{i,k} \right) - f_j \right] u_j \right\}
\]
\[\text{s.t } (22) - (24).\]

Although the above optimization problem has a nonlinear objective function, we can solve it through a sequence of linear programs in the following manner. We fix a product \(\tilde{j}\) as the one achieving the maximum value of \(\max_{j'} \gamma_{i,t} u_{j'}\). Since \(\tilde{j}\) achieves the maximum value, we must have \(u_{\tilde{j}} = 1\) and \(u_j = 0\) for \(j \neq \tilde{j}\) with \(\sum_{i \in I_{\tilde{j}}} \gamma_{i,t} > \sum_{i \in I_j} \gamma_{i,t}\). Letting \(\tilde{J} = \left\{ j \mid \sum_{i \in I_j} \gamma_{i,t} > \sum_{i \in I_{\tilde{j}}} \gamma_{i,t} \right\}\), we solve the following linear integer program for product \(\tilde{j}\):
\[
\Pi_t^{h_{AR}\tilde{j}}(\beta, \gamma) = \max_{q,u} \left\{ -\sum_i \gamma_{i,t} q_i - \sum_j w_j \left[ \beta_t + \lambda \left( \sum_{i \in I_{\tilde{j}}} \sum_{k=t+1}^\tau \gamma_{i,k} \right) - f_j \right] u_j \right\}
\]
\[\text{s.t } (22), (23) \]
\[u_{\tilde{j}} = 1\]
\[u_j = 0 \quad \forall j \in \tilde{J}\]
\[u_j \geq 0 \text{ integer } \forall j \in J \backslash \tilde{J}\]

Since the constraint matrix is totally unimodular, we can solve the above linear integer program as a linear program. So we solve the linear program for each product \(\tilde{j} \in J\) and obtain \(\Pi_t^{h_{AR}}(\beta, \gamma) = \max_{\tilde{j} \in J} \Pi_t^{h_{AR}\tilde{j}}(\beta, \gamma)\).
Since problem (32) can be solved in a tractable manner, separating the $hAR$ constraints is tractable, and $hAR$ can be solved in polynomial time by the ellipsoid method. However, unlike $wAR$ and $wAR^+$, $hAR$ does not seem to have a compact linear programming formulation. This is because the set $J_j$ depends on the values of the $\gamma$'s in a nonlinear fashion and the duality argument in §4.3 that we used to fold the separation problem back into the original LP does not hold.

Remark: One can get further relaxations by considering pairs of elements $j', j''$ for a level-2 relaxation (or triples for level-3, and so on) such that we find the $S$ that maximizes

$$- \left( 1 + \sum_{j \in S} w_j \right) \left( \max_{(j', j'' \in S)} \sum_{i \in I_{(j', j'')}} \gamma_{i,t} \right).$$

So we control the degree of approximation to the affine relaxation. We limit our numerical results to fixing a single element $j'$.

5 Extensions

In this section we describe how to extend the single-segment MNL weak affine relaxation of §4.3 to other choice models (the development for §4.4 and §4.5 is similar). In §5.1 we consider the MNL choice model with multiple customer segments and disjoint consideration sets. In §5.2 we consider the case where the consideration sets of the different segments may overlap. In §5.3 we consider the nested-logit choice model. We show that our formulation gives an upper bound that falls in between the CDLP and $AF$ bounds, as for the single-segment case.

5.1 Multiple segments with disjoint consideration sets

Let $I_t = \{i \in I \mid \exists j \in C_l$ and $j \in J_l \}$ and the set of segments that use $i$ as $L_i = \{l \in L | i \in I_t \}$ and recall that we use, for $S \subset J_l, i \in I_S$ if there is some $j \in S$ with $j \in J_l$.

We begin with the separation problem for $AF$. Using $\lambda Q_i(S) = \sum_l \lambda_l Q_i^l(S_l)$ and $\lambda R(S) = \sum_l \lambda_l R^l(S_l)$, where $S_l = S \cap C_l$, constraint (11) can be written as

$$\beta_{l,t} + \sum_i \mathbb{I}_{i \in I_S} \gamma_{i,t} + \sum_i \left( \sum_{k=t+1}^\tau \gamma_{i,k} \right) \sum_l \lambda_l Q_i^l(S_l) \geq \sum_l \lambda_l R^l(S_l). \quad (33)$$

We first split this constraint into $l$ separate constraints, one for each segment, by introducing variables $\beta_{l,t}$. The constraint for segment $l$ at time $t$ is that

$$\beta_{l,t} + \sum_{i \in I_S} \mathbb{I}_{i \in I_S} \gamma_{i,t} + \sum_i \sum_{k=t+1}^\tau \gamma_{i,k} \lambda_l Q_i^l(S_l) \geq \lambda_l R^l(S_l). \quad (34)$$

for each $S_l = S \cap C_l$. The proof of Proposition 4 below shows that the segment level constraints (34) imply (33) and that we obtain a looser upper bound by separating over (34) instead of (33).

Constraints (34) have the same form as constraints (11) and are therefore hard to separate. So we use the same relaxation as we did for the single-segment case to obtain a tractable separation
problem:

\[
\Pi_{t,t}^{swAR}(\beta, \gamma) = \max_{q,u} - \sum_{i \in I_t} \lambda_i \gamma_{i,t} q_i \\
- \sum_{j \in C_t} \lambda_j w_j \left[ \beta_{l,t} + \left( \sum_{i \in I_t} \sum_{k=t+1}^\tau \gamma_{i,k} \right) - f_j + \lambda_t \sum_{i \in I_t} \sum_{l' \in L_i} \gamma_{i,l'} \right] u_j \\
s.t. \quad u_j - q_i \leq 0 \quad \forall i \in I_t, j \in C_t \\
q_i \leq 1 \quad \forall i \in I_t \\
u_j \geq 0 \quad \forall j \in C_t.
\]

We replace constraints (34) with \( \Pi_{t,t}^{swAR}(\beta, \gamma) \) to obtain a segment-based weak affine relaxation (swAR):

\[
V_{swAR} = \min_{\beta, \gamma} \sum_{t} \sum_{l} \beta_{l,t} + \sum_{t} \sum_{i} \gamma_{i,t} r_i^1 \\
s.t. \quad \beta_{l,t} \geq \Pi_{t,t}^{swAR}(\beta, \gamma) \quad \forall l, t \\
\gamma_{i,t} \geq 0.
\]

By following the same steps as for the single-segment case, it is possible to show that \( swAR \) can be formulated as the compact LP

\[
V_{swAR} = \min_{\gamma, \beta, \pi} \sum_{t} \sum_{l} \beta_{l,t} + \sum_{t} \sum_{i} \gamma_{i,t} r_i^1 \\
s.t. \quad \beta_{l,t} \geq \sum_{j \in I_t} \left( \sum_{j'} \pi_{i,j,t} - \lambda_l \frac{\gamma_{i,t}}{\sum_{l' \in L_i} \lambda_{l'}} \right) \quad \forall l, t \\
(\text{swAR}) \quad \sum_{j \in I_t} \sum_{j' \in C_t} \pi_{i,j,t} \geq \lambda_l w_j \left[ f_j - \sum_{i \in I_t} \left( \sum_{k=t+1}^\tau \gamma_{i,k} + \frac{\gamma_{i,t}}{\sum_{l' \in L_i} \lambda_{l'}} \right) - \beta_{l,t} \right] \quad \forall j, t \\
\sum_{j \in I_t} \sum_{j' \in C_t} \pi_{i,j,t} - \lambda_l \frac{\gamma_{i,t}}{\sum_{l' \in L_i} \lambda_{l'}} \geq 0 \quad \forall i, l \in L, t \\
\gamma_{i,t}, \pi_{i,j,t} \geq 0,
\]

where \( \ell_j \) denotes the segment to which product \( j \) belongs. \( swAR \) can be viewed as an extension of \( wAR \) to the MNL model with multiple segments and disjoint consideration sets. Note that \( swAR \) is again tractable as it is a compact LP. Proposition 4 below shows that it also obtains an upper bound on the value function that is tighter than \( CDLP \).

**Proposition 4.** \( V^{AF} \leq V^{swAR} \leq V^{CDLP} \).

**Proof**

Using the MNL choice probabilities \( P_{0,l}(S_l) = \frac{1}{1 + \sum_{s \in S_l} w_k} \) and \( P_{l,j}(S_l) = \frac{w_j}{1 + \sum_{s \in S_l} w_k} \) and rearranging terms, the \( swAR \) constraint \( \beta_{l,t} \geq \Pi_{t,t}^{swAR}(\beta, \gamma) \) can be equivalently written as

\[
\beta_{l,t} \geq \lambda_l \left[ R_l(S_l) - \sum_{i \in I_t} \sum_{k=t+1}^\tau Q_{l,j}(S_l) \gamma_{i,k} \right] - \sum_{i \in I_t} \sum_{l' \in L_i} \gamma_{i,l'} \left( \sum_{j \in I_t} P_{l,j}(S_l) + P_{0,l}(S_l) \right). \quad (35)
\]
Consider now two intermediate problems:

\[
V = \min_{\beta, \gamma} \sum_{t} \sum_{l} \beta_{t,l} + \sum_{t} \gamma_{t} r_{t}^{1},
\]

s.t (34), \( \gamma_{t} \geq 0 \),

and

\[
\hat{V} = \min_{\hat{\beta}, \hat{\gamma}} \sum_{t} \sum_{l} \hat{\beta}_{t,l} + \sum_{t} \hat{\gamma}_{t} r_{t}^{1},
\]

s.t \( \beta_{t,l} \geq \lambda_{l} \left[ R_{l}(S_{l}) - \sum_{i} \sum_{k=t+1}^{t+\tau} Q_{l}^{i}(S_{l}) \gamma_{i,k} \right], \forall l, t, S_{l} \subseteq C_{l} \) (36)

\( \hat{\beta}_{t,l} \geq 0 \).

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 \( CDLP \).

We next show that \( V^{AF} \leq V \leq V^{swAR} \leq \hat{V} = V^{CDLP} \), which completes the proof of the proposition.

(i) \( V \leq V^{swAR} \leq \hat{V} \).

Since the objective functions of all the problems are the same, we only need to compare the corresponding constraints. Since \( \sum_{i} P_{j}^{i}(S_{l}) + P_{0}^{i}(S_{l}) \leq 1 \), it follows that constraint (35) implies constraint (34) and we have \( V \leq V^{swAR} \).

On the other hand, the right hand side of constraint (36) can be written as

\[
\lambda_{l} \left[ R_{l}(S_{l}) - \sum_{i} \sum_{k=t+1}^{t+\tau} Q_{l}^{i}(S_{l}) \gamma_{i,k} \right] - \sum_{i} \lambda_{i} Q_{l}^{i}(S_{l}) \gamma_{i,t}.
\]

Now note that

\[
\lambda_{l} Q_{l}^{i}(S_{l}) \gamma_{i,t} = \lambda_{l} \mathbb{I}_{[i \in I_{S_{l}}]} Q_{l}^{i}(S_{l}) \gamma_{i,t} = \lambda_{l} \mathbb{I}_{[i \in I_{S_{l}}]} \left[ \sum_{j \in J_{l}} P_{j}^{i}(S_{l}) \right] \gamma_{i,t}
\]

\( \leq \sum_{l' \in L_{l}} \mathbb{I}_{[i \in I_{S_{l}}]} \left[ \sum_{j \in J_{l}} P_{j}^{l'}(S_{l}) \right] \gamma_{i,t} \leq \sum_{l' \in L_{l}} \mathbb{I}_{[i \in I_{S_{l}}]} \left[ \sum_{j \in J_{l}} P_{j}^{l'}(S_{l}) + P_{0}^{l'}(S_{l}) \right] \gamma_{i,t}
\]

where the first equality holds since if \( \mathbb{I}_{[i \in I_{S_{l}}]} = 0 \), then \( Q_{l}^{i}(S_{l}) = 0 \) and the first inequality holds since \( \sum_{l' \in L_{l}} \lambda_{l'} \leq 1 \). Therefore constraint (36) implies constraint (35) and we have \( V^{swAR} \leq \hat{V} \).

(ii) \( V^{AF} \leq V \).

Suppose that \((\hat{\beta}, \hat{\gamma})\) satisfies constraints (34). We show that it satisfies constraints (33) as well.
Fix a set $S$ and let $S_i = S \cap C_i$. Adding up constraints (34) for all the segments

$$\sum_t \tilde{\beta}_i,t \geq \sum_t \left\{ \lambda_t \left[ R^l(S_i) - \sum_{i \in I_i} \sum_{k=t+1}^\tau Q_i^l(S_i) \tilde{\gamma}_{i,k} \right] - \sum_{i \in I_i} \frac{\lambda_t}{\sum_{l' \in \mathcal{L}_t} \lambda_{l'}} \tilde{\gamma}_{i,t} \right\}
$$

$$= \lambda \left[ R(S) - \sum_{i} \sum_{k=t+1}^\tau Q_i(S) \tilde{\gamma}_{i,k} \right] - \sum_{i} \sum_{l \in \mathcal{L}_i} \frac{\lambda_t}{\sum_{l' \in \mathcal{L}_i} \lambda_{l'}} \tilde{\gamma}_{i,t}$$

$$\geq \lambda \left[ R(S) - \sum_{i} \sum_{k=t+1}^\tau Q_i(S) \tilde{\gamma}_{i,k} \right] - \sum_{i} \sum_{l \in \mathcal{L}_i} \frac{\lambda_t}{\sum_{l' \in \mathcal{L}_i} \lambda_{l'}} \tilde{\gamma}_{i,t}$$

where the first equality uses the fact that $Q_i^l(S_i) = 0$ for $l \notin \mathcal{L}_i$ and hence $\lambda Q_i(S) = \sum_{i \in I_i} \lambda_t Q_i^l(S_i) = \sum_{i \in \mathcal{L}_i} \lambda_t$. The inequality holds since $\mathbb{1}_{[i \in I_i]} \leq \mathbb{1}_{[i \in \mathcal{L}_i]}$. Letting $\tilde{\beta} = \{ \tilde{\beta}_i \} = \{ \tilde{\beta}_{i,t} \} \forall t$, it follows that $(\tilde{\beta}, \tilde{\gamma})$ satisfies constraints (33). Therefore $V^A_F \leq \sum_t \tilde{\beta}_i + \sum_t \sum_i \tilde{\gamma}_{i,t} = V$.

Meissner et al. [11] prove the following that we include for completeness.

(iii) $V = V^{CDLP}. \quad \text{([11])}$

Constraints (5) in $dCDLP$ are equivalent to

$$\beta_t = \max_S \left\{ \lambda \left[ R(S) - \sum_{i} \sum_{k=t}^\tau Q_i(S) \gamma_{i,k} \right] \right\}
$$

$$= \max_S \left\{ \sum_{i} \lambda_t \left[ R^l(S) - \sum_{i \in I_i} \sum_{k=t}^\tau Q_i^l(S \cap C_i) \gamma_{i,k} \right] \right\}
$$

$$= \max_{S_i} \left\{ \lambda_t \left[ R^l(S_i) - \sum_{i \in I_i} \sum_{k=t}^\tau Q_i^l(S_i) \gamma_{i,k} \right] \right\}
$$

where the last inequality uses the fact that the consideration sets are disjoint. Therefore, the $dCDLP$ constraint is equivalent to the constraints $\beta_t = \sum_{S_i} \tilde{\beta}_{i,t}$ and

$$\beta_{i,t} = \max_{S_i} \left\{ \lambda_t \left[ R^l(S_i) - \sum_{i \in I_i} \sum_{k=t}^\tau Q_i^l(S_i) \gamma_{i,k} \right] \right\},
$$

which is exactly constraint (36).

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 swAR, which we give below,
V_{dswAR} = \max_{x, \rho} \sum_l \lambda_l \left[ \sum_{j \in C_l} f_j x_{j,t}^l \right]
\text{s.t.} \quad \sum_l \lambda_l \left[ \sum_{l' \in C_l} x_{0,t}^l + \sum_{s=1}^{t-1} \sum_{j \in J_{s}, j \in C_l} x_{j,s}^l + \sum_{l' \in C_l} \sum_{v} \lambda_v \sum_{l' \in C_l} \lambda_v \right] \leq r_i^l \quad \forall i, t
\quad x_{0,t}^l + \sum_{j \in C_l} x_{j,t}^l = 1 \quad \forall l, t
\quad \frac{x_{j,t}^l}{w_j^l} - x_{0,t}^l + \rho_i^l \leq 0 \quad \forall l, i, j \in J_{i}, j \in C_i, t
\quad x_{0,t}^l, x_{j,t}^l, \rho_i^l \geq 0. \quad (37)

The $x_{j,t}^l$ have the same interpretation as the variables in the compact formulation due to Gallego et al. [7] in §4.1.1: as the rate of sales of product $j$ at time period $t$.

5.2 Tightening the formulation for 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. [11] 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 [15] 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 $dswAR$, in the sense that after adding them, the value of the resulting linear program is an upper bound on $V_{DPLP}$.

5.2.1 PC-equalities

We first state the equalities for MNL, to be added to the formulation $SBLP$ of (12). They provide consistency conditions for the marginal distributions. The intuition behind them (taken verbatim from Meissner et al. [11]) is as follows:

For any product $j \in C_l \cap C_k$, the length of time that product $j$ is offered to segment $l$ must be equal to the length of time that it is being offered to segment $k$. Interpret $h_S = \sum_k h_{S,k}$ in CDLP as a distribution over subsets of $J$, or as a randomization rule—at each point choose a subset based on this distribution. The distribution in turn induces a distribution for each one of the segments $l$.

Let $X_j$ be a Bernoulli random variable which takes the value $X_j = 1$ if $j \in S$ for an offer set $S$ sampled from the $h_S$ distribution, and $X_j = 0$ otherwise. The expectation $E[X_j]$ is then the probability that product $j$ is offered under this randomized rule. Consider a similar sampling from another distribution given by $h_{S_l}^l$'s focusing only on segment $l$. This would also lead to a Bernoulli random variable, and if the $h_{S_l}^l$ are induced by the $h_S$'s, the expectations of these random variables

\begin{align*}
\end{align*}
should coincide across the segments; i.e., the $E[X_j]$ should be the same for two segments $l$ and $k$ whose consideration sets contain the product $j$, leading to the constraint:

$$
\sum_{\{S_l \subseteq \mathcal{C} | j \in S_l\}} h_{S_l}^l = \sum_{\{S_k \subseteq \mathcal{C} | j \in S_k\}} h_{S_k}^k = E[X_j].
$$

The product cut constraints can be specialized to MNL as follows; a fuller explanation for their specialization to MNL can be found in [17]:

$$\frac{x_j^l}{\tau w_k^l} = \lambda_l \sum_{\{S_{l,m} \subseteq (\mathcal{C} \cap \mathcal{C}_m) | k \in S_{l,m}\}} x_{S_{l,m}}^{l,m}, \forall k \in \mathcal{C}_l \cap \mathcal{C}_m, \forall l, m \quad (38)$$

$$x_{S_{l,m},k}^{l,m} \leq x_{S_{l,m}}^{l,m}, \forall S_{l,m} \subseteq \mathcal{C}_l \cap \mathcal{C}_m, k \in \mathcal{C}_l \cap \mathcal{C}_m, \forall l, m \quad (39)$$

$$\sum_{\{S_{l,m} \subseteq (\mathcal{C} \cap \mathcal{C}_m) | S_{l,m} \supseteq \tilde{S}_{l,m}\}} \left\{ \sum_{k \in C_l \setminus C_m} w_k^{l,m} x_{S_{l,m},k}^{l,m} + (1 + w_S^{l,m}) x_{S_{l,m}}^{l,m} \right\} = \quad (40)$$

$$\sum_{\{S_{l,m} \subseteq (\mathcal{C} \cap \mathcal{C}_m) | S_{l,m} \supseteq \tilde{S}_{l,m}\}} \left\{ \sum_{k \in C_m \setminus C_l} w_k^{m,l} x_{S_{l,m},l}^{m,l} + (1 + w_S^{m,l}) x_{S_{l,m}}^{m,l} \right\}, \forall \tilde{S}_{l,m} \subseteq \mathcal{C}_l \cap \mathcal{C}_m, \forall l, m$$

where $x_j^l = \lambda_l \sum_{t} x_{j,t}^{l} w_{i,m}^l = \sum_{j \in S_{l,m}} w_j^l$ and we have new variables of the form $x_{S_{l,m}}^{l,m}$ defined for all pairs of segments $l, m$ and for all $S_{l,m} \subseteq \mathcal{C}_l \cap \mathcal{C}_m$. We refer to $SBLP$ with (38–40) as $SBLP_{PC}$. While not compact, when the size of the intersections $(\mathcal{C} \cap \mathcal{C}_m)$ is small, this formulation is tractable.

Now we show that (38–40) can be added to $dswAR$ and the resulting linear program gives an upper bound on $V^DPLP$.

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

To show validity, as the feasible region of $DPLP$ is contained in the feasible regions of $dswAR1$ as well as that of $SBLP_+$ (the latter shown in [17]), 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 4, but we give a direct proof: We make the connection between $dswAR1$ and $SBLP$ variables via $x_j^l = \lambda_l \sum_{t} x_{j,t}^{l} \tau w_k^l$, and $x_0^l = \lambda_l \sum_{t} x_{0,t}^{l} \tau$. So $SBLP$ can be written in terms of the time-indexed variables $x_{j,t}^l$ and we consider $dswAR1$ as an extended formulation with new variables $\rho_{l,t}^j$, and the projection of $dswAR1$ into the space of the variables $x_{j,t}^l$ is now shown to be a subset of $SBLP$. Consider a feasible solution of $dswAR1$. The solution clearly satisfies $\sum_{t} \lambda_l (x_{0,t}^l + \sum_{j \in \mathcal{C}_l} x_{j,t}^l) = \lambda_l \tau$ and hence $x_0^l + \sum_{j \in \mathcal{C}_l} x_{j}^l = \lambda_l \tau$. Likewise, $\frac{x_0^{l}}{x_{j}^l} - x_0^l \leq -\lambda_l \sum_{t} \rho_{l,t}^j \leq 0$.

So the only remaining set of constraints to verify is (13). Consider the constraints of $dswAR1$ for period $\tau$:

$$\sum_{l} \lambda_l \left[ \frac{x_{0,t}^{l \tau}}{\sum_{t} \lambda_{l} \tau} + \sum_{s=1}^{\tau-1} \sum_{j \in \mathcal{J}_l, j \in \mathcal{C}_l} x_{j,s}^{l \tau} + \frac{\sum_{j \in \mathcal{J}_j, j \in \mathcal{C}_l} x_{j,t}^{l \tau}}{\sum_{j \in \mathcal{J}_l} \lambda_{l}} \right] \leq r_{l}^{t}$$

which can be rewritten as

$$\sum_{l} \sum_{j \in \mathcal{J}_1, j \in \mathcal{C}_l} (x_{j}^{l \tau} - \lambda_l x_{j}^{l \tau}) + \lambda_l \left[ \frac{x_{0,t}^{l \tau}}{\sum_{t} \lambda_{l} \tau} + \sum_{s=1}^{\tau-1} \sum_{j \in \mathcal{J}_l, j \in \mathcal{C}_l} x_{j,s}^{l \tau} + \frac{\sum_{j \in \mathcal{J}_j, j \in \mathcal{C}_l} x_{j,t}^{l \tau}}{\sum_{j \in \mathcal{J}_l} \lambda_{l}} \right] \leq r_{l}^{t}.$$

23
The objective function value of the above constraints can be equivalently written as

$$\sum_{i} \lambda_i \left[ \frac{x_{i,t}^{l}}{\sum_{l'} \lambda_{l'}} + \sum_{j \in J_t \cap C_i} x_{i,t}^{j} \sum_{l'} \lambda_{l'} - \frac{p_{i,t}^{l}}{\sum_{l'} \lambda_{l'}} \right] \geq \sum_{i} \lambda_i \sum_{j \in J_t \cap C_i} x_{i,t}^{j}$$

or

$$\sum_{i} \lambda_i x_{0,t}^{l} + \lambda_i \left( 1 - \sum_{l' \in C_i} \lambda_{l'} \right) \sum_{j \in J_t \cap C_i} x_{j,t}^{l} - \lambda_i p_{i,t}^{l} \geq 0,$$

which is true as (37) implies $p_{i,t}^{l} \leq x_{0,t}^{l}$ and $(1 - \sum_{l' \in C_i} \lambda_{l'}) \geq 0$.

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

**Proposition 5.** The objective function value of $SAR1$ with (38–40) added, with $x_{k}^{l} = \lambda \sum_{t=1}^{T} x_{k,t}^{l}$ in (38), is less than or equal to $V^{SBLP^+}$.

---

**5.3 Nested-Logit choice model**

In this section we briefly outline how our ideas on handling multiple segment MNL can also be applied to the nested logit model of choice; the nested-logit choice model generalizes MNL (but not mixed logit of the previous sections) and is considered in recent research in assortment optimization ([5]).

In the nested-logit model we have $L$ nests and $n$ products in each nest. We index nests by $l$ and products by $j$ and let $f_{i,j}$ denote the revenue associated with product $j$ in nest $l$ and $w_{l,j}$ be its preference weight. We let $w_{l,0}$ be the no-purchase preference weight for nest $l$ and $w_0$ be the no-purchase preference weight at the nest level (i.e., none of the nests are chosen). The probability that product $j$ in nest $l$ is chosen $P_{i,j}(S_l) = P(S)P_{j,l}(S_l)$, where $S_l$ is the set of products offered within nest $l$, $S = (S_1, \ldots, S_L)$ is the complete offer set, $W_i(S_l) = w_{l,0} + \sum_{j \in S_l} w_{l,j}$,

$$P_i(S_l) = \frac{W_i(S_l)^{g_l}}{(w_0 + \sum_k W_k(S_k)^{w_0})}$$

and $P_{j,l}(S_l) = \frac{w_{l,j}}{W_i(S_l)}$; see for example ([5]). We assume $g_l \leq 1$ for all $l$, which is a necessary condition for $CBLP$ to be tractable; see [5].

For the nested-logit choice model, AF constraint (11) can be written as

$$\beta_t + \sum_{i} \mathbb{1}_{[i \in S]} \gamma_{i,t} + \sum_{i} \left[ \sum_{j \in J_t} \lambda P_i(S)P_{j,l}(S_l) (\sum_{k=t+1}^{\tau} \gamma_{i,k}) \right] \geq \sum_{i} \sum_{j} \lambda P_i(S)P_{j,l}(S_l) f_{i,j}$$

for all $t$ and $S$. Using the form of the nested-logit choice probabilities and after some simplifications, the above constraints can be equivalently written as $\beta_t w_0 \geq \Pi_t^{AF,NL}(\beta, \gamma)$ for all $t$, where

$$\Pi_t^{AF,NL}(\beta, \gamma) = \max_{S} \left\{ \sum_{i} W_i(S_i)^{g_l} \left[ \sum_{j} \lambda P_{j,l}(S_l) \left[ f_{i,j} - \sum_{k=t+1}^{\tau} \sum_{l' \in J_t} \gamma_{i,k} \right] - \beta_t \right. \right.$$

$$\left. - \sum_{i} \gamma_{i,t} \mathbb{1}_{[i \in S]} \right) - w_0 \sum_{i} \gamma_{i,t} \mathbb{1}_{[i \in S]} \right\}.$$
Solving the above maximization is hard and so we look at tractable approximations.

Noting that $\mathbb{1}_{[\in\mathcal{I}_S]} \geq \mathbb{1}_{[i \in \mathcal{I}_S]}$ and $\mathbb{1}_{[i \in \mathcal{I}_S]} \geq \sum_i \mathbb{1}_{[i \in \mathcal{I}_S]} / L$,

$$\Pi_{t}^{AF,NL}(\beta, \gamma) \leq \max_{S} \left\{ \sum_{l} W_l(S_l)^{g_l} \left\{ \sum_{j} \lambda p_{j,l}(S_l) \left[ f_{l,j} - \sum_{k=t+1}^{r} \sum_{i \in \mathcal{I}_j} \gamma_{i,k} \right] - \beta_t - \sum_{i} \gamma_{i,t} \mathbb{1}_{[i \in \mathcal{I}_S]} \right\} \right\} \left(41\right)$$

Now the term on the right hand side of (42) separates by nest and we can maximize over the nests separately, the optimization problem for nest $l$ being

$$\tilde{\Pi}_{t,l}(\beta, \gamma) = \max_{S_l} \left\{ W_l(S_l)^{g_l} \left\{ \sum_{j} \lambda p_{j,l}(S_l) \left[ f_{l,j} - \sum_{k=t+1}^{r} \sum_{i \in \mathcal{I}_j} \gamma_{i,k} \right] - \beta_t - \sum_{i} \gamma_{i,t} \mathbb{1}_{[i \in \mathcal{I}_S]} \right\} \right\} \left(43\right)$$

Letting $u_{l,j}, q_i \in \{0,1\}$, respectively indicate product $j$ in nest $l$ being open and resource $i$ being open, problem (43) can be written as

$$\tilde{\Pi}_{t,l}(\beta, \gamma) = \max_{u,q} \frac{w_l,0 + \sum_j w_{l,j} u_{l,j} \lambda}{w_l,0 + \sum_j w_{l,j} u_{l,j}} \left\{ \sum_{j} w_{l,j} u_{l,j} \left[ f_{l,j} - \sum_{k=t+1}^{r} \sum_{i \in \mathcal{I}_j} \gamma_{i,k} \right] - \beta_t \left( w_l,0 + \sum_j w_{l,j} u_{l,j} \right) - \sum_i \gamma_{i,t} q_i \right\} - w_{\phi} \sum_i \gamma_{i,t} q_i / L$$

s.t. $u_{l,j} - q_i \leq 0 \quad \forall i \in \mathcal{I}_j, \forall j$

$q_i, u_{l,j} \in \{0,1\}$.

Since $\beta_t w_{l,0} \geq 0$, $\gamma_{i,t} q_i w_{l,0} \geq 0$, $q_i u_{l,j} = u_{l,j}$ for $j \in \mathcal{I}_i$ and $q_i u_{l,j} \geq 0$ for $j \notin \mathcal{I}_i$, we have $\tilde{\Pi}_{t,l}(\beta, \gamma) \leq \Pi_{t,l}^{AF,NL}(\beta, \gamma)$, where

$$\Pi_{t,l}^{AF,NL}(\beta, \gamma) = \max_{u} \frac{w_l,0 + \sum_j w_{l,j} u_{l,j} \lambda}{w_l,0 + \sum_j w_{l,j} u_{l,j}} \left\{ \sum_{j} w_{l,j} u_{l,j} \left[ f_{l,j} - \sum_{k=t+1}^{r} \sum_{i \in \mathcal{I}_j} \gamma_{i,k} \right] - \beta_t - \sum_{i \in \mathcal{I}_j} \gamma_{i,t} \right\} - \frac{w_{\phi} \sum_{i \in \mathcal{I}_j} \gamma_{i,t} w_{l,j} u_{l,j}}{L}$$

Putting everything together, we have $\Pi_{t,l}^{AF,NL}(\beta, \gamma) \leq \sum_t \Pi_{t,t}^{AF,NL}(\beta, \gamma)$ and we propose replacing the AF constraint $\beta_t w_{\phi} \geq \Pi_{t,t}^{AF,NL}(\beta, \gamma)$ with $\beta_t w_{\phi} \geq \sum_t \Pi_{t,t}^{AF,NL}(\beta, \gamma)$. Lemma 1 implies that by doing this, we get an upper bound on the value function. We show below that this upper bound can be computed in a tractable manner.

We need only show that $\Pi_{t,t}^{AF,NL}(\beta, \gamma)$ can be solved in a tractable manner. To this end, we
write

$$\Pi_{i,t}^{wAR,NL}(\beta, \gamma) = \max_u \frac{1}{(w_{i,0} + \sum_j w_{i,j}u_{i,j})^{1-n}} \left\{ \sum_j w_{i,j}u_{i,j} \left[ \lambda \left( f_{i,j} - \sum_{k=t+1}^{t} \sum_i \gamma_{i,k} \right) - \beta_t - \sum_i \gamma_{i,t} \right] \right\}$$

$$= \max_u \left\{ \sum_j \left( \frac{w_{i,j}u_{i,j}}{w_{i,0} + \sum_j w_{i,j}u_{i,j}} \right)^{1-g_i} w_{i,j}^{g_i} \left[ \lambda \left( f_{i,j} - \sum_{k=t+1}^{t} \sum_i \gamma_{i,k} \right) - \beta_t - \sum_i \gamma_{i,t} \right] \right\}$$

where we use $g_i \leq 1$ and $u_{i,j} = u_{i,j}^{1-g_i}$, since $u_{i,j} \in \{0, 1\}$. Now using the sales rate transformation (see Theorem 1 in [20]) we can write $\Pi_{i,t}^{wAR,NL}(\beta, \gamma)$ as

$$\Pi_{i,t}^{wAR,NL}(\beta, \gamma) = \max_x \sum_j x_{i,j}^{1-g_i} w_{i,j}^{g_i} \left[ \lambda \left( f_{i,j} - \sum_{k=t+1}^{t} \sum_i \gamma_{i,k} \right) - \beta_t - \sum_i \gamma_{i,t} \right] - \frac{w_{i,0}}{l} \sum_j \sum_i \gamma_{i,t} x_{i,j}$$

s.t.  

$$\frac{x_{i,j}}{w_{i,j}} \leq \frac{x_{i,0}}{w_{i,0}}$$

$$x_{i,0} + \sum_j x_{i,j} = 1$$

$$x_{i,j} \geq 0.$$ 

Since $1 - g_i \leq 1$, we have a concave maximization problem, which is tractable.

6 Computational Experiments

In this section, we compare the upper bounds and the revenue performance of the policies obtained by the different benchmark solution methods. We test the performance of our benchmark solution methods on a small airline network, a parallel flights network and a hub and spoke network, with a single hub serving multiple spokes. While comparing the revenue performance, we divide the booking period into five equal intervals. At the beginning of each interval, we resolve the benchmark solution methods to get fresh estimates for the marginal value of capacity on the resources. All of the benchmark methods give a solution of the form $(\hat{\beta}, \hat{\gamma})$ with $\sum_{k=t}^{t} \hat{\gamma}_{k,t}$ being an estimate for the marginal value of capacity on resource $i$ at time $t$. We use these marginal values to construct a value function approximation $\hat{V}_i(r) = \sum_i (\sum_{k=t}^{t} \hat{\gamma}_{k,t}) r_i$ and solve problem (7) to decide on the offer set. We continue to use this decision rule until the beginning of the next interval where we resolve the benchmark solution methods. 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.

Weak affine relaxation (wAR) This is the version of weak affine relaxation that applies to multiple segments and described in §5 (swAR).

Weak affine relaxation+ (wAR+) This is the version of wAR+ that applies to multiple segments. As mentioned, it is possible to extend the wAR+ method described in §4.4 to the setting with multiple segments by following the steps in §5.
Hierarchical affine relaxation (hAR) This is the version of the level-1 hierarchical affine relaxation that applies to multiple segments. As mentioned, it is possible to extend the hierarchical affine relaxation method described in §4.5 to the setting with multiple segments by following the steps in §5. In our computational experiments, we restrict attention to the level-1 relaxation. Since hAR does not admit a compact formulation, we solve hAR by generating constraints on the fly and stop when we are within 1% of optimality.

Alternative weak affine relaxation (awAR) We consider an alternative relaxation, that is complementary to weak affine relaxation. We consider replacing the right hand side of constraints (14) with
\[-\gamma_{s,t} - \sum_{i,j} \left[ \mathbb{1}_{i \in \mathcal{I}_S} + \mathbb{1}_{j \in \mathcal{S} \setminus I} \right] \gamma_{i,t} w_{j} = \sum_{i} \left[ 1 - \mathbb{1}_{i \in \mathcal{I}_S} \right] \left( \sum_{j \in \mathcal{J}_t} \gamma_{i,t} w_{j} \right).\]
We show in Appendix B that the resulting LP gives an upper bound that is tighter than the CDLP upper bound. It is possible to show that the awAR bound is weaker than the wAR bound and that awAR can be formulated as a compact LP; we omit the details.

Affine relaxation (AF) This is the solution method described in §3.2. We work with the reduced formulation RAF of [21]. 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.

6.1 Small Airline Network

We consider an airline network consisting of seven flights that connect three spokes with a hub. We note that the flight legs correspond to the resources in our network RM formulation. There are 22 products, of which half are high fare products whereas the remaining are low fare products. Each origin-destination pair is associated with two customer segments. The first segment is interested only in the high fare products connecting the origin-destination pair, while the second segment is interested only in the low fare products connecting the same origin-destination pair. This set of test problems is drawn from [10] and all the problem parameters are set as in [10]. Let \( w^H_0 \) and \( w^L_0 \) denote the no-purchase preference weights associated with the customer segments interested in the high fare and low fare products, respectively. We vary the no-purchase preference weights and scale the flight leg capacities by a parameter \( \tau \) to obtain different test problems. We label our test problems by the tuple \((\kappa, (w^H_0, w^L_0))\) \( \in \{0.6, 0.8, 1.0\} \times \{ (1, 4), (5, 10), (10, 20) \} \), which gives us a total of 9 test problems. We have \( \tau = 200 \) in all of our test problems.

Table 1 gives the upper bounds obtained by the different solution methods for the test problems on the small airline network. The first column in Table 1 gives the problem characteristics. The second to seventh columns, respectively, give the upper bounds obtained by CDLP, wAR, wAR\(^+\), hAR, awAR and AF. The last five columns give the percentage gap between the upper bounds obtained by CDLP and wAR, CDLP and wAR\(^+\), CDLP and hAR, CDLP and awAR, and CDLP and AF, respectively. The upper bounds obtained by wAR are on average 0.27% tighter than CDLP. wAR\(^+\), hAR, awAR and AF obtain upper bounds that are on average 0.3%, 0.28%, 0.18% and 0.89% tighter than CDLP, respectively. wAR provides a small but consistent improvement over awAR. wAR\(^+\) and hAR both further tighten the wAR bound by a small amount.

Table 2 gives the expected revenues obtained by the different solution methods for the test problems on the small airline network. The columns have a similar interpretation as in Table 1 except that they give the expected total revenues. We evaluate the revenue performance by simulation and use common random numbers in our simulations. In the last five columns, we use ✓ to indicate that the corresponding benchmark method generates higher revenues than CDLP.
at the 95% level, an ⊙ if the difference in the revenue performance of the benchmark method and CDLP is not significant at the 95% level and a × if the benchmark method generates lower revenues than CDLP at the 95% level. The performance gap between wAR and CDLP is around 0.19% on average. The corresponding numbers for wAR⁺, hAR, awAR and AF are 0.49%, 0.57%, -0.04% and 0.95%, respectively. Overall, CDLP generates the lowest revenues, while AF generates the highest. The performance of awAR is comparable with that of CDLP, while wAR, wAR⁺, hAR provide a small but consistent improvement over CDLP.

6.2 Parallel Flights

We consider $N$ parallel flights that operate between the same origin-destination pair. 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 sample the preference weights of the fare-products from a poisson distribution with a mean of 100 and set the no-purchase preference weight to be $0.5\sum_{j \in S_l} w_j$. So the probability that a customer does not purchase anything when all the products in the consideration set are offered is around 33%.

We measure the tightness of the leg capacities using the nominal load factor, which is defined in the following manner. Letting $\hat{S}_{l,t} = \text{argmax } S_l R_l(S_l)$ 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

$$\alpha = \frac{\sum_l \sum_t \lambda_{l,t} Q_l^S(\hat{S}_{l,t})}{\sum_t r^t},$$

where $\lambda_{l,t}$ 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. We refer to these test problems as stationary arrivals. For stationary arrivals, the total arrival rate in each period is 0.9. When the problem parameters are stationary, [9] show that the percentage gap between the CDLP and AF upper bounds vanishes in a fluid scaling of the problem where demand and capacity increases at the same rate. Therefore, we expect the CDLP and AF bounds to be close when all the problem parameters are stationary; see also the computational study in [10]. In this case, the benefit of tightening the CDLP bound through the tractable approximation methods is bound to be marginal. In order to understand settings where the approximation methods might be more beneficial, we also consider a second set of test problems with non-stationary arrival rates. 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 3 gives the upper bounds obtained by the different solution methods for the parallel flights test problems with stationary arrivals. The columns have the same interpretation as in Table 1. The upper bounds obtained by $wAR$ are on average 0.19% tighter than CDLP, while the AF bound is
on average 0.38% tighter than the \textit{CDLP} bound. \(wAR\) further tighten the \(wAR\) bound and obtain upper bounds that are on average 0.26% and 0.21% tighter than \textit{CDLP}, respectively. The \(awAR\) upper bound is on average 0.13% tighter than \textit{CDLP}. \(wAR\) in general obtains tighter bounds than \(awAR\), although we observe instances where the upper bound obtained by \(awAR\) is slightly tighter.

Table 4 gives the expected revenues obtained by the different solution methods for the parallel flights test problems with stationary arrivals. \(wAR\), \(wAR^+\) and \(hAR\) generate significantly higher revenues than \textit{CDLP} and their revenue performance is comparable with that of \textit{AF}. The average performance gap between \textit{AF} and \textit{CDLP} is around 10.1%. On the other hand \(wAR\), \(wAR^+\) and \(hAR\) generate revenues that are on average 11.2%, 11.5% and 9.9% higher than \textit{CDLP}. The performance gap with \textit{CDLP} seems to increase with the nominal load factor. The revenue performance of \(awAR\) is somewhat inferior compared to \(wAR\), but it still generates revenues that are on average 1.9% higher than \textit{CDLP}.

Table 5 gives the upper bounds obtained by the benchmark solution methods for the parallel flights test problems with non-stationary arrivals. The percentage gap between \textit{CDLP} and the other solution methods increases compared to the stationary arrivals case. \(wAR\), \(wAR^+\), \(hAR\), \(awAR\) and \(AF\) on average obtain upper bounds that are 0.7%, 0.86%, 0.72%, 0.2%, and 1.45% tighter than \textit{CDLP}, respectively. \(wAR\) obtains upper bounds that are noticeably tighter than \(awAR\) and roughly closes 50% of the gap between the \textit{CDLP} and \textit{AF} upper bounds. \(wAR^+\) and \(hAR\) further tighten the \(wAR\) bound by a small amount. Table 6 gives the corresponding expected revenues. The average performance gaps for \(wAR\), \(wAR^+\), \(hAR\), \(awAR\) and \(AF\) are 1.68%, 2.51%, 2.4%, 0.64% and 1.66%, respectively. The pattern is broadly similar to the case with stationary arrivals: The revenues generated by \(wAR\), \(wAR^+\) and \(hAR\) are in general comparable with that generated by \(AF\). \(awAR\) provides a slight revenue boost compared to \textit{CDLP}, but falls short of \(wAR\).

### 6.3 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 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. The parameters of the MNL model are set in a similar manner as in the parallel flights test problems. 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 7 gives the upper bounds obtained by the benchmark solution methods for the hub and
spoke network with stationary arrivals. AF generates the tightest upper bound and CDLP the weakest, with the remaining upper bounds sandwiched in between. wAR tends to generate tighter upper bounds than awAR, although we observe one instance where the awAR bound is tighter. The average percentage gap between wAR and CDLP is 1.59%, although we observe instances where the gap is as high as 2.73%. The percentage gap between wAR and CDLP seems to increase with the nominal load factor. wAR+ and hAR tighten the wAR bound and obtain bounds that are on average 1.81% and 1.63% tighter than the CDLP bound. AF generates bounds that are on average 2.16% tighter than the CDLP bound. Table 8 gives the corresponding expected revenues. wAR on average generates revenues that are 2.28% higher than CDLP, although we observe instances where the gap is as high as 7%. wAR+, hAR, awAR and AF, respectively, generate revenues that are on average 3.02%, 2.94%, 1.33% and 2.34%, higher than the CDLP revenues.

Table 9 compares the upper bounds obtained by the benchmark solution methods for the hub and spoke network with non-stationary arrivals, while Table 10 gives the expected revenues. The results display the same trends as before. For both the upper bounds and the expected revenues, the percentage gaps between CDLP and the other solution methods increases on average, compared to the stationary arrivals case. wAR, wAR+, hAR, awAR and AF produce upper bounds that are on average 2.75%, 3.08%, 2.80%, 1.6% and 3.43% tighter than the CDLP bound, respectively. wAR closes roughly 80% of the gap between the CDLP and AF bounds, while wAR+ is able to close almost 90% of the gap. In terms of the revenue improvement, wAR, wAR+, hAR, awAR and AF generate revenues that are on average 3.8%, 3.3%, 3.22%, 2.38% and 3.63% higher than the CDLP revenues. The nominal load factor and the number of the spokes in the network seem to be two factors which lead to larger gaps. Overall, all the tractable approximations generate significantly higher revenues than CDLP and the revenue performance of wAR, wAR+ and hAR is comparable with that of AF.

Table 11 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 the remaining solution methods are generally in minutes. wAR 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. wAR+ and hAR have additional computational overheads associated with them and can take longer than wAR. Overall, wAR seems to achieve a good balance between the quality of the solution and the computational burden.

### Table 1: Comparison of the upper bounds for the small airline network test problems.

<table>
<thead>
<tr>
<th>Problem</th>
<th>(n, (w_H^0, w_L^0))</th>
<th>CDLP</th>
<th>wAR</th>
<th>wAR+</th>
<th>hAR</th>
<th>awAR</th>
<th>AF</th>
<th>% Gap with CDLP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.6, (1, 5))</td>
<td>36,187</td>
<td>36,066</td>
<td>36,053</td>
<td>36,060</td>
<td>36,127</td>
<td>35,802</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>(0.6, (5, 10))</td>
<td>33,158</td>
<td>33,030</td>
<td>33,007</td>
<td>33,029</td>
<td>33,067</td>
<td>32,757</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>(0.6, (10, 20))</td>
<td>29,960</td>
<td>29,838</td>
<td>29,815</td>
<td>29,834</td>
<td>29,903</td>
<td>29,556</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>(0.8, (1, 5))</td>
<td>43,202</td>
<td>43,101</td>
<td>43,093</td>
<td>43,101</td>
<td>43,119</td>
<td>42,759</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.8, (5, 10))</td>
<td>38,900</td>
<td>38,793</td>
<td>38,778</td>
<td>38,793</td>
<td>38,837</td>
<td>38,468</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>(0.8, (10, 20))</td>
<td>34,678</td>
<td>34,571</td>
<td>34,557</td>
<td>34,560</td>
<td>34,586</td>
<td>34,431</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>(1.0, (1, 5))</td>
<td>48,822</td>
<td>48,731</td>
<td>48,727</td>
<td>48,731</td>
<td>48,753</td>
<td>48,440</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(1.0, (5, 10))</td>
<td>43,767</td>
<td>43,649</td>
<td>43,626</td>
<td>43,633</td>
<td>43,682</td>
<td>43,446</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>(1.0, (10, 20))</td>
<td>35,103</td>
<td>35,103</td>
<td>35,103</td>
<td>35,103</td>
<td>35,103</td>
<td>35,103</td>
<td>0.00</td>
</tr>
</tbody>
</table>

avg. 0.27 0.30 0.28 0.18 0.89

Table 1: Comparison of the upper bounds for the small airline network test problems.
Figure 1: Structure of the airline network with a single hub and eight spokes.

Table 2: Comparison of the expected revenues for the small airline network test problems.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Expected Revenue</th>
<th>% Gap with CDLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.6, (1, 5))</td>
<td>34,048 34,059 34,882 34,132 34,836</td>
<td>0.03 (\circ) -0.34 (\circ) 0.25 (\circ) -0.62 (\circ) 0.54 (\circ)</td>
</tr>
<tr>
<td>(0.6, (10, 20))</td>
<td>27,396 27,653 27,796 27,770 27,530</td>
<td>0.94 (\checkmark) 1.46 (\checkmark) 1.37 (\checkmark) 0.49 (\checkmark) 2.29 (\checkmark)</td>
</tr>
<tr>
<td>(0.8, (1, 5))</td>
<td>41,089 40,874 40,973 40,935 40,910</td>
<td>-0.52 (\times) -0.28 (\circ) -0.37 (\circ) -0.44 (\circ) 0.54 (\circ)</td>
</tr>
<tr>
<td>(0.8, (10, 20))</td>
<td>31,150 31,043 31,363 31,123 31,093</td>
<td>-0.34 (\circ) 0.68 (\circ) -0.09 (\circ) -0.83 (\times) 0.76 (\circ)</td>
</tr>
<tr>
<td>avg</td>
<td>0.19 0.49 0.57 -0.04 0.95</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Comparison of the upper bounds for the parallel flights test problems with stationary arrival rates.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Upper Bound</th>
<th>% Gap with CDLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4, 0.8)</td>
<td>11,101 11,079 11,064 11,075 11,077 11,050</td>
<td>0.20 (\circ) 0.33 (\circ) 0.23 (\circ) 0.22 (\circ) 0.46</td>
</tr>
<tr>
<td>(4, 1.0)</td>
<td>9,899 9,864 9,860 9,882 9,848</td>
<td>0.36 0.40 0.36 0.17 0.52</td>
</tr>
<tr>
<td>(4, 1.2)</td>
<td>8,342 8,342 8,342 8,342 8,342</td>
<td>0.00 0.00 0.00 0.00 0.01</td>
</tr>
<tr>
<td>(4, 1.6)</td>
<td>6,217 6,217 6,217 6,217 6,217</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
</tr>
<tr>
<td>(6, 0.8)</td>
<td>12,880 12,850 12,817 12,839 12,834 12,807</td>
<td>0.23 0.49 0.32 0.36 0.57</td>
</tr>
<tr>
<td>(6, 1.0)</td>
<td>11,667 11,599 11,584 11,592 11,633 11,548</td>
<td>0.58 0.70 0.64 0.28 1.02</td>
</tr>
<tr>
<td>(6, 1.2)</td>
<td>9,861 9,859 9,859 9,859 9,843</td>
<td>0.02 0.02 0.02 0.00 0.18</td>
</tr>
<tr>
<td>(6, 1.6)</td>
<td>7,460 7,460 7,460 7,460 7,460</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
</tr>
<tr>
<td>(8, 0.8)</td>
<td>12,695 12,690 12,681 12,682 12,684 12,673</td>
<td>0.04 0.11 0.10 0.09 0.18</td>
</tr>
<tr>
<td>(8, 1.0)</td>
<td>11,817 11,720 11,697 11,718 11,770 11,647</td>
<td>0.82 1.01 0.84 0.40 1.43</td>
</tr>
<tr>
<td>(8, 1.2)</td>
<td>10,070 10,063 10,062 10,063 10,070 10,049</td>
<td>0.06 0.08 0.06 0.00 0.20</td>
</tr>
<tr>
<td>(8, 1.6)</td>
<td>7,524 7,524 7,524 7,524 7,524 7,524</td>
<td>0.00 0.00 0.00 0.00 0.01</td>
</tr>
<tr>
<td>avg</td>
<td>0.19 0.26 0.21 0.13 0.38</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Comparison of the upper bounds for the parallel flights test problems with stationary arrival rates.
### Table 4: Comparison of the expected revenues for the parallel flights test problems with stationary arrival rates.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Expected Revenue</th>
<th>% Gap with CDLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N, α)</td>
<td>CDLP</td>
<td>wAR</td>
</tr>
<tr>
<td>(4, 0.8)</td>
<td>7,935</td>
<td>9,953</td>
</tr>
<tr>
<td>(4, 1.0)</td>
<td>7,862</td>
<td>9,838</td>
</tr>
<tr>
<td>(4, 1.2)</td>
<td>6,877</td>
<td>8,841</td>
</tr>
<tr>
<td>(4, 1.6)</td>
<td>5,504</td>
<td>5,467</td>
</tr>
<tr>
<td>(6, 0.8)</td>
<td>5,127</td>
<td>7,512</td>
</tr>
<tr>
<td>(6, 1.0)</td>
<td>5,819</td>
<td>5,768</td>
</tr>
<tr>
<td>(6, 1.2)</td>
<td>4,704</td>
<td>4,659</td>
</tr>
<tr>
<td>(8, 0.8)</td>
<td>7,043</td>
<td>6,981</td>
</tr>
<tr>
<td>(8, 1.0)</td>
<td>5,859</td>
<td>5,901</td>
</tr>
<tr>
<td>(8, 1.2)</td>
<td>4,225</td>
<td>4,180</td>
</tr>
<tr>
<td>avg.</td>
<td>0.70</td>
<td>0.86</td>
</tr>
</tbody>
</table>

### Table 5: Comparison of the upper bounds for the parallel flights test problems with non-stationary arrival rates.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Expected Revenue</th>
<th>% Gap with CDLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N, α)</td>
<td>CDLP</td>
<td>wAR</td>
</tr>
<tr>
<td>(4, 0.8)</td>
<td>6,975</td>
<td>6,978</td>
</tr>
<tr>
<td>(4, 1.0)</td>
<td>6,777</td>
<td>6,804</td>
</tr>
<tr>
<td>(4, 1.2)</td>
<td>6,109</td>
<td>6,181</td>
</tr>
<tr>
<td>(4, 1.6)</td>
<td>4,762</td>
<td>4,974</td>
</tr>
<tr>
<td>(6, 0.8)</td>
<td>5,794</td>
<td>5,814</td>
</tr>
<tr>
<td>(6, 1.0)</td>
<td>5,312</td>
<td>5,364</td>
</tr>
<tr>
<td>(6, 1.2)</td>
<td>4,851</td>
<td>4,913</td>
</tr>
<tr>
<td>(6, 1.6)</td>
<td>3,966</td>
<td>4,079</td>
</tr>
<tr>
<td>(8, 0.8)</td>
<td>5,217</td>
<td>5,259</td>
</tr>
<tr>
<td>(8, 1.0)</td>
<td>4,734</td>
<td>4,827</td>
</tr>
<tr>
<td>(8, 1.2)</td>
<td>4,208</td>
<td>4,297</td>
</tr>
<tr>
<td>(8, 1.6)</td>
<td>3,489</td>
<td>3,621</td>
</tr>
<tr>
<td>avg.</td>
<td>1.68</td>
<td>2.51</td>
</tr>
</tbody>
</table>

Table 6: Comparison of the expected revenues for the parallel flights test problems with non-stationary arrival rates.
Table 7: Comparison of the upper bounds for the hub and spoke test problems with stationary arrival rates.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Upper Bound</th>
<th>% Gap with CDLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N, n)</td>
<td>CDLP</td>
<td>wAR</td>
</tr>
<tr>
<td>(4, 0.8)</td>
<td>5,755</td>
<td>5,748</td>
</tr>
<tr>
<td>(4, 1.0)</td>
<td>5,263</td>
<td>5,242</td>
</tr>
<tr>
<td>(4, 1.2)</td>
<td>5,056</td>
<td>5,080</td>
</tr>
<tr>
<td>(6, 0.8)</td>
<td>5,487</td>
<td>5,531</td>
</tr>
<tr>
<td>(6, 1.0)</td>
<td>5,087</td>
<td>5,127</td>
</tr>
<tr>
<td>(6, 1.2)</td>
<td>4,605</td>
<td>4,764</td>
</tr>
<tr>
<td>(6, 1.6)</td>
<td>3,824</td>
<td>4,101</td>
</tr>
<tr>
<td>(8, 0.8)</td>
<td>4,829</td>
<td>4,888</td>
</tr>
<tr>
<td>(8, 1.0)</td>
<td>4,343</td>
<td>4,434</td>
</tr>
<tr>
<td>(8, 1.2)</td>
<td>3,969</td>
<td>4,091</td>
</tr>
<tr>
<td>(8, 1.6)</td>
<td>3,684</td>
<td>3,579</td>
</tr>
</tbody>
</table>

avg. 2.28 3.02 2.94 1.33 2.34

Table 8: Comparison of the expected revenues for the hub and spoke test problems with stationary arrival rates.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Upper Bound</th>
<th>% Gap with CDLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N, n)</td>
<td>CDLP</td>
<td>wAR</td>
</tr>
<tr>
<td>(4, 0.8)</td>
<td>4,400</td>
<td>4,396</td>
</tr>
<tr>
<td>(4, 1.0)</td>
<td>4,138</td>
<td>4,053</td>
</tr>
<tr>
<td>(4, 1.2)</td>
<td>3,796</td>
<td>3,711</td>
</tr>
<tr>
<td>(4, 1.6)</td>
<td>3,100</td>
<td>3,057</td>
</tr>
<tr>
<td>(6, 0.8)</td>
<td>4,311</td>
<td>4,256</td>
</tr>
<tr>
<td>(6, 1.0)</td>
<td>4,015</td>
<td>3,900</td>
</tr>
<tr>
<td>(6, 1.2)</td>
<td>3,628</td>
<td>3,508</td>
</tr>
<tr>
<td>(6, 1.6)</td>
<td>2,855</td>
<td>2,769</td>
</tr>
<tr>
<td>(8, 0.8)</td>
<td>3,802</td>
<td>3,678</td>
</tr>
<tr>
<td>(8, 1.0)</td>
<td>3,440</td>
<td>3,308</td>
</tr>
<tr>
<td>(8, 1.2)</td>
<td>3,082</td>
<td>2,940</td>
</tr>
<tr>
<td>(8, 1.6)</td>
<td>2,475</td>
<td>2,364</td>
</tr>
</tbody>
</table>

avg. 2.75 3.08 2.80 1.60 3.43

Table 9: Comparison of the upper bounds for the hub and spoke test problems with non-stationary arrival rates.
Table 10: Comparison of the expected revenues for the hub and spoke test problems with non-stationary arrival rates.

<table>
<thead>
<tr>
<th>No. of spokes</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spokes</td>
<td>CDLP</td>
<td>wAR</td>
<td>wAR+</td>
<td>hAR</td>
</tr>
<tr>
<td>(N, \alpha)</td>
<td>CDLP</td>
<td>wAR</td>
<td>wAR+</td>
<td>hAR</td>
</tr>
<tr>
<td>(4, 0.8)</td>
<td>3,556</td>
<td>3,553</td>
<td>3,547</td>
<td>3,532</td>
</tr>
<tr>
<td></td>
<td>-0.08 \circ \times -0.26 \circ \times -0.13 \circ \times -0.33 \circ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4, 1.0)</td>
<td>3,290</td>
<td>3,301</td>
<td>3,294</td>
<td>3,296</td>
</tr>
<tr>
<td></td>
<td>0.34 \circ \times 0.10 \circ \times 1.11 \circ \times 1.19 \circ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4, 1.2)</td>
<td>3,050</td>
<td>3,084</td>
<td>3,085</td>
<td>3,076</td>
</tr>
<tr>
<td></td>
<td>1.09 \checkmark \times 1.15 \checkmark \times 0.85 \checkmark \times 1.93 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4, 1.6)</td>
<td>2,468</td>
<td>2,618</td>
<td>2,621</td>
<td>2,504</td>
</tr>
<tr>
<td></td>
<td>6.09 \checkmark \times 6.20 \checkmark \times 1.46 \checkmark \times 6.00 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6, 0.8)</td>
<td>3,399</td>
<td>3,397</td>
<td>3,394</td>
<td>3,432</td>
</tr>
<tr>
<td></td>
<td>-0.06 \circ \times -0.14 \circ \times 0.99 \checkmark \times 0.59 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6, 1.0)</td>
<td>3,134</td>
<td>3,151</td>
<td>3,148</td>
<td>3,160</td>
</tr>
<tr>
<td></td>
<td>0.52 \checkmark \times 0.45 \checkmark \times 0.83 \checkmark \times 1.23 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6, 1.2)</td>
<td>2,800</td>
<td>2,861</td>
<td>2,864</td>
<td>2,843</td>
</tr>
<tr>
<td></td>
<td>2.20 \checkmark \times 1.57 \checkmark \times 1.53 \checkmark \times 3.53 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6, 1.6)</td>
<td>2,154</td>
<td>2,409</td>
<td>2,409</td>
<td>2,271</td>
</tr>
<tr>
<td></td>
<td>11.42 \checkmark \times 11.83 \checkmark \times 5.42 \checkmark \times 12.21 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8, 0.8)</td>
<td>2,963</td>
<td>3,030</td>
<td>2,985</td>
<td>3,027</td>
</tr>
<tr>
<td></td>
<td>0.74 \checkmark \times 0.55 \checkmark \times 2.15 \checkmark \times 1.96 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8, 1.0)</td>
<td>2,642</td>
<td>2,733</td>
<td>2,707</td>
<td>2,710</td>
</tr>
<tr>
<td></td>
<td>2.44 \checkmark \times 2.57 \checkmark \times 2.56 \checkmark \times 2.29 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8, 1.2)</td>
<td>2,330</td>
<td>2,424</td>
<td>2,417</td>
<td>2,414</td>
</tr>
<tr>
<td></td>
<td>3.76 \checkmark \times 2.92 \checkmark \times 3.61 \checkmark \times 2.63 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8, 1.6)</td>
<td>1,810</td>
<td>1,906</td>
<td>2,016</td>
<td>2,027</td>
</tr>
<tr>
<td></td>
<td>11.36 \checkmark \times 11.99 \checkmark \times 8.19 \checkmark \times 10.29 \checkmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg.</td>
<td>3.80</td>
<td>3.30</td>
<td>3.22</td>
<td>2.38</td>
</tr>
</tbody>
</table>

Table 11: CPU seconds for the benchmark solution methods as a function of the number of spokes in the airline network and the number of time periods in the booking horizon.

<table>
<thead>
<tr>
<th>No. of spokes</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spokes</td>
<td>CDLP</td>
<td>wAR</td>
<td>wAR+</td>
<td>hAR</td>
</tr>
<tr>
<td>(N, \alpha)</td>
<td>CDLP</td>
<td>wAR</td>
<td>wAR+</td>
<td>hAR</td>
</tr>
<tr>
<td>(4, 0.8)</td>
<td>3.1</td>
<td>46</td>
<td>25</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>16</td>
<td>23</td>
<td>30</td>
</tr>
<tr>
<td>(4, 1.0)</td>
<td>0.8</td>
<td>135</td>
<td>59</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>16</td>
<td>23</td>
<td>30</td>
</tr>
<tr>
<td>(4, 1.2)</td>
<td>1.2</td>
<td>105</td>
<td>127</td>
<td>213</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>21</td>
<td>62</td>
<td>41</td>
</tr>
<tr>
<td>(4, 1.6)</td>
<td>1.9</td>
<td>191</td>
<td>335</td>
<td>407</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>31</td>
<td>104</td>
<td>59</td>
</tr>
<tr>
<td>avg.</td>
<td>3.80</td>
<td>3.30</td>
<td>3.22</td>
<td>2.38</td>
</tr>
</tbody>
</table>

7 Contribution

CDLP and the affine relaxation are two methods in the literature that give upper bounds on the value function for choice network revenue management. While CDLP is known to be tractable for the MNL model with disjoint consideration sets, 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. [7] 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 multiple segments and nested logit model of choice. 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


Demand. MIT Press, Cambridge, MA.


Appendix A: A continuous relaxation of the affine separation problem

We show that $\text{wAR}$ gives a tighter upper bound than a continuous relaxation of the $AF$ separation problem. We first describe the continuous relaxation and then show that $\text{wAR}$ gives a tighter upper bound than the continuous relaxation. We use the following result, which follows from Theorem 1 in [20]. Throughout, we assume that the choice probabilities are according to the MNL choice model and that the no-purchase preference weight is normalized to 1.

**Theorem 1.** Let $\mathcal{H} = \{\sum_S h_S P_j(S) | 0 \leq h_S \leq 1, \sum_S h_S = 1\}$ and let $\mathcal{Y} = \{(y_0, \{y_j | \forall j\}) | y_0 + \sum_j y_j = 1, y_j / w_j - y_0 \leq 0, \forall j, y_j, y_0 \geq 0\}$. We have $\mathcal{H} = \mathcal{Y}$.

Note that in Theorem 1 above, we can interpret $y_j$ as being the sales rate of product $j$. We now give an equivalent formulation of the $AF$ separation problem in terms of the sales rates. Recall, from constraints (11) that the $AF$ separation problem is to check if

$$\Pi_t^{RAF}(\gamma) = \max_{h,q} \left\{ \sum_j \lambda P_j(S) \left[ f_j - \sum_{i \in I_j, k=t+1}^{\tau} \gamma_{i,k} \right] - \sum_i \gamma_{i,t} q_i \right\}.$$  

Introducing variables $h_S$ to indicate if set $S$ is offered and $q_i$ to indicate if resource $i$ is open, and noting that we write $i \in I_S$ whenever there is a $j \in S$ with $i \in I_j$, the above maximization problem can be equivalently written as

$$\Pi_t^{RAF}(\gamma) = \max_{h,q} \left\{ \sum_S h_S \left\{ \sum_j \lambda P_j(S) \left[ f_j - \sum_{i \in I_j, k=t+1}^{\tau} \gamma_{i,k} \right] \right\} - \sum_i \gamma_{i,t} q_i \right\}$$

s.t. \[\left( \sum_S P_j(S) h_S \right) / w_j \leq q_i, \quad \forall i \in I_j, j\]

\[\sum_S h_S = 1\]

\[h_S \in \{0,1\}, q_i \in \{0,1\}.\]
We note that since $P_j(S) > 0$ for $j \in S$ and $P_j(S) = 0$ for $j \notin S$, constraints (44) ensure that $q_i = 1$ for all $i \in I_S$ with $h_S = 1$. Now observe that we can replace the integrality constraint on $h_S$ with $0 \leq h_S \leq 1$ without affecting the optimal objective function value. After doing this and by appealing to Theorem 1, we can obtain $\Pi_t^{RAF}(\gamma)$, through the following equivalent, sales-rate based, formulation:

$$\Pi_t^{RAF}(\gamma) = \max_{y, q} \sum_j \lambda \left[ f_j - \sum_{i \in I_S} \sum_{k=t+1}^{\tau} \gamma_{i,k} \right] y_j - \sum_i \gamma_{i,t} q_i$$  \hspace{1cm} (45)$$

$$\text{s.t. } \begin{align*}
y_j / w_j &\leq q_i, \quad \forall i \in I_j, j \\
y_0 + \sum_j y_j &= 1 \\
y_j / w_j - y_0 &\leq 0 \quad \forall j \\
y_j, y_0 &\geq 0, q_i \in \{0, 1\}. \hspace{1cm} (46)\
\end{align*}$$

The above, sales-rate based formulation motivates our continuous relaxation of the AF separation problem, where we relax the integrality requirement on $q_i$ by replacing it with the constraint $0 \leq q_i \leq 1$ and solve an LP. In other words, if we let $\Pi_t^{AF_{CR}}(\gamma)$ denote the optimal objective function value of the continuous relaxation of (45)–(49), we replace the AF constraint $\beta_t \geq \Pi_t^{RAF}(\gamma)$ with $\beta_t \geq \Pi_t^{AF_{CR}}(\gamma)$. It is clear that by separating over the continuous relaxation, we get an upper bound that is weaker than $AF$. We show below that the resulting upper bound is weaker than the $wAR$ bound as well. We first take the dual of the continuous relaxation of (45)–(49), which after some simplifications is

$$\Pi_t^{AF_{CR}}(\gamma) = \min_{\pi, z} \sum_i \left[ \sum_{j \in J_i} \pi_{i,j} - \gamma_{i,t} \right] + \sum_j z_j$$

$$\text{s.t. } \begin{align*}
\sum_{i \in I_j} \pi_{i,j} / w_j + \sum_{j'} z_{j'} + z_j / w_j &\geq \lambda \left[ f_j - \sum_{i \in I_j} \sum_{k=t+1}^{\tau} \gamma_{i,k} \right] \quad \forall j \\
\sum_{j \in J_i} \pi_{i,j} - \gamma_{i,t} &\geq 0 \quad \forall i \\
\pi_{i,j}, z_j &\geq 0.
\end{align*} \hspace{1cm} (47)$$

Folding the dual back into $RAF$ by following the same steps as in §4.3, we obtain the following LP:

$$V^{AF_{CR}} = \min_{\beta, \gamma, \pi, z} \sum_t \beta_t + \sum_i \sum_{j \in J_i} \gamma_{i,t} r_i$$

$$\text{s.t. } \begin{align*}
\beta_t - \sum_i \sum_{j \in J_i} \pi_{i,j,t} - \sum_{j} \gamma_{i,t} - \sum_{j} z_{j,t} &\geq 0 \quad \forall t \\
\sum_{i \in J_t} \pi_{i,j,t} / w_j + \sum_{j'} z_{j',t} + z_{j,t} / w_j &+ \lambda \sum_{k=t+1}^{\tau} \sum_{i \in I_j} \gamma_{i,k} \geq \lambda f_j \quad \forall j, t \\
\sum_{i \in J_t} \pi_{i,j,t} - \gamma_{i,t} &\geq 0 \quad \forall i, t \\
\gamma_{i,t}, z_{j,t}, \pi_{i,j,t} &\geq 0.
\end{align*} \hspace{1cm} (AF_{CR})$$
Taking the dual of $AF_{CR}$ and after some simplifications we get

$$V^{dAF_{CR}} = \max_{x,\rho} \sum_t \sum_j \lambda_j x_{j,t}$$

s.t.

$$x_0,t + \frac{\lambda}{w_j} \sum_{s=1}^{t-1} \sum_{j \in J_s} x_{j,s} + \sum_j x_{j,t} - \rho_{i,t} \leq r^1_i \quad \forall i, t$$

$$(dAF_{CR}) \quad x_0,t + \sum_j x_{j,t} = 1 \quad \forall t$$

$$\frac{x_{j,t}}{w_j} - 1 + \rho_{i,t} \leq 0, \quad \forall i, j \in J_i, t$$

$$\frac{x_{j,t}}{w_j} - x_{0,t} \leq 0, \quad \forall j, t$$

$$x_{0,t}, x_{j,t}, \rho_{i,t} \geq 0.$$  

We have the following lemma.

**Lemma 4.** $V^{wAR} \leq V^{AF_{CR}}$

**Proof**

Note that the second constraint in $dAF_{CR}$ implies that $x_{0,t} \leq 1$. Therefore, $\frac{x_{j,t}}{w_j} - x_{0,t} + \rho_{i,t} \geq \frac{x_{j,t}}{w_j} - 1 + \rho_{i,t}$. Since $\rho_{i,t} \geq 0$, we also have $\frac{x_{j,t}}{w_j} - x_{0,t} + \rho_{i,t} \geq \frac{x_{j,t}}{w_j} - x_{0,t}$. Comparing the constraints of $dAF_{CR}$ with those of $dwAR$, we see that any feasible solution to $dwAR$ is also feasible to $dAF_{CR}$. The result follows by noting that $dAF_{CR}$ and $dwAR$ have the same objective function.

---

**Appendix B: Alternative weak affine relaxation**

In this section we describe a complementary, tractable relaxation that uses a different bound than weak affine relaxation. We refer to this method as alternative weak affine relaxation ($awAR$). We show that it also gives an upper bound on the value function and that this bound is weaker than the $AF$ bound. We describe $awAR$ for the single-segment MNL model, but all of the development goes through when there are multiple segments by following the steps in §5.

$awAR$ is based on the following identity:

**Proposition 6.** $\gamma_{S,t} \left( \sum_{j \in S} w_j \right) \geq \sum_{i,j} \left[ \mathbb{I}_{i \in I_s} + \mathbb{I}_{j \in S} - 1 \right] \gamma_{i,t} w_j + \sum_i \left[ 1 - \mathbb{I}_{i \in I_s} \right] \left( \sum_{j \in J_i} \gamma_{i,t} w_j \right)$.

**Proof**

The right hand side can be written as

$$\gamma_{S,t} \left( \sum_{j \in S} w_j \right) - \sum_{\{i,j \mid i \in I_s, j = 0, 1 \mid i \in s = 0\}} \gamma_{i,t} w_j + \sum_i \left[ 1 - \mathbb{I}_{i \in I_s} \right] \left( \sum_{j \in J_i} \gamma_{i,t} w_j \right).$$
So it is enough to show that
\[ -\sum_{\{i,j|1_{\{i\in\mathcal{I}\}}=0,1_{\{j\in\mathcal{S}\}}=0\}} \gamma_{i,t}w_j + \sum_i \left[1 - \mathbb{1}_{\{i\in\mathcal{I}\}}\right] \left(\sum_{j\in\mathcal{J}_i} \gamma_{i,t}w_j\right) \leq 0. \]

Examining the term,
\[ -\sum_{\{i,j|1_{\{i\in\mathcal{I}\}}=0,1_{\{j\in\mathcal{S}\}}=0\}} \gamma_{i,t}w_j + \sum_i \left[1 - \mathbb{1}_{\{i\in\mathcal{I}\}}\right] \left(\sum_{j\in\mathcal{J}_i} \gamma_{i,t}w_j\right) = -\sum_{\{i,j|1_{\{i\in\mathcal{I}\}}=0,1_{\{j\in\mathcal{S}\}}=0\}} \gamma_{i,t}w_j + \sum_{i\notin\mathcal{I}_S} \gamma_{i,t} \left(\sum_{j\in\mathcal{J}_i} w_j\right) \leq 0, \]

where the last equality follows from the fact that for all \(i \in \{i|j \notin \mathcal{I}\}\) we cannot have a \(j \in \mathcal{S}\) with \(i \in \mathcal{I}_j\).

Proposition 6 implies that \(\Pi_t^{AF}(\beta, \gamma) \leq \Pi_t^{awAR}(\beta, \gamma)\) where
\[ \Pi_t^{awAR}(\beta, \gamma) = \max_S -\gamma_{S,t} - \sum_{i,j} \left[1_{\{i\in\mathcal{I}\}} + \mathbb{1}_{\{j\in\mathcal{S}\}} - 1\right] \gamma_{i,t}w_j - \sum_i \left[1 - \mathbb{1}_{\{i\in\mathcal{I}\}}\right] \left(\sum_{j\in\mathcal{J}_i} \gamma_{i,t}w_j\right) \]
\[ -\sum_{j\in\mathcal{S}} w_j + \beta_t + \lambda \left(\sum_{i\in\mathcal{I}_t} \sum_{k=t+1}^{t+1} \gamma_{i,k} - f_j\right) \geq 0. \]

We replace the right hand side of constraint (16) with \(\Pi_t^{awAR}(\beta, \gamma)\) and solve the LP
\[ V^{awAR} = \min_{\beta, \gamma} \sum_t \beta_t + \sum_t \sum_i \gamma_{i,t}r_i^1 \]
\(awAR\) s.t. \(\beta_t \geq \Pi_t^{awAR}(\beta, \gamma) \quad \forall t\)
\(\gamma_{i,t} \geq 0.\)

Lemma 1 now implies that \(V^{AF} \leq V^{awAR}\).

Finally, we note that the \(awAR\) bound can be computed in a tractable manner since the separation problem (50) can be solved as the following integer program with a totally-unimodular
constraint matrix:

\[
\Pi^\text{w-AR}_t(\beta, \gamma) = \max_{q_i,t} \left[ \beta_t + \lambda \left( \sum_{i \in I} \sum_{k=t+1}^{\gamma_{i,k}} - f_j \right) \right] u_j - \sum_{i,j} [u_j + q_i - 1] \gamma_{i,t} w_j
\]

- \sum_{i} [1 - q_i] \gamma_{i,t} \left( \sum_{j \in J_i} w_j \right) - \sum_{i} \gamma_{i,t} q_i

s.t. (22) – (24).

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