View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by University of Essex Research Repository

Choice-Based Demand Management and Vehicle Routing in E-fulfilment

Xinan Yang[*], Arne K. Strauss[†], Christine Currie[‡], Richard Eglese[§]

1 July 2013

Abstract

Attended home delivery services face the challenge of providing narrow delivery time slots to ensure customer satisfaction, whilst keeping the significant delivery cost under control. To that end, the firm can try to influence customers when they are booking their delivery time slot so as to steer them towards choosing slots that are expected to result in cost-effective schedules. We estimate a multinomial logit customer choice model from historic booking data and demonstrate that this can be calibrated well on a genuine e-grocer data set. We propose dynamic pricing policies based on this choice model to determine which and how much incentive (discount or charge) to offer for which time slot at the time a customer intends to make a booking. A crucial role in these dynamic pricing problems is played by the delivery cost, which is also estimated dynamically. We show in a simulation study based on real data that anticipating the likely future delivery cost of an additional order in a given location can lead to significantly increased profit as compared to current industry practice.

1 Introduction

For many firms, the delivery of goods or services to a customer at an appointed time constitutes a crucial part of their business. Examples include e-grocers and e-tailers such as Sainsburys.com, Tesco.com, Ocado, parcel delivery (e.g. FedEx, VelocityExpress), repair or servicemen (e.g. BT), furniture delivery and many more. The growing demand for such services can be illustrated by the British online grocery sector which grew by 17.2% in 2011, with sector sales being forecast to rise by 79% over the next five years relative to 2012's levels [Mintel, 2012]. They all face the trade-off between the substantial fulfillment costs that are typically associated with making such deliveries on the one hand, and customer satisfaction particularly with regard to narrow delivery time windows on the other. The importance of getting this trade-off right was dramatically demonstrated by business failures such as Webvan (bankrupt in 2001) or Publix Direct (shut down 2003). Other companies learned from these failures and successfully provide delivery services with larger delivery time windows, using appropriate scheduling and vehicle routing software and, at the same time, often have become very good at collecting valuable customer information that is used for customer segmentation and targeted marketing. The sales process typically consists of collecting orders including delivery time requests (subject to available capacity) until a certain cut-off time, and subsequent planning of the delivery schedule using appropriate routing software. Clearly, the requested delivery times have a potentially large impact on scheduling and routing efficiency: for example, the given delivery time specifications might require much longer routes than if demand were geographically clustered for each time window. Furthermore, some delivery

[*]Lancaster University Management School, Lancaster LA1 4YX, United Kingdom
[†]Warwick Business School, University of Warwick, Coventry CV4 7AL, United Kingdom. Corresponding author: arne.strauss@wbs.ac.uk.
[‡]Mathematics, University of Southampton, Southampton SO17 1BJ, United Kingdom
[§]Lancaster University Management School, Lancaster LA1 4YX, United Kingdom

time slots might be much more popular than others, so that a large vehicle fleet would be needed for peak times while many vehicles would be idle in off-peak times.

A reservoir of great profit-generating potential is to manage demand (in the form of online bookings) by setting incentives in such a way that customers are steered towards selecting time slots that maximize the firm's expected profits. Such incentives can take a variety of forms: the size of the delivery charge for different time slots, rewards such as discounts or shopping points for choosing unpopular slots, or even indications of environmental impact as e.g. used by Ocado, who label certain time slots with a "green van" symbol representing delivery slots where an order would save fuel as compared to slots without this symbol. Note that the objective of maximizing profit implies knowledge of expected fulfilment costs, which links demand management to vehicle routing. The customers' choices of delivery times directly impact on delivery costs so that steering their choices could cause significant profit increases.

We consider the following problem: service requests for a fixed delivery day arrive at random over a finite time horizon until a cut-off time before the delivery day. A schedule/routing plan is subsequently needed for the delivery of the collected orders. The firm has a number of homogeneous delivery vehicles with a certain capacity. Each request for a specific delivery day consists of customer information (including post code); a revenue; and a capacity consumption corresponding to the number of standard-sized transport boxes that are required to accommodate the goods that this customer has ordered. The firm needs to check which delivery time slots are feasible for this post code given the orders collected to date (in the sense that there is capacity for delivery at that time), and has to decide in real-time which time slots receive which monetary incentive. In particular, the incentive can be positive or negative, representing either a cost or a discount for the customer. The resulting time slot/discount combinations are subsequently displayed to the customer who makes a decision on which time slot to request. The firm is then committed to delivering at the requested time.

Our main contribution lies in the formulation of pricing policies that are based on a more advanced customer choice model (namely, a multinomial logit model) than has been considered in this context before; moreover, our proposed policies also take dynamically estimated delivery cost into account and hence combine models from the traditional vehicle routing and the revenue management literature. We show in a simulation study that including the impact of future expected orders in the estimation of delivery cost produces higher profits "foresight policy" than only using orders accepted-to-date in this estimation "hindsight policy". The practical applicability of the proposed methodology including estimation of the entire demand model is demonstrated on real data that we obtained from our industrial collaboration partner in the online grocery sector.

The paper is organized as follows: in Section 2 we review the relevant literature, followed by a formal statement of the home delivery pricing problem in Section 3. Then we describe the methodology of how to fit a demand model of the realistic booking process that is using by the e-retailer, and conclude with estimations of Poisson Process for arrival time and an MNL choice model for customer selection in Section 4. The dynamic pricing policy that we propose is discussed in Section 5, and we compare it with benchmark policies in a simulation study based on real data in Section 6. Finally, we summarize and draw conclusions in Section 7.

2 Literature Review

While there is a rich literature for both vehicle routing and revenue management (RM), little research has been carried out at the interface between the two areas. Agatz et al. [2008] give a qualitative introduction into the use of RM techniques in managing demand for delivery services, when estimates of the costs of fulfilling the orders are taken into account. In their comparison of the use of RM in the airline sector and its potential use in home deliveries, they segment potential solution methodologies into the four categories of Table 1, corresponding to a similar categorization used in RM.

The static methods of differentiated slotting and differentiated pricing are fed by forecast data and can be used to solve problems such as deciding on the number, length and selection of delivery

Table 1: Classification of Demand Management.

	Time slot allocation	Pricing
Static (off-line, forecast-based)	Differentiated slotting	Differentiated pricing
Dynamic (real-time, order-based)	Dynamic slotting	Dynamic pricing

slots on offer or setting off-peak discounts to smooth out demand. In reality, static methods should be updated as soon as new forecasts are available; and in this sense, they become "dynamic". The opportunity cost is the expected loss in future revenue that will be incurred if the request is accepted, i.e. the revenue we will miss out on if we use the capacity now rather than saving it for later requests. In principle, the methods of dynamic slotting and dynamic pricing will re-evaluate opportunity cost estimates at each request arrival in order to make a decision; nonetheless they can (and should) also exploit knowledge on expected future demand using appropriate forecasts. We have found little research in the area of differentiated pricing of delivery slots, but structure the literature review around the remaining categories of Table 1.

Differentiated slotting: Agatz et al. [2011] consider the problem of selecting which time slots to offer in each post code of the delivery region in a case study with Dutch Internet grocer Albert.nl, given service requirements and average weekly demands for each post code. The objective is to minimize expected delivery costs while meeting the service requirements. Note that this time slot allocation takes place before receiving any orders (as opposed to real-time management of time slots in dynamic slotting). The resulting time slot schedule could then serve as a starting point for subsequent real-time adjustments. The optimization requires estimates of the routing costs from the time slot schedule; however, since calculating a full routing model is generally computationally too expensive, they instead propose two options: either to use the continuous approximation method, which represents demand by continuous functions and assesses system-wide costs by aggregating over local cost estimates, or to explicitly model routing decisions but on a more aggregate level than a full vehicle routing problem, using the seed-based scheme put forward in Fischer and Jaikumar [1981].

Their approach is interesting in that the problem decomposes by shift, but this is due to rather unrealistic demand assumptions:

1. The expected weekly demand for each post code is known and independent of the set of offered time slots

2. The expected weekly demand for each post code is divided evenly over the set of offered time slots.

The main motivation of demand management in this context is that many customers are not flexible in time so that changes in the offered time slot menu can have significant consequences in demand. Hence, the assumption of independence seems somewhat counter-intuitive. Likewise, the second assumption means that all time slots are equally popular. While this can be achieved in some cases (such as for Albert.nl) by means of appropriate differentiated pricing, it is not clear why this would be reasonable to assume in general. The work offers no real-time control. As future research directions, the authors point towards a better understanding of customer behavior, the aforementioned real-time controls, and use of incentives.

Dynamic slotting: Campbell and Savelsbergh [2005] were among the first to propose a framework for this problem: in their model, the company can decide whether to accept or to reject service requests, and, furthermore, in which time slot to choose to deliver, so as to maximize expected profits. Customer choice behavior is not taken into account and the fulfillment costs for a new request are estimated by evaluating the estimated incremental cost of inserting the new demand into the tentative delivery routes that encompass all accepted requests. The company's decision is to accept or to deny service requests. This can be problematic for industries where service denials create considerable "bad will" among the customers, which leads to lost sales and

potential brand damage if done on a large scale. In general, persuasive methods (such as dynamic pricing) seem preferable to coercive methods (not offering slots).

Dynamic pricing: The work the closest to ours is Campbell and Savelsbergh [2006] who expand on their previous work Campbell and Savelsbergh [2005] by adding a model of customer choice. The firm is assumed to receive a series of orders of known size and revenue, and by setting discounts is able to encourage customers to choose delivery slots corresponding to lower delivery costs. Customer preferences for delivery slots are assumed to be known in advance but by offering a discount in a particular slot the firm can increase a customer's probability of selecting this slot by an amount proportional to the discount. After a customer has selected a particular slot, the firm updates the set of delivery routes to accommodate the new request and these are then used for estimation of delivery costs when the next order request arrives.

We follow a similar approach to Campbell and Savelsbergh [2006] but make some significant improvements. First, in the Foresight Policy described in Section 5.1.2, we use stochastic information about future requests to estimate the expected delivery cost, rather than computing the incentives based only on the requests that have already been accepted and the request under consideration. If it is possible to construct reasonable forecasts, one would expect that this should significantly improve the optimisation, especially as the routing cost estimates at the beginning of the sales horizon are based on only a few arrivals, and therefore are unlikely to reflect the final routing costs. This is underlined by the results presented by Campbell and Savelsbergh [2006] who show that offering incentives can be critical even in the early stages. Second, we use the multinomial logit (MNL) model to describe customer behavior, an area of improvement that they highlight in their article.

Choice Modeling and RM: In an example from the airline sector, Vulcano et al. [2010] show that by incorporating customer-choice behavior into RM, in the form of an MNL model to describe customer behavior, revenues can be increased by 1-5% over RM methods where this choice behavior is not taken into account. While coming from a different sector, this result does emphasize the need for accurate modeling of customer choice behavior.

We use the MNL model to describe the behavior of customers who are choosing between different delivery time slots. The MNL model is a random utility, discrete choice model where the decision-maker is assumed to choose the option that maximizes his or her utility (see e.g. Ben-Akiva and Lerman [1994], Train [2009]). In such random utility models, the utility that a customer from a particular segment l places on option j, can be written as

$$U_j^l = u_j^l + \epsilon_j^l,$$

where u_j^l is termed the representative utility and is dependent on choice factors that are known to the person deriving the models. The error term, ϵ_j^l is defined as the difference between the true utility and the part of the utility that can be described by u_j^l. For an MNL model, the ϵ_j^l are assumed to be independent and identically distributed, following a Gumbel distribution. The MNL model is well-known to have the property of the independence of irrelevant alternatives (IIA), i.e. the relative odds of choosing an option i over a different option k are independent of the attributes of any of the alternative options that are available. This has some interesting consequences, as discussed in Chapter 3 of Train [2009]. In the application considered here, where the time slots that customers are choosing between are allowed to overlap, the assumption of independence is not strictly valid. While the IIA property is not a feature of other choice models such as the nested logit model, these more complex choice models have other problems associated with them. In particular, as discussed by Vulcano et al. [2010] and Vulcano et al. [2012], the estimation of the nested logit model is considerably harder than that of the MNL because of the higher number of parameters and the fact that the log-likelihood function is no longer globally concave.

Customer choice modeling is used by Asdemir et al. [2009] in determining how to dynamically price delivery fees for time slots in home delivery. They use the MNL model to describe choice behavior based on a RM formulation using dynamic programming. Their objective is to maximize expected profits subject to known costs (including delivery costs). Consequently, the entire delivery

routing aspect is ignored. For most practical situations, this seems to be problematic since one would expect that the delivery fees should affect demand for time slots and therefore the routing schedule and its associated costs. In particular, they assume independence of geographical regions so that the model can be solved for each single, small area. However, if we also want to consider routing costs, we should plan routes incorporating demand in neighboring areas as well.

The use of the MNL model in RM is widespread (e.g. Vulcano et al. [2010]; Suh and Aydin [2011]; Meissner et al. [2013]; Vulcano et al. [2012]) and although the applications come almost exclusively from the airline sector, the choice behavior follows a similar structure to the choice of delivery slots. The standard problem encountered in the literature considers the optimal pricing of multiple, substitutable flights offered by the same carrier, between the same origin and destination airports. Customers will choose between the flights or opt not to purchase a ticket based upon their own preferences and the prices being charged. A considerable amount of research has focused on the estimation of the parameters in the MNL model. This is a difficult problem in the sale of airline tickets because of the inability to distinguish between a period with no arrivals and a period with an arrival and no purchases. Maximum likelihood methods that take account of this missing data such as the Expectation-Maximization algorithm used by Vulcano et al. [2012] are needed. Due to the requirement to register a delivery location before starting the ordering process, this issue does not exist in the case of home deliveries and consequently the estimation is much more straightforward. Standard maximum likelihood methods can be used to fit the MNL model, as described in e.g. in McFadden [1974] or Train [2009], and this can be done separately from the fitting of the arrival process.

3 Delivery Pricing Problem

Our model of the delivery pricing problem is motivated by real-life home delivery applications where a customer reveals his identity to the firm by logging into the website before a delivery can be booked. Then the customer can select a delivery day from a booking horizon that extends for a fixed number of days into the future. We model these customer arrival events for a specific delivery day by a time-dependent Poisson process. Once a delivery day has been selected, the firm has to decide on both availability and pricing of slots before displaying them to the customer, who then chooses to pick one of these slots (or not to book) according to a MNL model. We first state the formal problem formulation that we investigate in this study, before providing more background on the demand model in Section 4.

Given the arrival of a customer interested in a specified delivery day, our aim is to decide dynamically how to set delivery charges for time slots on this day so as to steer the customer's choice to a slot that is likely to be cheap to serve, so that overall we intend to maximize profits after fulfillment costs for this delivery day. The pricing policy for a specified delivery day can be formulated based on a stochastic dynamic program, where the state $x_t = [x_{tas}]_{a,s}$ is a vector with components x_{tas} that represent the number of orders accepted in time slot s for area a until time t in the booking horizon. In the following, we omit the time index from x since it will be clear from the context. We consider a discretized booking horizon with T booking periods, each sufficiently small such that the probability of having more than one arrival in a period is negligible. The final time period T denotes the cut-off time after which no further bookings are taken. The stages of the dynamic program are the time periods $t \in [1, T]$. Let $V_t(x)$ denote the value function at stage t and state x; it represents the maximum profit obtainable from the sales process from time t until the cut-off time T. Furthermore, we let $C(x)$ represent an oracle that returns the minimum cost solution to the vehicle routing problem with time windows for the set of orders x given a fixed fleet of vehicles with known capacities. If there is no feasible solution for a given x, then $C(x) := \infty$. The value function after cut-off is then

$$V_{T+1}(x) = -C(x) \quad \forall\, x \in \mathcal{X}, \tag{1}$$

where $\mathcal{X}$ denotes the set of all states that allow a feasible delivery schedule. The set $F_a(x) := \{s : C(x + \mathbf{1}_{as}) < \infty\}$ contains all feasible time slots for area a into which order (a, s) can be feasibly

inserted given orders on the books x, where $\mathbf{1}_{as}$ is the unit vector with 1 in the (a,s)th position. There are a finite number of customer segments l for the fixed delivery day under consideration, each with an arrival rate of λ_l. For a given customer arrival of segment l, μ_{la} is the probability that this arrival will be for area a. Given an arrival, the probability that this customer chooses slot s when the firm offers the vector of delivery charges $\vec{d_a}$ in area a is denoted by $P^l_{s,F_a(x)}(\vec{d_a})$.

For a clearer exposition, we assume that all orders require the same known number of totes, say one, and that the profit of the order (before distribution cost), r, is likewise known. The assumption on the number of totes can be easily relaxed, and the assumption of a known profit at the point in time where a pricing decision needs to be made is violated in practice since a customer can often choose a time slot before shopping. Instead, we could e.g. use an estimated revenue based on the average historic order value either for that delivery area or for the individual customer, and multiply this revenue by the average profit margin before distribution cost.

With this notation, the dynamic programming recursion at stage $t \in [1,T]$ is

$$\begin{aligned} V_t(x) &= \max_{\vec{d}} \sum_{l,a} \lambda_l \mu_{la} \sum_{s \in F_a(x)} P^l_{s,F_a(x)}(\vec{d_a}) \Big[r + d_{as} + V_{t+1}(x + \mathbf{1}_{as}) \Big] + \\ &\qquad \Big[1 - \sum_{l,a} \lambda_l \mu_{la} \sum_{s \in F_a(x)} P^l_{s,F_a(x)}(\vec{d_a}) \Big] V_{t+1}(x) \\ &= \max_{\vec{d}} \sum_{l,a} \lambda_l \mu_{la} \sum_{s \in F_a(x)} P^l_{s,F_a(x)}(\vec{d_a}) \Big[r + d_{as} - \Big(V_{t+1}(x) - V_{t+1}(x + \mathbf{1}_{as}) \Big) \Big] + V_{t+1}(x), \\ &\forall\, x \in \mathcal{X}. \end{aligned} \tag{2}$$

The dynamic program (1–2) is intractable due to the large state space $\mathcal{X}$ and the fact that the computation of $C(x)$ alone is intractable since it requires solving a large vehicle routing problem with time windows. However, it offers some insights into the structure of a good policy: for a given request of segment l at time t for area a, it shows that the time slot pricing decision is a trade off between the immediate profit before distribution $(r + d_s)$ and expected opportunity cost $(V_{t+1}(x) - V_{t+1}(x + \mathbf{1}_{as}))$ arising from collecting an order (a,s) at time t. Therefore, given the arrival of a segment l customer at area a, a policy of the form

$$\vec{d}^*_a = \arg\max_{\vec{d_a}} \sum_{s \in F_a(x)} P^l_{s,F_a(x)}(\vec{d_a}) \Big[r + d_s - \mathcal{O}_{xtas} \Big], \tag{3}$$

should perform well if the expression $\mathcal{O}_{xtas}$ is close to the true opportunity cost $(V_{t+1}(x) - V_{t+1}(x + \mathbf{1}_{as}))$. In Section 5 we present an approximation of the opportunity cost in the form of order insertion cost estimates. In the following section, we discuss the demand model and how to estimate its parameters from real online shopping data.

4 Demand Model

This section is organized as follows: we first discuss the assumptions that we have to make in order to carry out the study in Section 4.1, and then characterize customer segmentation in Section 4.2. Estimation methodologies of the Poisson process and the MNL model are discussed in Section 4.3 and 4.4 respectively, each tested on artificial booking data with known parameters to evaluate their quality. Finally, we present statistics of the estimation process based on real booking data.

4.1 Assumptions and Justifications

Our model is based on a number of assumptions:

- A customer can be identified when a delivery booking request is being made and the order size is known when the delivery time selection is made. Although in practice the latter might not be the case as the delivery selection is often required to be made before products are

selected, e.g. at Tesco and Waitrose, we still can infer an expected revenue e.g. from past purchases of this customer; if the application is an e-grocer, it is likely that such estimates will be of good quality.

- Customer requests for a fixed delivery day arrive over time according to a time-dependent Poisson process. This arrival process is independent of time slot pricing and availability; we assume that this information is not known to the customer at this point. In practice, customers may be able to see the slot availabilities/pricing of several days. The assumption of choices being independent of alternative delivery days would hence be violated; however, both we and the partner firm felt that it is a reasonable assumption to make in order to obtain a choice model that can be calibrated in acceptable time. Removing the assumption of independence between delivery days would be an interesting feature for future research.

- We can observe both time slot selections and decisions not to choose any time slot on the fixed delivery day. This assumption is not restrictive in the home delivery business as orders are usually taken only over the company's own website where customers are required to identify themselves before being able to book a delivery. Moreover, we also assume that the firm records what time slot prices and availabilities were offered when a customer selects a delivery day, regardless of whether or not the customer subsequently chooses a time slot. This is indeed the case for our partner firm.

- Given the arrival of a customer, the probability that he selects a specific time slot on a fixed delivery day is independent from the time in the booking horizon when the order is placed. Time dependence is only captured via the arrival process.

- Customers are myopic, i.e. they do not anticipate future pricing/availability decisions. This seems reasonable since most grocery orders are booked within only three days prior to delivery, and the charges are usually small relative to the order value. In fact, the companies usually impose a limit on delivery charges because of competition.

- We ignore cancelation and re-scheduling of orders. This assumption seemed not overly restrictive since our real-life data set showed that most deliveries are booked only shortly in advance and are therefore usually fulfilled as ordered.

4.2 Customer Segments

The definition of customer segments is a major design decision for the choice model, since the demand model assumes that the choice behavior of all customers within a given segment is homogeneous. Remember that we assumed that customers only select time slots of a specific chosen day, rather than also choosing time slots among various delivery days. Hence, as far as the choice model is concerned, we can create customer segments as subsets of those customers who select a specific delivery day; however, we need to keep in mind that therefore a customer viewing delivery options on various delivery days will appear to the model as arrivals from several segments.

We investigated various definitions on the basis of a real data set from our industrial collaboration partner, ranging from aggregating all customers into a single group, to aggregation according to postcode or time slot preference. Eventually we chose to aggregate all customers into one segment for a fixed delivery day, and we assume that every customer considers all time slots on the fixed delivery day. The customer's decision is made based on availability and delivery charges. The no-booking option is always available and can either represent that the customer did not book at all, or that the customer chose to book delivery for a different day.

The reason for segmenting customers only by delivery day is that such a segmentation allows us to attribute each delivery request to exactly one segment and we also have sufficiently dense data available to calibrate a choice model for each segment. An intuitive alternative segmentation could be to include time preference for the delivery time slot, e.g. "Monday Morning", "Monday Afternoon" and "Monday Evening". However, this could result in two customers ordering simultaneously for the same post code area and for the same delivery day, having different prices

and availabilities quoted since their estimated choice preferences (and hence estimated time slot selection probabilities) are likely to differ, which in turn impacts on the dynamic pricing optimization. Price discrimination based on characteristics of individual customers is regarded as unacceptable by many customer, as e.g. Amazon experienced in 2000 when experimenting with customer-tailored DVD prices, see Ward [2000]. Our partner firm accordingly rejected models that may result in different quotes in the same time slot, same location and same order value because of potential damage to their reputation. For the same reason, analysis of an individual customer is undesirable, even if sufficient data would be available for every customer.

4.3 Customer Arrival Process

Given the definition of a customer segment, we need to model the arrival process for customers of a certain segment. A segment l arrival for delivery day d is defined as the event that a segment l customer selects delivery day d on the web site. Note that this does not imply that this customer books a time slot on this day; this arrival event takes place before any time slots are displayed, so the customer may book on a different day or not book at all. Therefore, there are no slot-specific arrival rates; slot-specific expected demand will be modeled in the next section with a multinomial logit choice model. As mentioned in the previous section, our model excludes cancelation or re-scheduling of orders. Given real data that does include occurrences of customers re-scheduling arrivals, we accordingly only considered the final decision process on the day that led to the actual delivery (or no delivery booking at all); all potential other arrival events on previous days in the booking horizon were ignored in the estimation of the arrival intensity function.

Based on our knowledge of past booking histories, we intend to identify a partition of the booking horizon with arrival intensities that are constant within each time period, such that the arising cumulative intensity function is likely to be a good estimate of the (unknown) cumulative intensity function $\Lambda_l(t) := \int_0^t \lambda_l(r)dr$ of the non-homogeneous continuous Poisson arrival process $\lambda_l(t)$. Note that although the arrival rate shall be constant for each time period by construction, the resulting estimated cumulative intensity function is still time-dependent since the length of time periods differs. Once an estimate of the cumulative intensity function has been found, we can use inversion to define time periods for specified expected arrivals. In particular, let us look at how to identify time periods in a way such that each time period has the same expected number of arrivals, and moreover such that the probability of more than one arrival in a given time period is smaller than a pre-specified value. The latter is not only required by the dynamic program (1–2), but is also important for the estimation of the choice model parameters as described in the next section.

There are various parametric and non-parametric models that could be used to estimate this function; the following non-parametric procedure by Leemis [1991] seems appropriate for our purpose because it does not require arbitrary decisions from the modeler. Furthermore, linear interpolation allows inversion to be used to identify arrival times (otherwise there may be ties). Let $(0, \tau]$ denote the booking horizon leading up to a specific cut-off day at time τ. We refer to all data of arrivals and choice processes associated with a fixed segment over the entire booking horizon as booking histories. Assuming that each booking horizon leading up to e.g. a Monday is based on the same demand model, we use all available H booking histories in the arrival intensity estimation. This approach can be justified by the fact that the booking patterns for each weekday in our data were fairly stable, even bank holidays did not appear to dramatically change the booking pattern. In our data set, each booking horizon has a length of 22 days, and we have $H = 26$ histories available. Leemis' estimator $\bar{\Lambda}_l(t)$ is the piecewise linear function that is obtained by linear interpolation of the superposition of the historic arrival events: a very simple example is depicted in Figure 1. We refer to Leemis [1991] for details on the implementation of this method.

Having estimated the cumulative intensity function from the booking histories, we can now determine time periods with a uniform arrival intensity λ_l by specifying the desired number of time periods T. The time periods will typically have different lengths, reflecting the time-dependent arrival intensity. The arrival intensity λ_l is given by the total expected number of arrivals divided

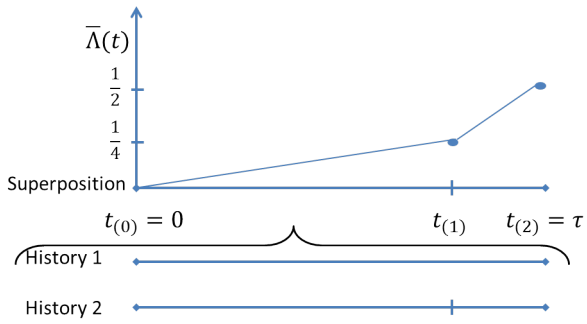

Figure 1: Nonparametric estimate of the cumulative arrival intensity function. Example with $H = 2$ histories, containing $n_1 = 0$ and $n_2 = 1$ observations, respectively.

by the number of time periods: $\lambda_l := \bar{\Lambda}_l(\tau)/T$. This defines a uniform grid on the interval $[0, \bar{\Lambda}_l(\tau)]$ with T subintervals of length λ_l. We can use the inverse of our estimate of the cumulative intensity function to define time periods $(t_i, t_{i+1}]$ with $t_i := \bar{\Lambda}_l^{-1}(i\lambda_l)$, $i \in \{0, \ldots, T\}$, that therefore have lengths such that the expected number of arrivals within each time period is equal to λ_l. Note that the inverse is well-defined since $\bar{\Lambda}_l$ is monotone and piecewise linear. Again we refer to Leemis [1991] for details on the inverse.

To obtain time periods that are sufficiently small such that the probability of having more than one arrival is negligible, we can use a result by Çinlar (1975) for a Poisson process $\lambda(t)$: the probability $P_{(a,b]}(X > 1)$ of more than one arrival in $(a, b]$ is

$$P_{(a,b]}(X > 1) = 1 - \exp(-\int_a^b \lambda(t)dt)[1 + \int_a^b \lambda(t)dt].$$

We can approximate the integral terms using $\int_a^b \lambda_l(t)dt \approx \bar{\Lambda}_l(b) - \bar{\Lambda}_l(a)$ in the formula for $P_{(a,b]}(X > 1)$ above. Note that by construction $\bar{\Lambda}_l(t_{i+1}) - \bar{\Lambda}_l(t_i) = \lambda_l$ for all $i \in \{1, \ldots, T-1\}$. Hence $P_{(t_i,t_{i+1}]}(X > 1) = 1 - \exp(-\lambda_l)(1 + \lambda_l)$ for all i. This gives us a rule for selecting the number of time periods T: since $\lambda_l := \bar{\Lambda}(\tau)/T$, define T such that

$$1 - \exp(-\frac{\bar{\Lambda}_l(\tau)}{T})(1 + \frac{\bar{\Lambda}_l(\tau)}{T}) \approx \zeta,$$

where $\zeta > 0$ represents the desired (approximate) probability of having more than one arrival per time period.

The choice of ζ will need to take into account that very small ζ will produce very small time periods, which in turn brings noise from the data into the estimated density function. On the other hand, ζ should not be chosen too large since this would violate the assumption for the probability of having more than one arrival per time period, i.e. ζ, being negligible. Based on a simulation study with arrivals that follow known arrival rates similar to those in real data, we found that $\zeta = 0.2$ provided a good trade-off.

4.4 Multinomial Logit Choice Model

Let us consider the time slot selection process that follows the delivery day selection. More precisely, we need to estimate the choice probability of a customer from each segment for selection of any particular time slot, given their respective availability and incentive. Time slot availability is considered pre-determined by whether or not the firm can feasibly insert the current order in that time slot given the current delivery schedules based on all accepted orders to date. We elaborate on schedule and pricing optimization in §5.

For a fixed delivery day, there is a fixed number of S (potentially partly overlapping) time slots $s \in \{1, \ldots, S\}$; we denote the delivery charge (or discount) for time slot s by d_s. Each customer

of a segment l is assumed to consider all delivery time slots for the chosen day, and the customer's utility perception is homogeneous across the segment population in that he attributes a utility of $U_s^l := u_s^l + \epsilon_s^l$ to time slot s. Customer behavior is assumed to follow an MNL model such that u_s^l is a known deterministic part that we can influence via the delivery charges, and ϵ_s^l is an i.i.d. random variable that follows a Gumbel distribution with zero mean and variance $\mu^2\pi^2/6$ for $\pi = 3.14\ldots$ and $\mu > 0$. The deterministic part is assumed to be a linear function of the time slots' attributes, namely $u_s^l := \beta_0^l + \beta_s^l + \beta_d^l d_s$, where β_0^l is a base utility across all options, β_s^l is the utility attributed to the slot itself, and β_d^l is the sensitivity of utility to the delivery charge. Let $u_0^l = 0$ denote the utility for the no-purchase option for segment l, normalized to zero. We can estimate β^l for each segment separately since we are able identify which segment a customer belongs to. We can assume that $\mu = 1$ as this parameter is absorbed in the scaling of the vector β. For a given segment l arrival in the booking horizon, we have given the set $F \subset \{1, \ldots, S\}$ of feasible time slots that are available for selection.

Under the assumption that customers are utility maximizers, the probability that time slot s is chosen given that time slots $s \in F$ are available at delivery charges $\vec{d} := [d_s]_{s\in F}$, can be shown to be the following:

$$P_s^l(\vec{d}) = \frac{\exp\left(\beta_0^l + \beta_s^l + \beta_d^l d_s\right)}{\sum_{k\in F} \exp\left(\beta_0^l + \beta_k^l + \beta_d^l d_k\right) + 1},$$

and the probability that no time slot for this delivery day is chosen is given by

$$P_0^l(\vec{d}) = \frac{1}{\sum_{k\in F} \exp\left(\beta_0^l + \beta_k^l + \beta_d^l d_k\right) + 1}.$$

We proceed with the estimation of the attribute sensitivity parameters β^l by maximizing the log-likelihood function with respect to β^l. Let $\mathcal{P}_h$ denote the set of time periods in history h where a time slot was selected, $\bar{\mathcal{P}}_h$ the set of time periods where there was an arrival but no time slot selection on the fixed delivery day, $\bar{\bar{\mathcal{P}}}_h$ the set of time periods without arrivals, F_{ht} the set of feasible time slots offered at delivery charges $\vec{d}_{ht}$ in history h and time period t, and $s(h,t)$ the index of the time slot s that was chosen in history h in time period t. The data that we were given by the firm contains records of which delivery days customers looked at and on which one they booked delivery, if at all. Hence we can observe directly when customers look at the fixed delivery day but then choose not to book on this day, and all the sets $\mathcal{P}_h$, $\bar{\mathcal{P}}_h$ and $\bar{\bar{\mathcal{P}}}_h$ are observable. This is a great advantage over e.g. airline data where $\bar{\mathcal{P}}_h$ can typically not be observed, and thus one requires methods to estimate this set.
The complete log-likelihood function is given by

$$\begin{aligned}
&\sum_h \sum_{t\in\mathcal{P}_h} \left[\log\lambda_l + \beta_0^l + \beta_{s(h,t)}^l + \beta_d^l d_{s(h,t)} - \log(\sum_{k\in F_{ht}} \exp(\beta_0^l + \beta_k^l + \beta_d^l d_k) + 1)\right] \\
&+ \sum_h \sum_{t\in\bar{\mathcal{P}}_h} (\log\lambda_l - \log(\sum_{k\in F_{ht}} \exp(\beta_0^l + \beta_k^l + \beta_d^l d_k) + 1) \\
&+ \sum_h \sum_{t\in\bar{\bar{\mathcal{P}}}_h} \log(1-\lambda_l).
\end{aligned}$$

We discussed the estimation of the arrival rate λ_l for a segment l above; remember that we constructed the time grids for each segment in a way such that λ_l is the same for each time period. Note that the log-likelihood function is separable in β^l and λ_l, so that we only need to maximize:

$$\begin{aligned}
\mathcal{L}(\beta^l) = &\sum_h \sum_{t\in\mathcal{P}_h} \left[\beta_0^l + \beta_{s(h,t)}^l + \beta_d^l d_{s(h,t)} - \log(\sum_{k\in F_{ht}} \exp(\beta_0^l + \beta_k^l + \beta_d^l d_k) + 1)\right] \\
&+ \sum_h \sum_{t\in\bar{\mathcal{P}}_h} (-\log(\sum_{k\in F_{ht}} \exp(\beta_0^l + \beta_k^l + \beta_d^l d_k) + 1).
\end{aligned}$$

Table 2: Runtime of MNL estimation procedure. Runtime in minutes and average percentage differences between true and recovered parameters for various sample sizes.

Sample Size	Runtime (min)	Average Difference (%)
20,000	36.9	18
40,000	69.6	13
80,000	121.5	10
160,000	226.5	9

Train [2003], p. 65, states that this function is concave, hence we can tackle it with standard nonlinear programming solvers to find the global maximizer. E.g., we can use Quasi-Newton methods; they only require the first-order derivatives but achieve superlinear convergence rates.

Example 4.1 *We test our ability to recover the choice parameters depending on sample size and the associated required runtime in a simulation study. Again, in our choice of number of time slots, sensitivity parameters and range of sample choice observation we intend to be close to real-life data, so that the example provides a feel for how many samples are required for achieving a certain accuracy. We defined $S = 27$ partly overlapping one hour time slots, each with a pre-defined sensitivity parameter β_s, specified the base parameter β_0 and delivery charge sensitivity β_d. The estimation is done for each customer segment individually, hence we focus on a single segment in this simulation. We disregard the point in time when the booking is made in the simulation since we assume that the time slot choice behavior (given a selected delivery day) is independent from the time of booking.*

To create the sample events, we generate randomly a set of available time slots by considering a time slot to be open with a probability of 0.7, and otherwise to be closed. Next, for each open time slot we sample a delivery charge from the discrete uniform distribution over three price points as reported in Table 7 in the appendix. Finally, the customer's decision is drawn from the probability distribution defined by the MNL model for the given β parameters over all available time slots and sampled price points. We repeat this process for sample sizes ranging from 20,000 to 160,000. Again, such sample sizes are realistic in the real-life application as we concluded from our real data set.

We use the attribute of time slot 23 with its associated sensitivity β_{23} as the reference variable since this attribute featured most frequently in the observed choice sets ($> 90\%$), re-define the β parameters (we refer to Vulcano et al. [2010] for details on this standard transformation) and attempt to recover the β parameters from the sampled observations using MATLAB *R2012a to maximize the log-likelihood function. The maximization procedure used is* `fminunc` *in the large-scale mode with user-supplied gradients, and the runtimes are based on solving it on an Intel Xeon CPU X5650 @ 2.67GHz system.*

The results and original parameters are reported in Tables 8–11 in the Appendix. It is interesting to observe that for the large samples of 80,000 or 160,000 observations, most parameters can be recovered with great accuracy; particularly the price sensitivity β_d is recovered well. This is encouraging since our real data set features around 160,000 observations in the booking histories for any given weekday. We also report the quasi t-statistics, computed as the ratio between the estimated value of the parameter and the asymptotic standard error (ASE), which has the critical value of ± 1.96 at the 0.05 significance level [Newey and McFadden, 1994]. The result shows that we can reject the null hypothesis that the true parameter value is zero (meaning that the decision maker is indifferent to this attribute) for all parameters for the 160,000 and 80,000 samples, and even for 40,000 observations there is only one case β_{10} where it cannot be rejected, and in this case the original β_{10} is indeed zero.

Runtime increases linearly with sample sizes as can be seen from Table 2; doubling the sample size increases runtime by a factor of approximately 1.8. The runtime seems to be acceptable given that this estimation process runs offline.

5 Dynamic Pricing Policies

The main issue at stake is the question of how to control the booking process so as to maximize the firm's profit; more precisely, how to control time slot booking incentives so as to influence the customers' expected booking decisions so that the final routing schedules are likely to be cost-effective, and to potentially earn auxiliary revenue from the delivery charges.

Under the premise that a time slot is made available for a given delivery request if and only if that order can be feasibly inserted into the current routing schedule, the only decision that we can make to control the booking process is the pricing of the time slots. As mentioned beforehand, we can also offer negative prices as incentives, which would then be discounts coming off the order's total value. We mean by the term "pricing policy" a decision rule used to decide on these incentives given a customer arrival for a particular delivery day.

The dynamic program (2) motivates an approximation $\mathcal{O}_{xtas}$ of the opportunity cost that not only depends on the cost of inserting an order in slot s for area a based on the orders x collected until time t, but also based on the *future* orders we expect to receive. This insight is reflected in our proposed so-called hindsight and foresight policies that we discuss in the following.

5.1 Approximations to opportunity cost

By dynamic policy we mean a way of controlling delivery charges as we observe bookings over time. Examples in industry are Ocado and Tesco Groceries who appear to change individual time slot prices dynamically.

Before we can optimize pricing of time slots for a given order and given delivery day, we firstly need to check whether insertion of this order in any specific time slot allows a feasible schedule to be built, and secondly what cost would be affiliated with inserting this order in each time slot (the so-called *insertion cost*). For a given set of accepted orders and a current order j^* under consideration, we use the following heuristic approach to check feasibility as described in Algorithm 1, which is loosely based on the approach chosen by Campbell and Savelsbergh [2006]. We label the time required to drive from the location of order i to the location of the next order $i+1$ by $t_{i,i+1}$. We maintain a pool of at most 10 schedules for the orders accepted so far. This includes the best schedule available when the last order arrived (empty at the outset), and the 9 others are randomly generated to increase the likelihood of finding feasible schedules. Naturally, the final schedule will differ significantly from a schedule built earlier in the booking process since it is constructed based on partial information on the accepted orders.

The cost of inserting a given order into a time slot is computed by examining the cost of inserting the order into each schedule $\sigma \in \Sigma$, and selecting the cheapest as the insertion cost estimate. We calculate these cost estimates as a function of the distance traveled, so $f(d_{i,i+1})$ represents the fuel cost for traveling the distance $d_{i,i+1}$ from location of order i to that of the next order $i+1$ on the route. We do not take the driver's salaries, vehicle maintenance or other costs into account. Under this assumption, the cost of inserting order j between an order $i-1$ and another order i for a given schedule $\sigma \in \Sigma$ can be calculated by

$$C^s_\sigma \leftarrow f(d_{i-1,j}) + f(d_{j,i}) - f(d_{i-1,i}) + C(\sigma) - C^*,$$

where $C(\sigma)$ and C^* are the total delivery cost of schedule σ and of the cheapest schedule in Σ, respectively. The latter two terms need to be included in the calculation of C^s (as remarked by Campbell and Savelsbergh [2006]) because otherwise an order might be inserted in a schedule where $f(d_{i-1,j}) + f(d_{j,i}) - f(d_{i-1,i})$ is small but that is otherwise very expensive. More details on how to define the cost function $f(\cdot)$ are given in Section 6.3.

Campbell and Savelsbergh [2006] propose a dynamic pricing policy where the opportunity cost is approximated by an estimate of the cost of inserting the order into a given time slot for a given current order state x as seen in Algorithm 1. Recall that dynamic policies are likely to perform well if they are in the form of (3) with a good opportunity cost approximation, meaning that the impact of future expected deliveries should be taken into account as well when estimating the insertion cost. Otherwise, we might make potentially bad decisions very early in the booking

Algorithm 1 Procedure to find feasible schedules given a set of orders U.

Initialization: set of schedules $\Sigma \leftarrow \{\text{best previous schedule}\}$.
for $\sigma = 1 : 9$ **do**
  Initialize candidate list of yet unscheduled orders, U, to contain all accepted orders to date.
  while $U \neq \emptyset$ **do**
    Pick an order $j \in U$ at random, check feasibility of inserting it in schedule σ its time slot s $(\text{begin}^s, \text{end}^s)$ between already inserted orders $i-1$ and i by computing the earliest and latest insertion times e_j and l_j:

$$e_j = \max(e_{i-1} + \text{service time} + t_{i-1,j}, \text{begin}^s),$$

$$l_j = \min(l_i - \text{service time} - t_{j,i}, \text{end}^s).$$

    The insertion is feasible if and only if $e_j \leq l_j$.
    if Insertion is feasible **then**
      Insert order j in schedule σ. Remove order j from candidate list U.
    else
      Record that schedule σ is infeasible. Break from the while loop.
    end if
  end while
  If schedule σ is feasible, add it to the set Σ.
end for

horizon. Since Campbell and Savelsbergh [2006] emphasize that "the use of incentives can be critical even in the early stages of building a delivery schedule", it seems important to address this issue. To that end, we propose two approaches to constructing a decision policy based on our demand model: the Hindsight Policy and the Foresight Policy, thus labeled to reflect the omission/inclusion of information on the future demand in the optimization process.

5.1.1 Choice-based: Hindsight Policy

At a given point in time during the booking horizon, the Hindsight Policy only takes the information into account that has been revealed in the sales process so far. It maintains a set of schedules only for the accepted orders and updates them once a new order is accepted. In other words, no information about future demand is included. These schedules are used to determine the increase in delivery cost when accepting a new order, but since no knowledge on future demand is taken into account, it does not tell us the cost of having potentially displaced a more valuable order arriving later. In this regard, the Hindsight Approach is similar to the method proposed by Campbell and Savelsbergh [2006] except that they use a simpler choice model where the reduction of the delivery charge for a particular time slot results in a linear increase of selection probability for this time slot and an equal reduction of selection probability across all alternative time slots.

As a rule, we make all time slots available that are feasible to the vehicle routing problem; this is motivated by the fact that our partner firm wants to serve customers as long as there is a feasible way of doing so; customer retention is very important in this line of business. The main purpose of a delivery charge is to incentivize customers to choose time slots that allow cheaper delivery schedules and to secure high-value orders.

The lowest insertion costs C^s need to be computed for each time slot s since they are required in the subsequent optimization of delivery charges. Every time that an order (a, s) for a delivery to area a and in time slot s arrives we compute the lowest insertion costs for all time slots, and then carry out the pricing optimization. We maximize the immediate profit from (a, s) by deciding on the vector $\vec{d} \in \mathbb{R}^S$, where S is the number of time slots. Depending on the sign of the decision variable d_s, it represents a delivery charge or a discount to be deducted from the order value. The

online decision problem that we need to solve is

$$\max \sum_{s \in F_a(x)} (r + d_s - C^s) P^l_{s,F_a(x)}(\vec{d}) \tag{4}$$
$$\vec{d_a} \in \mathbb{R}^S,$$

where $F_a(x)$ denotes the set of feasible time slots in area a for given set of accepted orders x (determined by the routing schedule optimization).

The problem (4) is not quasi-concave in price as shown by Hanson and Martin [1996], but Dong et al. [2009] show that the problem is concave in the purchase probabilities. We apply their result to (4) in what follows, where we omit the segment index for the sake of a clearer presentation. Note that it follows from the MNL model that

$$\frac{P_s(\vec{d})}{P_0(\vec{d})} = \exp(\beta_0 + \beta_d d_s + \beta_s) \quad \forall\, s;$$

remember that the utility of the no-selection option u_0 is normalized to zero. It follows that

$$d_s = -\frac{1}{\beta_d}(\beta_0 + \beta_s + \ln P_0 - \ln P_s) \quad \forall\, s,$$

where $\vec{P} = [P_0, P_1, \ldots, P_S]$ is the vector of selection probabilities for the walk-away option and each time slot option, and $\sum_{s=0}^{S} P_s = 1$. Substituting this into (4) results in the following concave optimization problem (assuming $\beta_d < 0$):

$$\max \sum_{s \in F_a(x)} [-\frac{1}{\beta_d}(\beta_0 + \beta_s + \ln P_0 - \ln P_s) + r - C^s] P_s$$
$$\sum_{s=0}^{S} P_s = 1$$
$$\vec{P} \in [0,1]^{S+1}.$$

Theorem 5.1 *The optimal solution to (4) is given by*

$$d^*_s = C^s - r - \frac{m}{\beta_d} \quad \forall\, s \in F_a(x),$$

where m is the unique solution of

$$(m-1)\exp(m) = \sum_{s \in F_a(x)} \exp(\beta_0 + \beta_s + \beta_d(C^s - r)). \tag{5}$$

Proof Proof. Theorem 1 in Dong et al. [2009].

Therefore, we can solve the online pricing problem efficiently by e.g. using a standard Newton root search to find m in (5). We restrict the incentive range in our numerical study to $[-10, 10]$, i.e. if d^*_s is outside this interval, we project d^*_s onto the interval's boundary. The choice of this interval is motivated by the fact that delivery charges of our industry partner are always less than £10, and promotional discounts on the order of about £10 are sometimes given e.g. to new customers. Upper and lower limits should be specified to avoid occasional very large charges or discounts that could not realistically be offered since customers will likely not accept large variations as fair pricing.

5.1.2 Choice-based: Foresight Policy

A shortcoming of the previous approach is the implicit assumption that the insertion cost estimates are good approximations of the true opportunity costs of accepting orders. This assumption is questionable particularly in the beginning of the booking horizon because the estimation only takes observed orders into account. The Foresight Policy attempts to improve on that by creating and maintaining two pools of schedules Σ_H and Σ_F, one for the accepted-to-date orders only, and the other consisting of the most recent ten final schedules for the same delivery weekday. When a new order arrives, we attempt to insert it into each schedule in each of the two pools so as to obtain insertion cost estimates. For the pool Σ_H we compute the insertion cost C_H^s for each time slot s as described in the previous section.

For the pool Σ_F, we try to insert the new order into every schedule. If it is feasible to insert the new order, then we keep the corresponding insertion cost, otherwise we set the insertion cost to a big value (say £10). We define the foresight insertion cost estimate C_F^s as the average of all the insertion costs over the schedules in Σ_F. If the new order appeared regularly in those histories, the foresight insertion cost would be zero for those slots in which it had been delivered. On the other hand, for some very popular time slots which are highly likely to be full in those schedule histories, the foresight insertion cost will be fairly high.

The final insertion cost is estimated as a linear combination of C_H^s and C_F^s. For orders arriving early in the booking horizon, the foresight insertion cost is likely to be a good estimate, whereas for the late arrivals the hindsight insertion cost is more accurate. Hence we increase the weight w_j of C_H^s continually as orders come in by defining $w_j := j/J$, assuming that we have accepted $j-1$ orders so far, and defining J as the maximum expected number of bookings over the entire booking horizon. Thus the insertion cost for the jth order and time slot s is defined by

$$C_j^s = w_j C_H^s + (1 - w_j) C_F^s.$$

We solve the pricing problem (4) with these new insertion cost estimates using Theorem 5.1 as before.

6 Numerical Results: Real Booking Data

To test our demand estimation methodology as well as the dynamic policies, we apply them to the data set of our industrial collaboration partner. In this section, we first describe the data set

6.1 Description of the Data

We obtained anonymized booking data from a major e-grocer in the United Kingdom, ranging over 6 months from beginning of June to the end of November 2011. All bookings were made over the firm's website, and the customers had to be registered and logged into their account when booking a delivery time slot. A delivery could be booked up to 22 days in advance, though most bookings occurred within three days to the delivery day. For a given weekday, say Monday, we have 26 booking histories over the entire corresponding booking horizon available.

All deliveries were served by a single depot. The firm collects relevant data via their website: customer arrivals are recorded with customer number, order number, arrival time, advisory period and delivery day selected for order (if any). This arrival event is linked with detailed offer set information including when which slots were offered at which price, and which (if any) slot was selected by the customer. Finally, relevant order details included order number, postcode, customer number, total number of totes, and delivery day and slot. All these entries are linked via system event numbers. Our data does not include the revenue of any orders though the firm does record this as well. We excluded cancelation and re-scheduling from our analysis.

Table 3: Data Sample. Data extract for a specific order by a specific customer who re-schedules the delivery time slot several times.

AdviseStart	AdviseEnd	Selected	Creation
15/05/2011	17/05/2011		14/05/2011, 20:52
22/05/2011	23/05/2011		14/05/2011, 20:52
24/05/2011	25/05/2011	24/05/2011, 22:00–23:00	14/05/2011, 20:53
26/05/2011	27/05/2011		14/05/2011, 20:53
15/05/2011	17/05/2011		14/05/2011, 20:55
24/05/2011	25/05/2011	25/05/2011, 22:00–23:00	14/05/2011, 20:55
25/05/2011	27/05/2011		24/05/2011, 14:12
01/06/2011	02/06/2011	02/06/2011, 22:00–23:00	24/05/2011, 14:12
26/05/2011	28/05/2011		25/05/2011, 14:01
31/05/2011	01/06/2011	01/06/2011, 22:00–23:00	25/05/2011, 14:02

Table 4: Estimation results. Arrival rates and time grids based on the e-grocer data set using $\zeta = 0.2$. "Total arrivals" contains all arrivals over all 26 booking histories.

	Mon	Tue	Wed	Thu	Fri	Sat
Total arrivals	149,803	150,249	173,087	163,991	165,771	150,836
T	6,990	7,010	8,076	7,651	7,734	7,038
λ_l	0.824	0.824	0.824	0.824	0.824	0.824

6.2 Demand Estimation from E-Grocer Data

Example 6.1 *We explain how we define a customer arrival using the example of a real booking record for a specific order. All required information can be obtained if the firm sells over its own web site, and it is available in our real-world data set. More specifically, our data contains the so-called AdviseStart and AdviseEnd dates that represent the span of delivery days that the customer looked at. If the customer selects a time slot, it is recorded in the Selected column. The firm allows customers to change their selection both within the current session as well as later on up to a certain cut-off time. Consider the data extract in Table 3: this is one customer who re-scheduled repeatedly the delivery time slot for the same order and eventually settled for the first of June. How should we count arrivals here for the delivery day* 26 May*, say? This date appears three times in the booking history within the advisory period, but as we ignore re-scheduling in our model, we only look at the decision process that occurred on the final creation day 25/05/2011. On that day, the customer considers delivery on* 26–28 May*,* 31 May *and* 1 June*. Therefore, for each of these days we would record an arrival for the corresponding segment; in particular, for* 26 May *we record an arrival occurring at 14:01 on 25/05/2011. Note that this is a rather unusual example since the vast majority of customers do not re-schedule their slot.*

According to this definition of customer arrivals, there are around 6,000 to 7,000 arrivals in each booking history for a given weekday. We chose the probability of having more than one arrival per time periods as $\zeta = 0.2$, which corresponds to an arrival rate of $\lambda_l \approx 0.824$. The number of time periods that achieve this rate are reported in Table 4.

Based on these time grids, we estimated the MNL parameters from the data. The β parameters for Monday deliveries are reported in Table 5. There are 27 (overlapping) time slots, and we chose time slot 23 as the reference point so that the parameter is eliminated. Based on the quasi t-statistics we can reject the null hypothesis that the true parameter value is zero for each parameter. When comparing time slot preference parameters with the price preference parameter β_p, we remark that there is a difference in scale since x_s is on a scale from 0 to 1, whereas the prices in the data ranged from £0 to £7. Based on the t-statistics, we can again reject the null hypothesis that the true parameter value is zero for all parameters.

Table 5: MNL parameters estimated from real sales data. ASE: asymptotic standard error.

Slot ID	Monday Delivery		
	β	ASE	t-stat
1	-0.8230	0.0955	-8.6178
2	-0.7436	0.0596	-12.4765
3	-0.5746	0.0509	-11.2888
4	-0.3181	0.0415	-7.6651
5	0.1529	0.0034	44.9706
6	0.1897	0.0319	5.9467
7	0.7656	0.0264	29.0000
8	0.9941	0.0237	41.9451
9	0.4561	0.0255	17.8863
10	0.9091	0.0226	40.2257
11	0.1340	0.0272	4.9265
12	-0.2514	0.0341	-7.3724
13	-1.2908	0.0522	-24.7280
14	-0.3500	0.0317	-11.0410
15	-0.6213	0.0341	-18.2199
16	-0.3435	0.0308	-11.1526
17	-0.5251	0.0326	-16.1074
18	-0.1118	0.0291	-3.8419
19	-0.5093	0.0330	-15.4333
20	0.2316	0.0263	8.8061
21	-0.2854	0.0293	-9.7406
22	-0.3950	0.0296	-13.3446
23	0.0000	n/a	n/a
24	0.6395	0.0227	28.1718
25	-1.1516	0.0374	-30.7914
26	0.3912	0.0255	15.3412
27	-1.1656	0.0428	-27.2336
β_0	-2.8618	0.0222	-128.9099
β_d	-0.0880	0.0033	-26.6788

6.3 Evaluation of Policies

As a benchmark that represents industry practice of various online grocers, we use two static price policies that do not require any optimization at all:

- Single static price: we simply define a fixed delivery charge across all time slots and all delivery areas. For example, Boots (pharmacy-led health and beauty company) charges £3.95 for a named weekday up to 14 days in advance. In general, the delivery charges are often around £3 to £5, so we experiment with fixed prices at £3, £4 and £5.
- Order-value-based static price: we define a fixed delivery charge depending on the order volume, namely £3 for delivery of groceries worth £50 or more, and £5 charge for smaller orders (this also corresponds to current industry practice, e.g. Waitrose).

We test the performance of the various pricing policies with regard to profit in simulations based on the real data provided by our industry partner. We simulate customer arrivals according to the rate and time grid that were estimated for Monday deliveries as described above. The location of the customer is subsequently sampled according to the proportional arrival-by-location table that we extracted from the data. Next, we sample the order size in terms of profit before distribution by sampling the number of totes requested from the normal distribution, where mean and standard deviation are obtained from historical bookings in the corresponding location. We multiply the resulting number of totes with the average revenue per tote and the profit margin that we set at 30%. This percentage was estimated simply by the ratio of gross profits over gross revenues as reported on the firm's balance sheet, and company representatives felt that it is an adequate rule of thumb.

The respective pricing policy is used to obtain a vector of delivery prices, based on which the customer's time slot selection is sampled from the corresponding distribution given by the MNL model. The MNL parameters are estimated from Monday delivery data as discussed above. If the customer selects a time slot, the corresponding profit is recorded. We assume orders are not being canceled. This process is repeated for each of the $T = 6990$ time periods until the cut-off time is reached. We generate 40 sample booking histories and use their average generated profit as a measure of the quality of the respective pricing policy.

The number of vans is varied in the simulations between 100, 110, 120 and 150 to observe the impact of various degrees of capacity tightness. The homogeneous van capacity, service and reloading times, average speed and fuel cost that were used in the simulations were provided by our industrial partner in the understanding that they remain confidential. Our real data set contained post code sector information that we used to approximate the driving time $t_{i,i+1}$ and cost $f(d_{i,i+1})$ between two orders i and $i+1$. To do this, we approximated the road distance between two locations by multiplying the distance as the crow flies by a constant factor γ. If two locations are in one district, the factor was set to $\gamma = 1.35/(1000 * 1.6)$, and $\gamma = 1.25/(1000 * 1.6)$ otherwise. This reflects that the straight point-to-point distance is a better approximation of the road distance if the locations are further apart. Hence, $f(d_{i,i+1}) = \gamma(\text{lineDistance})(\text{costPerMile})$.

We consider the typical situation of all goods being stored in a central warehouse, from which all orders are delivered. We simplify the model by assuming that there is only one warehouse, and that all vans have only one compartment. In reality, grocery home delivery vans have different compartments for frozen, chilled and non-chilled totes.

Table 6 reports the mean profit results of the various pricing policies, along with the average number of orders accepted, the average total cost of the resulting schedules and the percentage gap to the order-value-based static pricing policy (VS). The dynamic pricing policies both outperform the static policies in most cases. In particular, the foresight policy improves on the VS policy consistently by on average 3.8%. The latter is an important result as two-tier static price policies such as the one that we use here as a benchmark are often found in industry. As expected, the fixed price policies do not perform well, particularly under tight capacity limits, since they do not even favor larger orders over smaller ones.

It becomes particularly interesting when investigating the hindsight policy: for the scenario with only 100 vans, it is 3.5% worse than VS. This is due to the fact that the number of vans in

Table 6: Policy performance. "#Deliv" is the average number of deliveries under the respective policy, "TotalCost" the average total delivery cost, "MeanCost" is TotalCost/#Deliv, "Total-Profit" the average profit after distribution, "StdDev" is the standard deviation of profits, and "Gap" is the percentage gap to the total profit achieved by policy "VS", which stands for order-**V**alue based **S**tatic pricing policy.

#Vans	Policy	#OrdersTaken	TotalCost	MeanCost	Total Profit	Stddev	Gap (%)
100	VS	2,953	7,609	2.58	105,743	269.4	0.0
	$d_s = 3$	2,957	7,681	2.60	105,123	254.5	-0.6
	$d_s = 4$	2,880	7,550	2.62	105,493	250.4	-0.2
	$d_s = 5$	2,792	7,446	2.67	104,817	336.1	-0.9
	Hindsight	3,183	8,059	2.53	102,125	282.5	-3.5
	Foresight	3,052	7,941	2.63	109,139	300.6	3.1
110	VS	3,042	8,061	2.65	108,334	252.0	0.0
	$d_s = 3$	3,063	8,018	2.62	109,225	258.4	0.8
	$d_s = 4$	2,950	7,800	2.64	107,784	308.6	-0.5
	$d_s = 5$	2,844	7,614	2.68	106,398	316.1	-1.8
	Hindsight	3,390	8,709	2.57	110,680	215.5	2.1
	Foresight	3,165	8,356	2.64	112,215	298.8	3.5
120	VS	3,101	8,222	2.65	110,687	312.3	0.0
	$d_s = 3$	3,091	8,121	2.63	110,039	352.3	-0.6
	$d_s = 4$	2,997	7,886	2.63	109,309	343.7	-1.3
	$d_s = 5$	2,874	7,778	2.71	107,584	308.9	-2.9
	Hindsight	3,486	8,821	2.53	113,504	236.5	2.5
	Foresight	3,375	8,754	2.60	113,661	333.9	2.6
150	VS	3,120	8,204	2.63	110,910	396.2	0.0
	$d_s = 3$	3,145	8,209	2.61	111,673	301.2	0.7
	$d_s = 4$	3,026	7,964	2.63	110,171	340.7	-0.7
	$d_s = 5$	2,883	7,708	2.67	108,031	323.9	-2.7
	Hindsight	3,511	8,942	2.55	116,016	302.3	4.4
	Foresight	3,530	8,947	2.53	118,192	343.0	6.2

this scenario severely limits the number of orders that can be feasibly accepted, so it becomes very difficult to find feasible slots later in the booking horizon if many bookings have been accepted early on. Figure 2 shows a box plot of all delivery prices used in the simulations for the 100 vans example. The prices tend to be dramatically lower than what the foresight policy recommends; in fact, most of the time the hindsight policy offers discounts so as to increase the probability that a given customer books a time slot. This can also be seen from Figure 4 that shows how the average delivery charge changes over time for the two respective policies: the hindsight policy starts already with discounts from early in the time horizon since it does not anticipate the high future demand (relative to available capacity). Therefore, towards the end of the booking horizon, the hindsight policy must offer even more discounts because all popular time slots are filled, and it becomes very difficult to move customers to the remaining few unpopular slots. The foresight policy however does anticipate future demand and hence starts off with positive charges, and only towards the end offers discounts to fill remaining unpopular slots. Figures 3 and 5 illustrate that in the situation of having plenty of capacity, both policies behave in a similar way as one would expect. Still, even in the scenario of 150 vans, foresight produces about 2% more profit than hindsight. This underlines the importance of basing the expected insertion cost not only on orders hitherto accepted, particularly when demand is larger than delivery capacity. The strong performance of the foresight policy versus the static order-value-based pricing policy also demonstrates the benefit of using customer choice modeling in time slot pricing, rather than only basing the decision on order value.

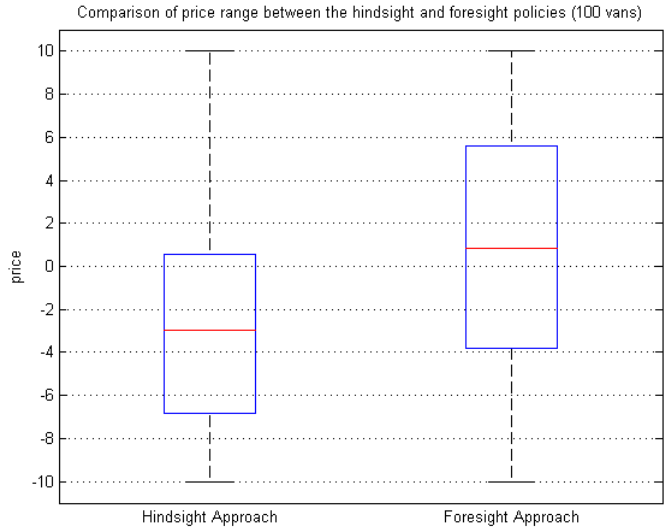

Figure 2: Box plot of delivery price range for the 100 vans example.

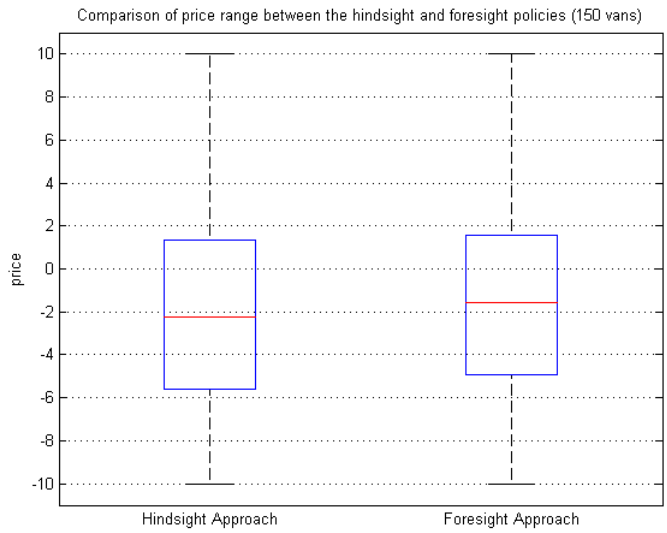

Figure 3: Box plot of delivery price range for the 150 vans example.

Figure 4: Average of delivery prices over intervals with 20 time periods for the 100 vans example, with beginning of the time horizon at $t = 0$.

Figure 5: Average of delivery prices over intervals with 20 time periods for the 150 vans example, with beginning of the time horizon at $t = 0$.

6.4 Practical Aspects and Managerial Insights

For practical implementation, computational effort needs to be reasonable relative to whether the task is carried out online or offline. The estimation of parameters of the arrival process and MNL model can be done offline. Computational efforts regarding the arrival process is negligible, but the MNL estimation takes approximately one hour with around 210,000 sample observations for each customer segment using MATLAB on an Intel Xeon CPU X5650 2.67GHz machine. However, since it is computed offline, it is not a practical limitation.

The evaluation of feasibility of inserting an order and the estimation of insertion cost however needs to be done online, and the optimization of dynamic pricing policies as well. With Theorem 5.1 the solution of problem (4) is easy; however the insertion cost calculation is not so straightforward. In order to dynamically estimate the insertion cost, we have to keep a sample schedule pool of 10 delivery schedules consisting of all of the orders that have been accepted so far. For every order arrival we evaluate the best possible delivery cost for every time slot by inserting the considered order into every feasible place in all of the sample schedules. This calculation dominates the optimization of slot-prices in the online algorithm which takes around 97% percent of running time.

When a customer selects a delivery day in practice, the system needs to display delivery slot prices and availabilities virtually immediately. In our simulation, the calculation of feasibility and optimal prices takes on average 0.18 seconds per arrival, which seems reasonable. Nevertheless, as a potential improvement, parallel computing can be applied in managing a larger pool of sample schedules in order to give more accurate estimations of insertion cost in practice, as the insertion step is independent across sample schedules/vans. Also, the firm could reduce the size of the insertion problem by using decomposition heuristics for the vehicle routing problem (e.g. by area) to reduce runtime.

It is interesting to observe in the box plots in Figures 2 and 3 that delivery charges or discounts are usually fairly small, confirming the conjecture that customers' time slot choices can be influenced by such small (dis-)incentives. Moreover, it means that most of the time the restriction to the range [-10,10] does not impact on the price optimization problem. Relaxing the allowable range can improve results somewhat, but may be less realistic due to customer acceptance issues as mentioned before.

7 Conclusions and Future Research

We propose a framework for the dynamic pricing of delivery time slots for home delivery services, which adds to the existing literature in that we include customer choice behavior and dynamically estimate delivery costs based on both the accepted orders to date as well as on the orders that we still expect to come; hence we call this the "foresight approach". Customer choice is modeled via the multinomial logit model, and we show how this can be estimated from data that e-tailers have readily available, demonstrating the estimation on real data from a British e-grocer. Furthermore, in a simulation study we show that our approach can outperform static two-tier delivery pricing policies similar to those often found in practice, generating on average a 3.8% increase in profits. In an industry that operates on very small margins (e.g. Ocado reported operating margins of 0.55% in H1 2012 according to Mintel [2012]), this profit potential is remarkable. We also show numerically that dynamic pricing without taking future expected demand into account can produce even worse results than static pricing when delivery capacity is scarce. Given the small margins that most home delivery services (particularly e-grocers) work with, we conclude that adoption of such a foresight dynamic pricing policy is a crucial means for improving profitability.

As for future research directions, there are various additions that would be of interest: e.g., given that usually around 10% of orders are rescheduled, it would be useful to extend the model to include a probability that an order is canceled or re-scheduled. Furthermore, our model currently approximates the opportunity cost of accepting a customer order in a time slot only with regard to delivery cost, but not with regard to lost profit due to reduced capacity for future orders. Also, if

past final delivery schedules are not a good approximation of future schedules, the foresight policy could instead be based on final schedule predictions derived from the demand model. Major firms such as Amazon, Ebay or Walmart are moving towards same-day delivery services according to recent media reports, see e.g. Wirthman [2013]. In the light of companies such as WebVan who went bankrupt when trying to offer such a service, the question of how to address this added complexity in the model would be highly relevant to industry. In principle, our foresight approach could be adapted to allow for same-day deliveries by determining feasibility and insertion cost of time slots based on orders accepted-to-date and future anticipated orders, but taking into account the on-going delivery operations.

APPENDIX

Table 7: Price points in Example 4.1.

Time Slot Start	Price 1	Price 2	Price 3
06:00 AM	$3.50	$4.00	$5.00
06:30 AM	$4.50	$5.00	$6.00
07:00 AM	$4.50	$5.00	$6.00
07:30 AM	$4.50	$5.00	$6.00
08:00 AM	$4.50	$5.00	$6.00
08:30 AM	$4.50	$5.00	$6.00
09:00 AM	$4.50	$5.00	$6.00
09:30 AM	$4.50	$5.00	$6.00
10:00 AM	$4.50	$5.00	$6.00
10:30 AM	$3.50	$4.00	$4.50
11:00 AM	$3.50	$4.00	$4.50
11:30 AM	$3.50	$4.00	$4.50
12:00 PM	$4.00	$5.00	$6.50
12:30 PM	$4.00	$5.00	$6.50
01:00 PM	$4.00	$5.00	$6.50
01:30 PM	$4.00	$5.00	$6.50
02:00 PM	$4.00	$5.00	$6.50
02:30 PM	$4.00	$5.00	$6.50
03:00 PM	$4.00	$5.00	$6.50
03:30 PM	$4.00	$5.00	$6.50
04:00 PM	$3.50	$4.00	$5.00
04:30 PM	$3.50	$4.00	$5.00
05:00 PM	$2.50	$3.00	$4.00
05:30 PM	$2.00	$3.00	$4.00
06:00 PM	$2.00	$3.00	$4.00
06:30 PM	$1.50	$2.00	$2.50
07:00 PM	$1.50	$2.00	$2.50

Table 8: Recovery of MNL parameters based on 20,000 simulations.

	Original	Modified Original	Recovered	proportional diff	ASE	t-stat
β_1	-1.96	-1.86	-1.8700	1%	0.1286	-14.5412
β_2	-1.60	-1.50	-1.3759	-8%	0.1153	-11.9332
β_3	-1.52	-1.42	-1.3785	-3%	0.1146	-12.0288
β_4	-1.10	-1.00	-0.9770	-2%	0.0994	-9.8290
β_5	-0.98	-0.88	-0.7488	-15%	0.0917	-8.1658
β_6	-0.52	-0.42	-0.4271	2%	0.0839	-5.0906
β_7	-0.35	-0.25	-0.2929	17%	0.0804	-3.6430
β_8	0.10	0.20	0.1891	-5%	0.0721	2.6227
β_9	-0.25	-0.15	-0.0197	-87%	0.0756	-0.2606
β_{10}	0.00	0.10	0.0747	-25%	0.0649	1.1510
β_{11}	-0.50	-0.40	-0.3671	-8%	0.0728	-5.0426
β_{12}	-1.10	-1.00	-1.0789	8%	0.0924	-11.6764
β_{13}	-1.50	-1.40	-1.4505	4%	0.1170	-12.3974
β_{14}	-0.90	-0.80	-0.6622	-17%	0.0889	-7.4488
β_{15}	-1.00	-0.90	-0.9004	0%	0.0958	-9.3987
β_{16}	-0.81	-0.71	-0.7653	8%	0.0921	-8.3094
β_{17}	-0.85	-0.75	-0.6934	-8%	0.0895	-7.7475
β_{18}	-0.70	-0.60	-0.5625	-6%	0.0870	-6.4655
β_{19}	-0.95	-0.85	-0.6958	-18%	0.0900	-7.7311
β_{20}	-0.30	-0.20	-0.1818	-9%	0.0782	-2.3248
β_{21}	-0.53	-0.43	-0.4337	1%	0.0754	-5.7520
β_{22}	-0.20	-0.10	-0.0790	-21%	0.0683	-1.1567
β_{23}	-0.10	0.00	n/a	n/a	n/a	n/a
β_{24}	-0.20	-0.10	-0.2156	116%	0.0656	-3.2866
β_{25}	-1.13	-1.03	-0.9584	-7%	0.0829	-11.5609
β_{26}	-0.30	-0.20	-0.3369	68%	0.0683	-4.9327
β_{27}	-1.30	-1.20	-1.2000	0%	0.0885	-13.5593
β_0	-2.50	-2.60	-2.4181	-7%	0.0701	-34.4950
β_d	-0.10	-0.10	-0.1399	40%	0.0171	-8.1813

Table 9: Recovery of MNL parameters based on 40,000 simulations.

	Original	Modified Original	Recovered	proportional diff	ASE	t-stat
β_1	-1.96	-1.86	-0.0585	3%	0.0922	-20.8080
β_2	-1.60	-1.50	0.0414	-3%	0.0819	-17.8095
β_3	-1.52	-1.42	-0.0282	2%	0.0815	-17.7693
β_4	-1.10	-1.00	-0.0789	8%	0.0715	-15.0895
β_5	-0.98	-0.88	0.0108	-1%	0.0666	-13.0511
β_6	-0.52	-0.42	-0.0983	23%	0.0600	-8.6383
β_7	-0.35	-0.25	0.0099	-4%	0.0562	-4.2722
β_8	0.10	0.20	-0.0810	-41%	0.0517	2.3017
β_9	-0.25	-0.15	-0.0274	18%	0.0551	-3.2196
β_{10}	0.00	0.10	-0.0530	-53%	0.0464	1.0129
β_{11}	-0.50	-0.40	0.0610	-15%	0.0513	-6.6082
β_{12}	-1.10	-1.00	0.0054	-1%	0.0635	-15.6630
β_{13}	-1.50	-1.40	-0.0630	5%	0.0822	-17.7981
β_{14}	-0.90	-0.80	-0.0363	5%	0.0657	-12.7291
β_{15}	-1.00	-0.90	-0.0751	8%	0.0693	-14.0707
β_{16}	-0.81	-0.71	-0.0633	9%	0.0646	-11.9706
β_{17}	-0.85	-0.75	-0.0441	6%	0.0650	-12.2169
β_{18}	-0.70	-0.60	-0.0192	3%	0.0614	-10.0847
β_{19}	-0.95	-0.85	-0.0918	11%	0.0680	-13.8500
β_{20}	-0.30	-0.20	0.0008	0%	0.0549	-3.6284
β_{21}	-0.53	-0.43	-0.0365	8%	0.0538	-8.6710
β_{22}	-0.20	-0.10	-0.0697	70%	0.0495	-3.4283
β_{23}	-0.10	0.00	n/a	n/a	n/a	n/a
β_{24}	-0.20	-0.10	-0.0097	10%	0.0459	-2.3900
β_{25}	-1.13	-1.03	-0.0876	9%	0.0631	-17.7116
β_{26}	-0.30	-0.20	-0.0601	30%	0.0486	-5.3519
β_{27}	-1.30	-1.20	-0.0292	2%	0.0649	-18.9399
β_0	-2.50	-2.60	0.0055	0%	0.0497	-52.2032
β_d	-0.10	-0.10	0.0057	-6%	0.0121	-7.7934

Table 10: Recovery of MNL parameters based on 80,000 simulations.

	Original	Modified Original	Recovered	proportional diff	ASE	t-stat
β_1	-1.96	-1.86	-1.8282	-2%	0.0628	-29.1115
β_2	-1.60	-1.50	-1.4992	0%	0.0589	-25.4533
β_3	-1.52	-1.42	-1.4206	0%	0.0572	-24.8357
β_4	-1.10	-1.00	-0.9594	-4%	0.0484	-19.8223
β_5	-0.98	-0.88	-0.9570	9%	0.0486	-19.6914
β_6	-0.52	-0.42	-0.4727	13%	0.0418	-11.3086
β_7	-0.35	-0.25	-0.3058	22%	0.0402	-7.6070
β_8	0.10	0.20	0.1436	-28%	0.0365	3.9342
β_9	-0.25	-0.15	-0.2451	63%	0.0397	-6.1738
β_{10}	0.00	0.10	0.0827	-17%	0.0325	2.5446
β_{11}	-0.50	-0.40	-0.4248	6%	0.0372	-11.4194
β_{12}	-1.10	-1.00	-1.0018	0%	0.0449	-22.3118
β_{13}	-1.50	-1.40	-1.4574	4%	0.0579	-25.1710
β_{14}	-0.90	-0.80	-0.9214	15%	0.0476	-19.3571
β_{15}	-1.00	-0.90	-0.9406	5%	0.0478	-19.6778
β_{16}	-0.81	-0.71	-0.7735	9%	0.0457	-16.9256
β_{17}	-0.85	-0.75	-0.8149	9%	0.0460	-17.7152
β_{18}	-0.70	-0.60	-0.6564	9%	0.0441	-14.8844
β_{19}	-0.95	-0.85	-0.8908	5%	0.0474	-18.7932
β_{20}	-0.30	-0.20	-0.2156	8%	0.0390	-5.5282
β_{21}	-0.53	-0.43	-0.4452	4%	0.0378	-11.7778
β_{22}	-0.20	-0.10	-0.1306	31%	0.0347	-3.7637
β_{23}	-0.10	0.00	n/a	n/a	n/a	n/a
β_{24}	-0.20	-0.10	-0.0976	-2%	0.0325	-3.0031
β_{25}	-1.13	-1.03	-1.0517	2%	0.0437	-24.0664
β_{26}	-0.30	-0.20	-0.2266	13%	0.0342	-6.6257
β_{27}	-1.30	-1.20	-1.2365	3%	0.0464	-26.6487
β_0	-2.50	-2.60	-2.6069	0%	0.0353	-73.8499
β_d	-0.10	-0.10	-0.0913	-9%	0.0086	-10.6163

Table 11: Recovery of MNL parameters based on 160,000 simulations.

	Original	Modified Original	Recovered	proportional diff	ASE	t-stat
β_1	-1.96	-1.86	-1.9698	6%	0.0472	-41.7331
β_2	-1.60	-1.50	-1.5788	5%	0.0430	-36.7163
β_3	-1.52	-1.42	-1.4794	4%	0.0413	-35.8208
β_4	-1.10	-1.00	-1.0672	7%	0.0356	-29.9775
β_5	-0.98	-0.88	-0.8955	2%	0.0336	-26.6518
β_6	-0.52	-0.42	-0.4508	7%	0.0295	-15.2814
β_7	-0.35	-0.25	-0.3060	22%	0.0284	-10.7746
β_8	0.10	0.20	0.1730	-14%	0.0255	6.7843
β_9	-0.25	-0.15	-0.1863	24%	0.0275	-6.7745
β_{10}	0.00	0.10	0.0713	-29%	0.0230	3.1000
β_{11}	-0.50	-0.40	-0.4383	10%	0.0263	-16.6654
β_{12}	-1.10	-1.00	-0.9851	-1%	0.0315	-31.2730
β_{13}	-1.50	-1.40	-1.4133	1%	0.0402	-35.1567
β_{14}	-0.90	-0.80	-0.8002	0%	0.0324	-24.6975
β_{15}	-1.00	-0.90	-0.9430	5%	0.0339	-27.8171
β_{16}	-0.81	-0.71	-0.7429	5%	0.0318	-23.3616
β_{17}	-0.85	-0.75	-0.7533	0%	0.0319	-23.6144
β_{18}	-0.70	-0.60	-0.5968	-1%	0.0304	-19.6316
β_{19}	-0.95	-0.85	-0.8753	3%	0.0331	-26.4441
β_{20}	-0.30	-0.20	-0.2434	22%	0.0277	-8.7870
β_{21}	-0.53	-0.43	-0.4588	7%	0.0268	-17.1194
β_{22}	-0.20	-0.10	-0.1440	44%	0.0245	-5.8776
β_{23}	-0.10	0.00	n/a	n/a	n/a	n/a
β_{24}	-0.20	-0.10	-0.1186	19%	0.0230	-5.1565
β_{25}	-1.13	-1.03	-1.0582	3%	0.0308	-34.3571
β_{26}	-0.30	-0.20	-0.2440	22%	0.0241	-10.1245
β_{27}	-1.30	-1.20	-1.1945	0%	0.0320	-37.3281
β_0	-2.50	-2.60	-2.5728	-1%	0.0249	-103.3253
β_d	-0.10	-0.10	-0.0979	-2%	0.0060	-16.3167

References

Niels Agatz, Ann Campbell, Moritz Fleischmann, Jo van Nunen, and Martin Savelsbergh. Demand management opportunities in e-fulfillment: What internet retailers can learn from revenue management. Working Paper, 2008.

Niels Agatz, Ann Campbell, Moritz Fleischmann, and Martin Savelsbergh. Time slot management in attended home delivery. *Transportation Science*, 45:435–449, 2011.

Kursad Asdemir, Varghese S. Jacob, and Ramayya Krishnan. Dynamic pricing of multiple home delivery options. *European Journal of Operational Research*, 196:246–257, 2009.

M. Ben-Akiva and S. Lerman. *Discrete Choice Analysis: Theory and Applications to Travel Demand*. The MIT Press, Cambridge, 6th edition, 1994.

Ann Melissa Campbell and Martin Savelsbergh. Incentive schemes for attended home delivery services:. *Transportation Science*, 40:327–341, 2006.

Ann Melissa Campbell and Martin W.P. Savelsbergh. Decision support for consumer direct grocery initiatives. *Transportation Science*, 39:313–327, 2005.

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

M.L. Fischer and R. Jaikumar. A generalized assignment heuristic for vehicle-routing. *Networks*, 11:109–124, 1981.

W. Hanson and K. Martin. Optimizing mulitnomial logit profit functions. *Management Science*, 42:992–1003, 1996.

L. Leemis. Nonparametric estimation of the cumulative intensity function for a nonhomogeneous Poisson process. *Management Science*, 37:886–900, 1991.

D. McFadden. *Frontiers in Econometrics*, chapter Conditional logit analysis of qualitative choice behavior, pages 105–142. Academic Press, 1974.

J. Meissner, A. Strauss, and K. Talluri. An enhanced concave program relaxation for choice network revenue management. *Production and Operations Management*, 22:71–87, 2013.

Mintel. Online Grocery Retailing - UK - September 2012. Technical report, Mintel, 2012.

W. Newey and D. McFadden. *Large sample estimation and hypothesis testing*, chapter 36, pages 2111–2245. Elsevier, 1994.

M. Suh and G. Aydin. Dynamic pricing of substitutable products with limited inventories under logit demand. *IIE Transactions*, 43:323–331, 2011.

K. Train. *Discrete Choice Methods with simulation*. Cambridge University Press, New York, 2003.

K. Train. *Discrete Choice Methods with simulation, second edition*. Cambridge University Press, New York, 2009.

G. Vulcano, G. van Ryzin, and W. Chaar. Choice-based revenue management: An empirical study of estimation and optimization. *Manufacturing & Service Operations Management*, 12:371–392, 2010. Working Paper.

G. Vulcano, G. van Ryzin, and R. Ratliff. Estimating primary demand for substitutable products from sales transaction data. *Operations Research*, 60:313–334, 2012.

Mark Ward. Amazon's old customers 'pay more', September 2000. URL `http://news.bbc.co.uk/1/hi/business/914691.stm`.

Lisa Wirthman. Amazon, Ebay, Walmart same-day deliver, but should you? Forbes.com, April 2013. URL http://www.forbes.com/sites/ups/2013/04/04/amazon-ebay-walmart-same-day-deliver-but-should-you/