European Journal of Operational Research 292 (2021) 456–468

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Discrete Optimisation

A dynamic programming framework for optimal delivery time slot pricing

Denis Lebedev*, Paul Goulart, Kostas Margellos

Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, United Kingdom

ARTICLE INFO

Article history:
Received 25 October 2019
Accepted 5 November 2020
Available online 11 November 2020

Keywords:
Dynamic programming
Revenue management
Discrete convex analysis

ABSTRACT

We study the dynamic programming approach to revenue management in the context of attended home delivery. We draw on results from dynamic programming theory for Markov decision problems to show that the underlying Bellman operator has a unique fixed point. We then provide a closed-form expression for the resulting fixed point and show that it admits a natural interpretation. Moreover, we also show that – under certain technical assumptions – the value function, which has a discrete domain and a continuous codomain, admits a continuous extension, which is a finite-valued, concave function of its state variables, at every time step. Furthermore, we derive results on the monotonicity of prices with respect to the number of orders placed in our setting. These results open the road for achieving scalable implementations of the proposed formulation, as it allows making informed choices of basis functions in an approximate dynamic programming context. We illustrate our findings on a low-dimensional and an industry-sized numerical example using real-world data, for which we derive an approximately optimal pricing policy based on our theoretical results.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

The expenditure of US households on online grocery shopping could reach $100 billion in 2022 according to the Food Marketing Institute (2018). Although growth forecasts vary and more conservative estimates lie, for example, at $30 billion for the year 2021 (Pitchbook, 2017), the overall trend is clear: The online grocery sector is likely to grow if some of its main challenges can be overcome.

One of these challenges is managing the logistics as one of the main cost-drivers. In particular, one can seek to exploit the flexibility of customers by offering delivery options at different prices to create delivery schedules that can be executed in a cost-efficient manner. To achieve this, recent proposals include giving customers the choice between narrow delivery time windows for high prices and *vice versa* (Campbell & Savelsbergh, 2006) or charging customers different prices based on the area and their preferred delivery time (Asdemir, Jacob, & Krishnan, 2009; Yang & Strauss, 2017; Yang, Strauss, Currie, & Eglese, 2016).

In this paper, we focus on the latter. We refer to the problem of finding the profit-maximising delivery slot prices as the *revenue management problem in attended home delivery*, where “attended” refers to the requirement that customers need to be present upon delivery of the typically perishable goods, which is in contrast to, for example, standard mail delivery. Note that attended home delivery problems are more complex than standard delivery services, since goods need to be delivered in time windows that are pre-agreed with the customers.

We adopt a dynamic programming (DP) model of an expected profit-to-go function, the value function of the DP, given the current state of orders and time left for customers to book a delivery slot. This DP was initially devised in the fashion industry (Gallego & van Ryzin, 1994), but subsequently adopted and refined by the transportation sector and the attended home delivery industry (Yang et al., 2016). This formulation could be thought of as an instance of a network revenue management problem with customer choice, which finds various applications, e.g. in transportation, hospitality and appointment scheduling problems (see Meissner & Strauss, 2012; Sauré, Patrick, Tyldesley, & Puterman, 2012; Zhang & Adelman, 2009).

To find the (approximately) optimal delivery slot prices, we need to compute the value function (at least approximately) for all states and times. The main challenge is that the state space of the DP grows exponentially with the set of delivery time slots, i.e. it suffers from the “curse of dimensionality”. This means that for industry-sized problems, due to the prohibitively large number of states, the value function cannot be computed exactly, even offline. Our ultimate objective is to compute improved value func-

* Corresponding author.
E-mail addresses: denis.lebedev@eng.ox.ac.uk (D. Lebedev), paul.goulart@eng.ox.ac.uk (P. Goulart), kostas.margellos@eng.ox.ac.uk (K. Margellos).

https://doi.org/10.1016/j.ejor.2020.11.010
0377-2217/© 2020 Elsevier B.V. All rights reserved.

tion approximations. Therefore, we study in this paper how the value function of the exact DP behaves mathematically in time and across state variables.

We show that the underlying DP operator has a unique fixed point. We then provide a closed-form expression of the resulting fixed point and derive a natural interpretation. Furthermore, we show that – under certain technical assumptions – for all time steps in the dynamic program, the value function admits a continuous extension, which is a finite-valued, concave function of its state variables.

Ultimately, our results open the road for achieving scalable implementations of the proposed formulation, as it becomes possible to make informed choices of basis functions in an approximate dynamic programming context. We illustrate our findings on a low-dimensional and an industry-sized numerical example using real-world data from a case study by Yang and Strauss (2017), for which we derive an approximate value function based on our theoretical results and a stochastic dual DP algorithm presented in Zhang and Sun (2019).

Improved value function approximations could finally be used for calculating approximately optimal delivery slot prices. For example, for continuous decision variables and under the multinomial logit customer choice model, Dong, Kouvelis, and Tian (2009) show that a unique set of optimal delivery slot prices exists, which can be found using Newton root search algorithms or using the Lambert W function as shown in B.2 if estimates of the value function are known for all states and times. Our mathematical results have immediate implications on the monotonicity of (approximately) optimal prices with respect to changes in the number of placed orders, which we also characterise in this paper. This analysis complements the research on the price-inventory relationship under multinomial logit customer choice (see e.g. Akçay, Natarajan, & Xu, 2010; Chen & Chen, 2015; Suh & Aydin, 2011).

Our paper is structured as follows: In the remainder of Section 1, we introduce some notation. In Section 2, we define the revenue management problem in attended home delivery and its DP formulation. In Section 3, we present our main results, Theorem 1, which analytically characterises the fixed point of the DP, and Theorem 2, which shows that there exists a continuous extension of the value function that is a finite-valued, concave function in its state variables at every time step. Section 4 contains reformulations of the DP into mathematically more convenient forms and develops supporting results leading to the proofs of the main results. We also develop a result on the monotonicity of prices with respect to the number of placed orders. Section 5 presents a numerical illustration of our theoretical results on a low-dimensional example and on an industry-sized problem, while Section 6 concludes the paper and suggests directions for future research. The Appendix contains the proofs of results not included in the main body of the paper.

Notation: Let $\mathbf{1}$ denote a vector with all elements equal to 1. Given some s, let 1_s be a vector of all zeros apart from the s-th entry, which equals 1. Furthermore, we define the convention that 1_0 is a vector of zeros. Let $\mathbb{R}_{+(+)}$ be the non-negative (positive) real numbers, let $\mathbb{Z}$ be the integers and let $\dim(\cdot)$ denote the dimension of its argument. Let $\mathrm{conv}(\cdot)$ denote the convex hull of its argument. We say that a function exhibits a monotonic behaviour if the monotonicity property holds element-wise, e.g. a function $f : \mathbb{R}^N \mapsto \mathbb{R}$ is monotonically increasing over its domain if $f(y) > f(x)$ for all (x, y), such that at least one element of y is greater than the corresponding element of x.

2. Revenue management problem formulation

In this section, we derive a discrete-state formulation of the revenue management problem in attended home delivery.

2.1. Problem statement

We model an online business that delivers goods to locations of known customers. We consider a local approximation of the revenue management problem by dividing the service area geographically into a set of non-overlapping rectangular sub-areas, where the customers in each sub-area operate independently by being served by one delivery vehicle. This model resembles the setting in the work of Yang and Strauss (2017). Due to this independence, we only consider a single sub-area, while our development directly extends to the case of multiple sub-areas. To cover all sub-areas in practice, it is possible to simply replicate our approach for every delivery sub-area, which would increase computational complexity linearly in the number of sub-areas, but which is easily parallelised.

We consider a finite booking horizon with possibly unequally-spaced time steps indexed by $t \in T := \{1, 2, \ldots, \bar{t}\}$. Based on the development of Yang et al. (2016, Section 4.3), we obtain a customer arrivals model using a Poisson process with *time-invariant* event rate $\lambda \in (0, 1)$ for all $t \in T$ from a Poisson process with homogeneous time steps, but *time-varying* event rate.

Customers can choose from a number of (typically 1-hour wide) delivery time windows, which we call slots $s \in S$, where $S := \{1, 2, \ldots, \bar{s}\}$. Let $s = 0$ correspond to a customer not choosing any slot. Each delivery slot s is assigned a delivery charge $d_s \in \left[\underline{d}, \bar{d}\right] \cup \{\infty\}$, for some minimum allowable charge $\underline{d} \in \mathbb{R}$ (which is typically, though not necessarily, non-negative) and some maximum allowable charge $\bar{d} \geq \underline{d}$. The role of $d_s = \infty$ is a convention to indicate that slot s is not offered. This is explained in more detail when introducing the customer choice model below. We define the delivery charge vector $d := [d_1, d_2, \ldots, d_{\bar{s}}]^\intercal$. Let the set of admissible delivery charge vectors be $D := \left\{d \;\middle|\; d_s \in \left[\underline{d}, \bar{d}\right] \cup \{\infty\} \text{ for all } s \in S\right\}$.

For each delivery slot $s \in S$, we denote the number of placed orders by $x_s \in \mathbb{Z}$. We also define $x := [x_1, x_2, \ldots, x_{\bar{s}}]^\intercal \in \mathbb{Z}^{\bar{s}}$ as well as $X := \{x \mid 0 \leq x_s \leq \bar{x}_s \text{ for all } s \in S\}$, where $\bar{x}_s$ is a scalar indicating the maximum number of deliveries that can be fulfilled in slot s. In general, we do not require the maximum number of deliveries to be the same for all slots, e.g. since this will depend on traffic patterns in the delivery area. Examples of computing this quantity can be found in Yang and Strauss (2017, Section 4). Let us define $\bar{x} := [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{\bar{s}}]^\intercal$ as well as the set of feasible slots $F(x) = \{s \in S \mid x + 1_s \in X\}$. Let $r \in \mathbb{R}$ denote the expected net revenue of an order, i.e. expected revenue minus costs prior to delivery. This is assumed to be invariant across all orders. We define

$$C(x) := \begin{cases} C_+(x), & \text{if } x \in X \\ \infty & \text{otherwise,} \end{cases} \tag{1}$$

where $C_+ : X \to \mathbb{R}_+$ is a given function. The function C approximates the delivery cost to fulfil the set of orders x. The delivery cost cannot be computed exactly, as it is the solution to a vehicle routing problem with time windows, which is intractable for industry-sized applications (Toth & Vigo, 2014). For the DP approach that we adopt in this paper and introduce in the next section, we need to know this function at the start of the booking horizon, i.e. before any orders are placed. Therefore, it is also prohibitive to include additional details like customer locations in this function, since this would increase the state-space to an intractable size. In principle, if a fast enough algorithm to compute approximately optimal delivery cost prices was available, one could update the delivery cost function as orders are placed and adjust the approximately optimal prices accordingly. In this paper, we mainly focus on the first part of the problem i.e. given one approximate delivery cost function C, we would like to compute the approximately optimal delivery prices. We further discuss the possibility of updating the terminal condition with additional details like customer locations and updating the approximate DP solution between orders in the context of our numerical example in Section 5.3.

Let the probability that a customer chooses delivery slot s if offered prices d be $\Pi_s(d)$, such that $d \mapsto \Pi_s(d) \in [0, 1)$ for all $s \in S$. Note that $\sum_{s \in S} \Pi_s(d) = 1 - \Pi_0(d)$, where $\Pi_0 > 0$ denotes the probability of a customer leaving the online ordering platform without choosing any delivery slot. A typical choice for Π_s is the multinomial logit model that was also used in Yang and Strauss (2017):

$$\Pi_s(d) := \frac{\exp(\beta_c + \beta_s + \beta_d d_s)}{\sum_{k \in S} \exp(\beta_c + \beta_k + \beta_d d_k) + 1}, \tag{2}$$

where $\beta_c \in \mathbb{R}$ denotes a constant offset, $\beta_s \in \mathbb{R}$ represents a measure of the popularity for all delivery slots and $\beta_d < 0$ is a parameter for the price sensitivity. Note that the no-purchase utility is normalised to zero, i.e. for the no-purchase "slot" $s = 0$, we have a no-delivery "charge" $d_0 = 0$, such that $\beta_c + \beta_0 + \beta_d d_0 = \beta_c + \beta_0 = 0$ and hence, the 1 in the denominator of (2) arises from $\exp(\beta_c + \beta_0) = 1$. Furthermore, note that the constant offset β_c is not necessary, since it can be encompassed within the $\{\beta_s\}_{s \in S \cup \{0\}}$ parameters. However, β_c is often preserved to normalise one of the $\{\beta_s\}_{s \in S \cup \{0\}}$ parameters to zero. This is also performed in Yang and Strauss (2017).

Note that our results on the fixed point computation do not depend on the particular form of the customer choice model. We require only that it is a probability density function and that it satisfies one mild technical assumption stated further below in Section 3.

For convenience, let the probability that a customer arrives and chooses slot s given prices d be denoted by $p_s(d) := \lambda \Pi_s(d)$. We define $p(d) := [p_1(d), p_2(d), \ldots, p_{\bar{s}}(d)]^\intercal$ and $P := \{p(d) \mid d \in D\}$. Finally, it is to be understood that all sums over s are always computed over the entire set S.

2.2. Dynamic programming formulation

We can express the problem described above as a DP. The expected profit-to-go closely resembles the DP formulation in Yang and Strauss (2017) and we define it as

$$V_t(x) := \max_{d \in D} \left\{ \sum_s p_s(d)[r + d_s + V_{t+1}(x + 1_s) - V_{t+1}(x)] + V_{t+1}(x) \right\},$$
$$\text{for all } (x, t) \in X \times T, \text{ where } V_{\bar{t}+1}(x) = -C(x) \quad \forall x \in X, \tag{3}$$

i.e. $C(\cdot)$ denotes the terminal condition. This representation is accurate for all feasible orders, i.e. for all $x \in X$, such that $x + 1_s \in X$, for all $s \in S$. For infeasible orders, i.e. when for some $s \in S$, $x + 1_s \notin X$, the term in square brackets can be undefined, since $V_{t+1}(x + 1_s) = -\infty$ and $d_s = \infty$. Since in this case $p_s(d) = 0$, we adopt the convention that the product of $p_s(d)$ and the expression in square brackets equals zero. The interpretation of this is that we assign zero additional value to an infeasible slot.

The difference $V_{t+1}(x) - V_{t+1}(x + 1_s)$ in (3) represents the value foregone by accepting an additional (discrete spatial) order, which in economic terms is the *opportunity cost* of an order. Note that – similar to Yang and Strauss (2017) – we ignore any vehicle load capacity constraints in the problem, as they are much less restricting than the time constraints on the delivery slots. Therefore, including the vehicle load capacity constraints would only increase computational costs, but would not substantially improve the decision policy. For convenience in the sequel, we define the DP operator $\mathcal{T}$ to express (3) in a more compact form as

$$V_{t-1} := \mathcal{T} V_t, \text{ for all } t \in T. \tag{4}$$

3. Infinite and finite time horizon results

3.1. Infinite time horizon result

We first consider the infinite horizon case, i.e. going backwards infinitely many time steps. In this scenario, we can find a fixed point of the DP described by (4) based on the following assumptions.

Assumption 1. The marginal cost of an additional, feasible order is always smaller than the maximum marginal profit, i.e. $C(x + 1_s) - C(x) \le \bar{d} + r$, for all $(x, s) \in X \times F(x)$.

Assumption 2. We assume that the transition probability density function has the following properties. For any $s \in S$:

(a) $p_s(d) > 0$, if $d_s \in [\underline{d}, \bar{d}]$.
(b) $p_s(d) d_s = 0$, if $d_s = \infty$.

Assumption 1 is not restrictive, since it offers the means to ensure that every additional, feasible order can generate profit. Otherwise, the delivery slot price, which maximises (3), would be $d_s = \infty$ for some $s \in S$, even if additional orders would still be feasible for that slot. Assumption 2(a) is not restrictive either, since we can change $\bar{d}$ to a value for which the choice probability for all slots is modelled to be arbitrarily small, yet positive. Assumption 2(b) is also not restrictive, since it effectively ensures that infinite prices do not generate infinite expected returns. Furthermore, it is easy to show that Assumption 2(b) holds for the adopted multinomial logit model. Based on the aforementioned definitions and Assumption 1, we formulate our first result, proof of which is deferred to Section 4.1.

Theorem 1. *Under Assumptions 1–2, the unique fixed point of (4) is given by*

$$V^*(x) := (\bar{d} + r)\mathbf{1}^\intercal(\bar{x} - x) - C(\bar{x}), \text{ for all } x \in X. \tag{5}$$

There is a natural interpretation of this perhaps surprisingly compact result: The fixed point of the DP is a hyperplane in x, where each element of the gradient of V^* is equal to $-(\bar{d} + r)$, or equivalently, the opportunity cost of an order is $V_t(x) - V_t(x + 1_s) = \bar{d} + r$ for all $(x, s) \in X \times F(x)$. Therefore, the only optimal selection of delivery slot prices is to choose $\bar{d}$ for all delivery slots $s \in F(x)$. For any other choice the opportunity costs would be larger than the revenue generated by any order. This result makes intuitive sense as in the limit as $t \to -\infty$, there will always arrive enough customers who will be willing to pay $\bar{d}$ for a delivery. Therefore, in the infinite time horizon case, it is best to always charge the maximum admissible delivery charge. Note that this result is not of direct practical use as a pricing policy. However, it serves as a useful upper bound to the value function for all time steps, which we exploit as an initial value function approximation in the non-linear stochastic dual DP example in Section 5.3, to speed up computation.

3.2. Finite time horizon result

For finite $\bar{t}$, we establish a geometric property of the value function V_t, $t \in T$, related to concavity of a continuous function. As the domain of V_t is discrete, it is not possible to establish this property from convexity theory. We provide some definitions before stating our main result.

Let the opportunity cost of an order in slot s at time t be denoted by $\gamma_{s,t}(x) := V_t(x) - V_t(x + 1_s) \ge 0$ for all $(x, s, t) \in X \times S \times T$. Let $\gamma_t(x) := [\gamma_{1,t}(x), \gamma_{2,t}(x), \ldots, \gamma_{\bar{s},t}(x)]$. We define the set of stochastic vectors in a set A as

$$\mathcal{V}_A := \left\{ v \in \mathbb{R}_+^{|A|} \;\middle|\; \sum_{i=1}^{|A|} v_i = 1 \right\}, \tag{6}$$

where v_i denotes the i-th component of v. Let $x \in X$ and let $Q \subseteq \mathbb{Z}^{\bar{s}}$ be a finite set. Then Q is defined to be an *enclosing set* of x if $x \in \operatorname{conv}(Q)$. We define $\mathcal{Q}(x)$ as the set of all sets Q enclosing x. The following two definitions are frequently used in discrete convex analysis:

Definition 1 (cf. Murota and Shioura (2001, (2.1))). Let $a \in \mathbb{R}^N$ and $b \in \mathbb{R}$. Then the concave closure $\tilde{f} : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty\}$ of a function $f : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ is defined as

$$\tilde{f}(x) := \inf_{a,b}\left\{a^\top x + b \;\middle|\; a^\top y + b \geq f(y) \quad \forall y \in \mathbb{Z}^N\right\}. \tag{7}$$

Definition 2 (cf. Murota and Shioura (2001, Lemma 2.3) and (Rockafellar & Wets, 1998, Proposition 2.31)). A function $f : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ is concave extensible if and only if any of the following equivalent conditions hold:

(a) The evaluations of f coincide with the evaluations of its concave closure $\tilde{f}$, i.e. $f(x) = \tilde{f}(x)$ for all $x \in \mathbb{Z}^N$.
(b) For all $x \in X$ and for all $Q \in \mathcal{Q}(x)$, the evaluation of f at x does not lie below any possible linear interpolation of f on the points $q \in Q$, i.e. for all $x \in X$, for all $Q \in \mathcal{Q}(x)$ and for all $\mu \in \mathcal{V}_Q$, such that $x = \sum_{q \in Q} \mu_q q$, it holds that

$$f(x) \geq \sum_{q \in Q} \mu_q f(q). \tag{8}$$

Based on these definitions, we impose the following assumptions on our finite time horizon result.

Assumption 3. We assume that the opportunity cost at the terminal condition $\gamma_{s,\bar{t}+1}$ of all orders is increasing in x for all unit hypercubes in X, i.e. $\gamma_{s,\bar{t}+1}(x) < \gamma_{s,\bar{t}+1}(x + 1_{s'})$ for all $(x, s, s') \in X \times F(x) \times F(x)$, such that $s \neq s'$.

Assumption 4. The function $-C$ is concave extensible.

Assumption 3 is satisfied if $V_{\bar{t}+1}$ is strictly submodular. This is since for all strictly submodular functions $f : \mathbb{Z}^n \to \mathbb{R}$ we have

$$f(\max(y, z)) + f(\min(y, z)) < f(y) + f(z) \tag{9}$$

for all y and $z \in \operatorname{dom}(f)$, where the maximum and minimum are taken componentwise (e.g. see Bertsimas and Weismantel (2005, Definition 3.2)). This means that f has increasing opportunity costs, since, for all $(x, s, s') \in X \times S \times S$, such that $s \neq s'$, we can set $f = V_{\bar{t}+1}$, $y = x + 1_s$, $z = x + 1_{s'}$, which yields the desired inequality:

$$\begin{aligned} &V_{\bar{t}+1}(x + 1_s + 1_{s'}) + V_{\bar{t}+1}(x) < V_{\bar{t}+1}(x + 1_s) + V_{\bar{t}+1}(x + 1_{s'}) \\ \iff &V_{\bar{t}+1}(x) - V_{\bar{t}+1}(x + 1_s) < V_{\bar{t}+1}(x + 1_{s'}) - V_{\bar{t}+1}(x + 1_s + 1_{s'}) \\ &\iff \gamma_{s,\bar{t}+1}(x) < \gamma_{s,\bar{t}+1}(x + 1_{s'}). \end{aligned} \tag{10}$$

Since $V_{\bar{t}+1}$ needs to be strictly submodular, this requires that C is strictly supermodular as $V_{\bar{t}+1}(x) = -C(x)$ for all $x \in X$. This is not the case for all C used in the literature. For example, Yang and Strauss (2017) use an affine cost function. However, our results are also relevant for situations with affine cost functions, since – as we show numerically in Section 5 – the value function can reach a state where Assumption 3 is satisfied in a small number of iterations of the Bellman operator. Assumption 4 is weak as it is satisfied by any convex cost function, which also includes the aforementioned affine cost functions. We can now state our second main result.

Theorem 2. *Under Assumptions 1, 3 and 4, there exists a sufficiently small $\lambda > 0$, such that V_t is finite-valued, concave extensible in x for all $t \in T$.*

In the following section, we prove our main two results. Furthermore, we quantify a range of values for λ such that Theorem 2 always holds. This condition is reported in B.2.

4. Proofs of main results

4.1. Proof of infinite time horizon theorem

To prove Theorem 1, we first note that the DP in (4) can be reformulated as a so-called stochastic shortest path problem (see Lebedev, Goulart, and Margellos, 2019, Section 4.1.1). The Bellman operator mapping of this class of problems is known to be contractive (see Bertsekas (2012, Chapters 1 and 3) and Lebedev et al. (2019, Lemma 5)). Therefore, the DP in (4) admits a unique fixed point. We start with the necessary and sufficient condition for $\mathcal{T}$ to have a fixed point V^*, which is $V^* = \mathcal{T}V^*$. Setting $V_t(x) = V_{t+1}(x) = V^*(x)$ in (3) yields

$$\max_{d \in D}\left\{\sum_s p_s(d)[r + d_s + V^*(x + 1_s) - V^*(x)]\right\} = 0. \tag{11}$$

Substituting the candidate V^* from (5) into (11) results then in

$$\max_{d \in D}\left\{\sum_s p_s(d)[d_s - \bar{d}\,]\right\} = 0. \tag{12}$$

Fix any $s \in S$ and consider the following two possible cases:

Case I: Suppose that $d_s \in [\underline{d}, \bar{d}]$. Then by Assumption 2(a), $p_s(d) > 0$. Furthermore, the value of $[d_s - \bar{d}\,]$ is non-positive and 0 only if $d_s = \bar{d}$. Hence, the maximum value that $p_s(d)[d_s - \bar{d}]$ can take in this case is 0, namely when $d_s = \bar{d}$.

Case II: Suppose that $d_s = \infty$. Then by Assumption 2(b), $p_s(d)[d_s - \bar{d}] = 0$.

Since in both cases the maximum attainable value of each term in the sum over s is 0, the equality in (12) holds.

Finally, notice that $V_t(\bar{x}) = C(\bar{x})$ for all $t \in T$. Since the candidate fixed point satisfies $V^*(\bar{x}) = V_{\bar{t}+1}(\bar{x}) = -C(\bar{x})$, V^* is a fixed point of $\mathcal{T}$ for all $x \in X$. □

4.2. Proof of finite time horizon theorem

In this section, we prove Theorem 2. We start by reformulating (3) as a maximisation over $p \in P$ instead of $d \in D$. As shown by Dong et al. (2009), this is possible, since, for all $s \in S$, the following unique mapping between p and d exists:

$$\frac{p_s}{p_0} = \exp(\beta_c + \beta_s + \beta_d d_s), \tag{13}$$

where we recall that $p_0 = \lambda \Pi_0 > 0$. We solve this equation with respect to d_s to obtain

$$d_s = \beta_d^{-1}\left[\ln\left(\frac{p_s}{p_0}\right) - \beta_c - \beta_s\right]. \tag{14}$$

We will prove the theorem by induction. To this end, we fix an arbitrary $t \in T$, assume for an induction hypothesis that V_t is concave extensible in x and now show that $V_{t-1} = \mathcal{T}V_t$ is also concave extensible. Note that the base case in our induction proof is captured by Assumption 4. By substituting (14) into (3) we obtain

$$\begin{aligned} \mathcal{T}V_t(x) &= \max_{p \in P} \sum_s p_s \left\{ r + \beta_d^{-1}[\ln(\frac{p_s}{p_0}) - \beta_c - \beta_s] + V_t(x + 1_s) \right. \\ &\qquad \left. - V_t(x) \right\} + V_t(x) \\ &= \max_{p \in P}\{f(p) + g_t(x, p)\}, \end{aligned} \tag{15}$$

where we have defined

$$\begin{aligned} f(p) &:= \sum_s p_s\left\{r + \beta_d^{-1}\left[\ln\left(\frac{p_s}{p_0}\right) - \beta_c - \beta_s\right]\right\}, \\ g_t(x, p) &:= \sum_s p_s\{V_t(x + 1_s) - V_t(x)\} + V_t(x) \end{aligned} \tag{16}$$

for all $(x, p) \in X \times P$. This allows us to formulate the following result, whose parts we prove in Appendices A and B, respectively. Recall that $p_s = \lambda \Pi_s$ for all $s \in S$.

Lemma 3. *For all $t \in T$, the functions f and g_t have the following properties:*

(i) *The function f is concave in p.*
(ii) *Under Assumption 3 and if V_t is concave extensible in x, there exists a sufficiently small $\lambda > 0$ such that the function g_t is concave extensible in (x, p).*

The proof of Lemma 3(i) is given in Appendix A. In Lemma 3(ii), we assume that V_t is concave extensible in x as this is embedded within our induction proof for Theorem 2, where this corresponds to our induction hypothesis. The proof of Lemma 3(ii) depends on the following self-contained result. Let us consider a relaxation of the constraint on the optimisation variable d in (3) and optimise over $\mathbb{R}^{\bar{s}}$ instead of D. We refer to this problem as the *unconstrained* DP as opposed to the original, *constrained* DP.

Proposition 4. *The constrained and unconstrained version of the DP share the following property:*

(i) *Consider the unconstrained DP. Under Assumption 3, the opportunity cost $\gamma_{s,t}$ of all orders is increasing for all unit hypercubes in X, i.e.*

$$\gamma_{s,t}(x + 1_{s'}) > \gamma_{s,t}(x) \tag{17}$$

for all $(x, s, s', t) \in X \times F(x) \times F(x) \times T$, such that $s \neq s'$.
(ii) *Property (17) also holds for the constrained DP.*

We prove the two parts of Proposition 4 in B.2 and B.3, respectively. These results are then used in B.4 to prove Lemma 3(ii). An interesting implication of Proposition 4 is that increasing opportunity costs in x imply that unconstrained optimal prices exhibit monotonic behaviour in x:

Lemma 5. *Under Assumptions 1, 3 and 4, there exists a sufficiently small $\lambda > 0$, such that for all slots $s \in S$, the optimal price in the unconstrained DP for this slot, for any state $x \in X$, denoted by d_s^* is non-decreasing in x_s and non-increasing in $x_{s'}$, for all $s' \in S \setminus \{s\}$.*

We prove this result in Appendix C. Lemma 5 also complements the research on the price-inventory relationship under multinomial logit customer choice, i.e. the function describing prices in terms of remaining order capacity, i.e. d^* as a function of $\bar{x} - x$, for all time steps in the booking horizon (see Chen & Chen, 2015 for a review). It would be desirable to understand if the function exhibits monotonic behaviour, e.g. if prices for a slot increase or decrease monotonically as orders for that slot (or another slot) increase. For the case of multinomial logit customer choice, such statements cannot be made without additional assumptions. For example, Akçay et al. (2010, Section 6.2) show that prices for a slot are not necessarily non-increasing as orders in a different slot increase. For the two-dimensional case, i.e. $\bar{s} = 2$ in our notation, Suh and Aydin (2011, Proposition 2) show that prices for a slot are non-decreasing in the number of orders for that slot. Our results, however, provide the necessary assumptions and setting so that monotonicity statements can be made for the multi-dimensional case. We illustrate this numerically in the example of Section 5.2.

By Lemma 3, there exists a sufficiently small λ such that g_t has a continuous extension $\tilde{g}_t$, which is jointly concave in (x, p). By inspection, f is only a function of the continuously-valued variable p. Therefore, $f(p) + \tilde{g}_t(x, p)$ is also jointly concave in (x, p). We define $U(x) := \max_{p \in P} \{f(p) + \tilde{g}_t(x, p)\}$. By Rockafellar and Wets (1998, Proposition 2.22) or Boyd and Vandenberghe (2004, Section 3.2.5), partial maximisation with respect to some variables of a continuous multivariate function that is jointly concave in all its variables, preserves concavity in the resulting function. Therefore U is a concave function of x.

It remains to show that $U(x) = \mathcal{T}V_t(x)$ for all gridpoints $x \in X$. Repeating the same calculation, now with the discrete $f(p) + g_t(x, p)$ in place of $f(p) + \tilde{g}_t(x, p)$, i.e. $\mathcal{T}V_t(x) = \max_{p \in P} \{f(p) + g_t(x, p)\}$, note that $\tilde{f}(p) + \tilde{g}_t(x, p) = f(x) + g_t(x, p)$ for all $x \in X$ by Definition 2(a). Therefore, $\mathcal{T}V_t(x) = U(x)$ for all $x \in X$. This shows that U is a continuous extension of $\mathcal{T}V_t$, which is concave in x. Hence, $\mathcal{T}V_t$ is concave extensible in x. This concludes our induction argument and shows that the value function V_t is concave extensible in x for all $t \in T$. □

5. Numerical examples

In the following two sections, we illustrate the validity and practical utility of our results. We first show a low-dimensional numerical example of a 2-slot problem and then how our results allow the application of a non-linear stochastic dual DP algorithm to a 17-slot problem.

5.1. Increasing opportunity costs in an illustrative 2-slot example

Table 1
The parameters of the numerical example.

(a) Fixed parameters.	
$(\lambda, \bar{t}, \bar{s})$	(0.5,200,2)
$(\underline{d}, \bar{d}, r)$	(0,2,2)
$\bar{x}$	$[4, 4]^\top$
$(\beta_c, \beta_d, \beta_1, \beta_2)$	$(0, -1, 1, -1)$
(b) Variable terminal conditions.	
Terminal condition 1	$C_+(x) = 2 + 2x_1 + x_2$
Terminal condition 2	$C_+(x) = 2 + 2x_1 + 2x_2$
Terminal condition 3	$C_+(x) = 2 + 2x_1 + 3x_2$
Terminal condition 4	$C_+(x) = 2 + 2x_1 + 4x_2$

To illustrate our results, consider a 2-slot problem, where we fix the parameters listed in Table 1(a) and vary the terminal conditions as listed in Table 1(b).

Notice that all terminal conditions violate Assumption 3. However, after a few iterations of the Bellman operator, the opportunity costs become strictly increasing by inspection for all terminal conditions, thus satisfying Assumption 3 if $\bar{t}$ is set to that time instance. This can be seen in Fig. 1(a), where we plot the quantity

$$\min_{(x,s,s') \in X \times F(x) \times F(x),\, s \neq s'} \gamma_{s,t}(x + 1_{s'}) - \gamma_{s,t}(x) \tag{18}$$

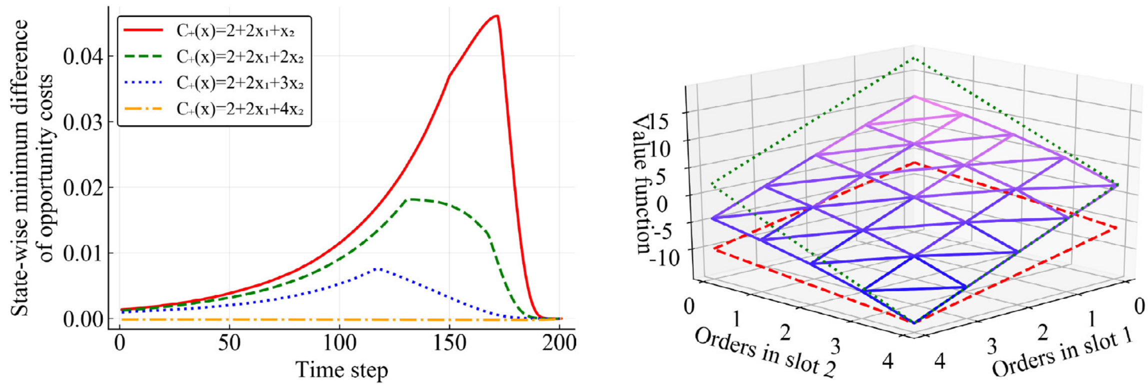
for all time steps $t \in T$. Observing that this quantity is always non-negative shows numerically that the result in Proposition 4 with non-strict inequality holds if the terminal condition has non-decreasing opportunity costs, i.e. constant in our case, since all C_+ are affine functions of x. Furthermore, it is easy to verify that the resulting value function is concave extensible by computing the concave closure of V_t for all $t \in T$ and checking that the concave closure $\tilde{V}_t(x) = V_t(x)$ for all $x \in X$. For this example, we can verify by direct observation that the function is concave extensible by plotting V_t against (x_1, x_2) as shown in Fig. 1(b) for the first terminal condition in Table 1(b) at time step $t = \bar{t} - 10$.

In Fig. 1(b), we also include the terminal condition (red, dashed line)

$$V_{\bar{t}+1}(x) := -C_+(x) = -2 - 2x_1 - x_2 \tag{19}$$

and, from Theorem 1, we also plot the fixed point (green, dotted line)

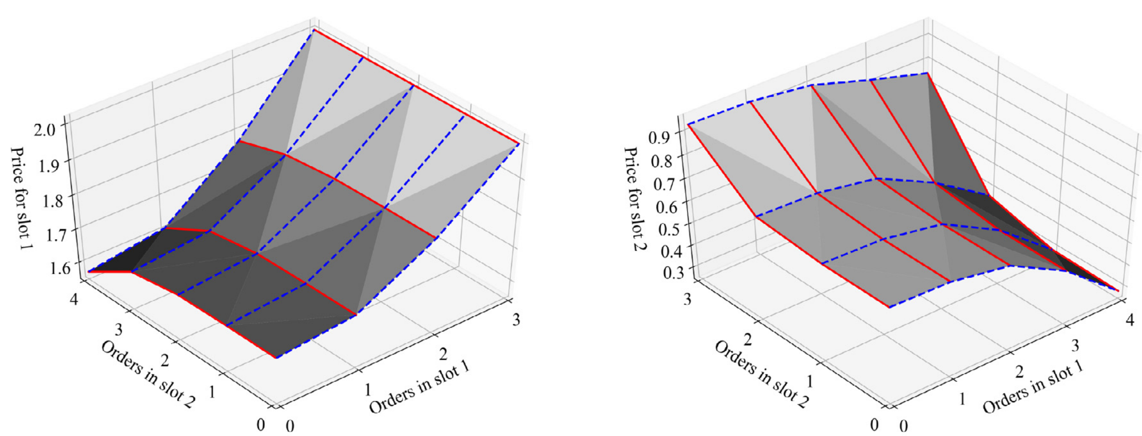
$$\begin{aligned} V^*(x) &:= (\bar{d} + r)\mathbf{1}^\top(\bar{x} - x) - C(\bar{x}) \\ &= 18 - 4(x_1 + x_2), \end{aligned} \tag{20}$$



(a) The state-wise minimum difference of opportunity costs is non-negative for all time steps t in the booking horizon and for all four terminal conditions of Table 1(b).

(b) The value function at $t = \bar{t} + 1$ (red, dashed line), at the fixed point $t = -\infty$ (green, dotted line) and at $t = \bar{t} - 10$ (blue/violet colour gradient).

Fig. 1. Illustrative example of a 2-slot problem. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) Optimal price of slot 1 at time $t = \bar{t} - 5$ is non-decreasing in orders in slot 1 (indicated by the blue, dashed lines) and non-increasing in orders in slot 2 (indicated by the red, solid lines).

(b) Optimal price of slot 2 at time $t = \bar{t} - 5$ is non-decreasing in orders in slot 2 (indicated by the red, solid lines) and non-increasing in orders in slot 1 (indicated by the blue, dashed lines).

Fig. 2. Monotonicity of prices. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

which corresponds to $V^* = \lim_{t\to-\infty} V_t$. When it comes to approximating V_t, e.g. by means of basis functions, we can use the observation that V_t always lies between the terminal condition and the fixed point to limit the range of basis functions, such that the approximated version of V_t also falls between these lower and upper bounds. Also notice that the value function at $t = \bar{t} - 10$ is concave extensible and has increasing opportunity costs.

5.2. Monotonicity of prices in the illustrative 2-slot example

We now illustrate the results of Lemma 5, namely that there exists a small enough $\lambda > 0$ such that, for all time steps $t \in T$, the optimal price of a slot is increasing in the number of orders of that slot and decreasing the number of orders of any other slot.

Consider the parameters of the numerical example from the previous section, Table 1(a) and the first terminal condition of Table 1(b). We obtain optimal prices for both slots by direct computation of the value function of the DP in Eq. (3). We show these prices for the time step $t = \bar{t} - 5$ in Fig. 2.

It should be remarked that the monotonicity property of Lemma 5 was shown to hold for unconstrained problems. However, it appears to hold for this numerical example even in the presence of constraints on admissible prices. Proving such a property for constrained problems constitutes a direction of current research.

Table 2
The parameters of the non-linear stochastic dual DP example.

$(\lambda, \bar{t}, \bar{s})$	(0.8,53,17)
$(\underline{d}, \bar{d}, r)$	(0,10,34.53)
$\bar{x}$	$[12, \dots, 12]^\intercal$
(β_c, β_d)	(−2.5087, −0.0766)
$\{\beta_s\}_{s\in S}$	{−1.0305, −0.3591, 0.3107, 0.5922, 0.6154, 0.0796, 0.5356, −0.2415, −0.6286, −1.6736, −0.4351, −0.161, 0, 0.2533, 0.0736, 0.562, 0.2346}
$C(x)$	$0.1042 \times \mathbf{1}^\intercal x$

5.3. Exploiting concave extensibility in non-linear stochastic dual DP

We now consider a more realistic problem described by the parameters in Table 2, which we have adapted from a real-world data, multi-subarea case study from Yang and Strauss (2017) to a single subarea. Notice that due to the large state space, $|X| = (\bar{x}+1)^{\bar{s}} \approx 8.65 \times 10^{18}$, it is impossible to compute the value function by direct computation.

Based on the ideas of Zou, Ahmed, and Sun (2019) for dynamic programming problems with binary state variables, (Zhang & Sun, 2019, Algorithm 3) have developed a stochastic dual dynamic pro-

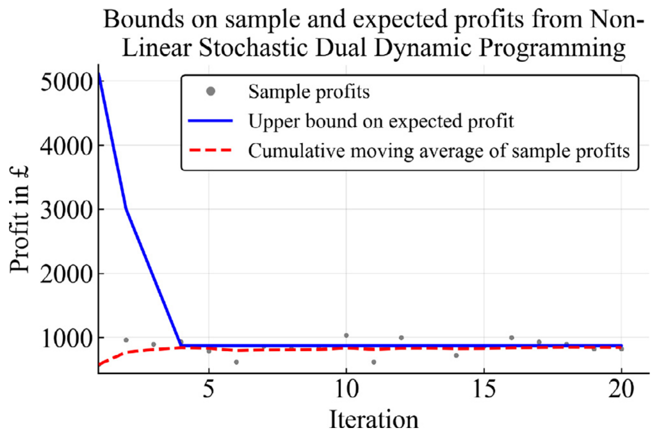


Fig. 3. Plots of sample profits (grey dots), upper bound on expected profit (blue solid line) and cumulative moving average of sample profits, i.e. the stochastic lower bound on expected profit (red dashed line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

gramming algorithm for multi-stage stochastic mixed-integer non-linear optimisation problems. We implement this algorithm for the above problem parameters. The exact details of this algorithm are beyond the scope of this paper and we refer the interested reader to Zhang and Sun (2019).

For this example, it suffices to state that the algorithm produces a value function approximation of the form of the pointwise minimum of hyperplanes in x for all $t \in T$. Since this approximate function is concave extensible in x, the algorithm can only converge to the exact value function if the original value function is also concave extensible. Through an iterative procedure, hyperplanes are added to refine the representation of the value function. Approximation progress across iterations is quantified by means of:

1. The algorithm produces a deterministic upper bound u on the total expected profit, i.e. $u \geq V_1(0)$.
2. The algorithm produces a stochastic lower bound l on the total expected profit, such that $\mathbb{E}l \leq V_1(0)$, which is computed from sample profits obtained in simulations of the DP forward in time, while pricing is based on the approximate value function instead of the (unavailable) exact value function. Note that $\mathbb{E}$ denotes the associated expected value operator. The stochastic lower bound is then computed simply as the cumulative moving average of the sample profits obtained in all previous iterations.

Fig. 3 shows how these bounds converge for the simulated scenario.

The deterministic upper and stochastic lower bounds of the algorithm only converge if the exact value function is concave extensible. Our theoretical analysis, which guarantees concave extensibility, allows us to employ this algorithm and provides theoretical support for its convergent behaviour. A similar conclusion holds for the algorithm in Lebedev, Goulart, and Margellos (2020b), which is derived based on ideas from stochastic dual DP theory (see Shapiro, 2011). Hence, the concave extensibility preserving properties of the DP in (3), even in cases where we cannot numerically verify them due to the prohibitive problem size, open the road to employ techniques derived from stochastic dual DP theory.

Furthermore, since this approach is quite fast – we compute 20 iterations on an i7-8565U CPU with 1.80 gigahertz base frequency and with 16 gigabyte in 1 minute, 28 seconds – it is also possible to update the delivery cost function and hence the terminal condition of the DP using additional information like location of customers that have already placed an order. One could then re-run the DP to obtain approximately optimal prices that have adjusted to the perturbed delivery cost function before the next customer places an order.

As a final remark, we also use the infinite time horizon result from Theorem 1 in the initialisation of this non-linear stochastic dual DP algorithm: The user-defined, initial value function approximation needs to be an upper bound to the exact value function. We can use the result of Theorem 1 for this purpose, since the fixed point provides a non-trivial upper bound on the value function for all time steps $t \in T$ and can easily be implemented, due to its low complexity, i.e. only a single hyperplane is needed.

Further algorithmic developments exploiting our theoretical results are also possible. For example, we develop an algorithm, termed gradient-bounded DP, based specifically on the theoretical properties derived in this paper in Lebedev, Goulart, and Margellos (2020a,b). Furthermore, we conduct an extensive theoretical and numerical case study comparing gradient-bounded DP with the non-linear stochastic dual DP algorithm above and the affine value function approximation algorithm from Yang and Strauss (2017). We refer the interested reader to this comparative study in Lebedev, Margellos, and Goulart (2020c).

6. Conclusions and future work

6.1. Summary of contributions

We have studied the mathematical properties of the value function of a dynamic program modelling the revenue management problem in attended home delivery exactly. We have shown that the recursive dynamic programming mapping has a unique, finite-valued fixed point and concavity-preserving properties. Hence, we have derived our main result stating that – under certain assumptions – for all time steps in the dynamic program, the value function admits a continuous extension, which is a finite-valued, concave function of its state variables. We have illustrated our findings on a low-dimensional numerical example and an industry-sized problem.

6.2. Managerial insights and implications

Monotonicity of optimal prices under multinomial logit customer choice: In our theoretical analysis, we identify conditions under which prices are monotonic in the number of placed orders (see Lemma 5). Furthermore, our numerical examples show that these conditions may even be violated while the monotonicity properties still hold. We verify this directly on a low-dimensional example and indirectly by showing that an algorithm relying on these properties converges for a high-dimensional example. The key managerial insight on the pricing of substitute goods under multinomial logit customer choice is that “optimal” prices may exhibit monotonic behaviour in many cases. Even when monotonic prices are not “optimal” from a theoretical standpoint, they may still perform indistinguishably well to an “optimal”, yet practically unobtainable, pricing policy (for a prohibitively large problem).

Diminishing returns to inventory under multinomial logit customer choice: A second important object for pricing in a dynamic programming framework is the value function describing the expected profit-to-go for all time steps in the booking horizon and all states of orders. Having computed the value function, or an approximation for prohibitively large problems, (approximately) optimal prices can be computed from the marginal value function, i.e. unit order differences of the expected profit-to-go, which represent the opportunity cost of this order. Our characterisation of the value function (Theorem 2 and Proposition 4) implies that there are diminishing marginal returns to inventory, which corresponds to the order capacity of the delivery time slots in our setting. As inventory increases, total expected profit increases sub-proportionally.

One intuitive explanation for this is that no matter how much the maximum order capacity is increased, total expected profit is ultimately limited by the expected number of customer arrivals. Hence, in addition to the *tactical* decisions on dynamic pricing, our insights on the structure of the value function can be used for *strategical* planning of inventory capacities. For example, our analysis allows the computation of an approximate value function. The value of this function at the beginning of the booking horizon and around the zero order state reveals the marginal value of additional inventory. This insight can then help decide on the feasibility of expanding slot capacities.

6.3. Directions for future research

Recent approaches have estimated V_t as an affine function of x for each $t \in T$ (Yang & Strauss, 2017). Based on our result, we believe that closer approximations can be found by pursuing different approximation strategies. One such strategy would be to adapt approximate DP algorithms from stochastic dual dynamic programming, also known as SDDP. The idea is to use a cutting plane algorithm to successively form tighter upper bounds to the value function described as the point-wise minimum of affine functions (Pereira & Pinto, 1991; Shapiro, 2011). One such algorithm is presented in Lebedev et al. (2020a,b).

A second possible direction of future research involves investigating the use of parametric models comprising concave basis functions. This idea can be exploited directly by using the given DP formulation – as suggested in Powell (2007, Section 8.2) – or by reformulating the problem as a linear program – as shown by de Farias and Roy (2003). Note that *a priori* knowledge of concave extensibility of V_t for all $t \in T$ creates some intuitive regularity. Therefore, it can be expected to get good approximations of V_t from a relatively small sample size even with simple models.

Another possible direction would be to adapt techniques that fit convex functions (or equivalently concave functions for our purposes) to multidimensional data. For example, Kim, Lee, Vandenberghe, and Yang (2004) and Magnani and Boyd (2009) show how data can be fitted by a function defined as the maximum of a finite number of affine functions. More sophisticated examples of convex (concave) function fitting techniques include adaptive partitioning (Hannah & Dunson, 2013) and Bayesian non-parametric regression (Hannah & Dunson, 2011).

Finally, another possible direction for future research involves investigating pricing policies for delivery cost terminal conditions which violate the imposed assumptions on increasing opportunity costs and concave extensibility. In some cases, it may be possible to find an approximation to the delivery cost function that satisfies the necessary assumptions and hence, makes it possible to use the results presented in this paper. At the same time, we aim at investigating the potential of relaxing the assumptions on the terminal condition and including additional information like customer location in the delivery cost function.

Acknowledgements

This work was supported by SIA Food Union Management. The authors are grateful for this financial support.

Appendix A. Proof of Lemma 3(i)

It is shown in Dong et al. (2009) that a structurally similar function to f is concave in its variables. We adopt a similar approach – computing the Hessian and showing that it is negative definite – to verify that f is jointly concave in all components of the vector p. We first compute the first-order partial derivatives of f:

$$\frac{\partial f}{\partial p_i} = [r + \beta_d^{-1}(\ln(p_i/p_0) - \beta_c - \beta_i)] + \beta_d^{-1}, \text{ for all } i \in S,$$
$$\frac{\partial f}{\partial p_0} = \sum_s -p_s \beta_d^{-1} p_0^{-1}. \tag{A.1}$$

The second-order partial derivatives are:

$$\frac{\partial^2 f}{\partial p_i^2} = \beta_d^{-1} p_i^{-1}, \text{ for all } i \in S,$$
$$\frac{\partial^2 f}{\partial p_0^2} = \sum_s p_s \beta_d^{-1} p_0^{-2},$$
$$\frac{\partial^2 f}{\partial p_i \partial p_0} = -\beta_d^{-1} p_0^{-1}, \text{ for all } i \neq 0,$$
$$\frac{\partial^2 f}{\partial p_i \partial p_j} = 0, \text{ for all } (i, j) \in S \times S, \text{ such that } i \neq j. \tag{A.2}$$

The resulting Hessian H of f with its second partial derivatives with respect to $\{p_i\}$ for all $i \in S \cup \{0\}$ is:

$$H := \beta_d^{-1} \begin{bmatrix} p_1^{-1} & \dots & 0 & -p_0^{-1} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & p_{\bar{s}}^{-1} & -p_0^{-1} \\ -p_0^{-1} & \dots & -p_0^{-1} & p_0^{-2}\sum_s p_s \end{bmatrix}$$
$$=: \begin{bmatrix} A & B \\ B^\intercal & C \end{bmatrix}, \tag{A.3}$$

where we have defined block sub-matrices $A, B, B^\intercal$ and C of appropriate dimension, such that C is a scalar corresponding to the last entry of H. Note that A is negative definite, because $p_s \geq 0$ for all $s \in S \cup \{0\}$ and $\beta_d^{-1} < 0$. We compute the Schur complement of A in H:

$$C - B^\intercal A^{-1} B = \beta_d^{-1}\left(\sum_s p_s p_0^{-2} - p_0^{-2}\sum_s p_s\right) = 0. \tag{A.4}$$

As a result of $\beta_d < 0$, A is negative definite and as H/A is non-positive, H is negative semi-definite (see e.g. Boyd and Vandenberghe (2004, Appendix A.5.5)). This implies that f is concave in p. □

Appendix B. Proof of Lemma 3(ii)

The proof of Lemma 3(ii) requires several intermediate results which we present in the following sections before returning to the proof of Lemma 3(ii).

B1. Auxiliary function definitions and properties

In this section, we define some auxiliary functions and establish some of their properties that are needed in the subsequent sections. We define $W : \mathbb{R}_+ \mapsto \mathbb{R}_+$ as the inverse function of $f : \mathbb{R}_+ \mapsto \mathbb{R}_+$, such that $f(x) := x\exp(x)$, i.e. implicitly defined through the relationship

$$x = W(x)e^{W(x)}. \tag{B.1}$$

Note that W, as defined above, is the principal branch of the so-called Lambert W function, which is uniquely defined over the non-negative real numbers. To simplify notation in the following proof, we define two more functions: We define $\psi_s : \mathbb{R}_+ \mapsto \mathbb{R}_{++}$ as

$$\psi_s(z) := \exp(\beta_c + \beta_s + \beta_d(z - r) - 1) \tag{B.2}$$

for all $s \in S$. We define the function $\phi : \mathbb{R}_+^{\bar{s}} \mapsto \mathbb{R}_{++}$ as

$$\phi(z) := -\lambda\beta_d^{-1} W\left(\sum_s \psi_s(z_s)\right), \tag{B.3}$$

where z_s indicates the s-th component of z. We can now establish some properties of ϕ that are instrumental for the subsequent proof of Lemma 3(ii).

Lemma 6. *The function ϕ has the following properties:*

(i) *It is decreasing over its domain.*
(ii) *It satisfies the inequality:*

$$\begin{aligned}&\phi(\gamma_t(x+1_{s'}))-\phi(\max\{\gamma_t(x+1_s),\gamma_t(x+1_{s'})\})\\&\quad+\phi(\gamma_t(x+1_s))\\&\geq\phi(\min\{\gamma_t(x+1_s),\gamma_t(x+1_{s'})\})\end{aligned} \tag{B.4}$$

for all $(x,s,s')\in X\times S\times S$, such that $s\neq s'$.

Proof.

(i) It is useful to state the first derivative of W, which is

$$\frac{\mathrm{d}W}{\mathrm{d}y}(y)=\frac{W(y)}{y(1+W(y))}. \tag{B.5}$$

It suffices to show that the first partial derivative of W with respect to the components z_i for all $i\in S$ is negative. To this end, fix any $i\in S$. Setting $y=\sum_s\psi_s(z_s)$, gives

$$\begin{aligned}\frac{\partial\phi}{\partial z_i}(z)&=-\lambda\beta_d^{-1}\frac{\mathrm{d}W}{\mathrm{d}y}(y)\frac{\partial y}{\partial z_i}(z)\\&=-\lambda\beta_d^{-1}\frac{W(\sum_s\psi_s(z_s))}{[\sum_s\psi_s(z_s)](1+W(\sum_s\psi_s(z_s))}\\&\quad\times\beta_d\exp(\beta_c+\beta_i+\beta_d(z_i-r)-1)\\&=-\lambda\frac{W(\sum_s\psi_s(z_s))}{1+W(\sum_s\psi_s(z_s))}\frac{\psi_i(z_i)}{\sum_s\psi_s(z_s)},\end{aligned} \tag{B.6}$$

where the last equality follows from the definition of ψ_i in (B.2). The customer arrival rate $\lambda\in(0,1)$. The first fraction in (B.6) lies in (0,1) as $W(y)\geq0$ for all $y\in\mathrm{dom}(W)$, while the second one lies in (0,1] as $\psi_s(y)>0$ for all $y\in\mathrm{dom}(\psi_s)$ for all $s\in S$ and $\bar{s}\geq1$. Therefore, $\partial\phi/\partial z_i(z)\in(-1,0)$ for all $i\in S$ and hence, it is negative.

(ii) Fix any $i\in S$. Let $\alpha_{i,t}:=\max\{\gamma_{i,t}(x+1_s),\gamma_{i,t}(x+1_{s'})\}$ and $\beta_{i,t}:=\min\{\gamma_{i,t}(x+1_s),\gamma_{i,t}(x+1_{s'})\}$ and notice that $\alpha_{i,t}\geq\beta_{i,t}$. Let us distinguish two cases.
Case I: Suppose that $\gamma_{i,t}(x+1_s)\geq\gamma_{i,t}(x+1_{s'})$. For all $j\in S, j\neq i$, define $\epsilon^{\alpha}_{j,t}:=\gamma_{j,t}(x+1_s)$, $\epsilon^{\beta}_{j,t}:=\gamma_{j,t}(x+1_{s'})$ and $\epsilon^{\alpha\beta}_{j,t}:=\max\{\gamma_{j,t}(x+1_s),\gamma_{j,t}(x+1_{s'})\}$. Under these assignments, the left-hand side of (B.4) can be equivalently written as

$$\begin{aligned}&\phi(\gamma_t(x+1_{s'}))-\phi(\max\{\gamma_t(x+1_s),\gamma_t(x+1_{s'})\})\\&\quad+\phi(\gamma_t(x+1_s))\\&\quad=\phi(\epsilon^{\beta}_{1,t},\dots,\beta_{i,t},\dots,\epsilon^{\beta}_{\bar{s},t})-\phi(\epsilon^{\alpha\beta}_{1,t},\dots,\alpha_{i,t},\dots,\epsilon^{\alpha\beta}_{\bar{s},t})\\&\quad+\phi(\epsilon^{\alpha}_{1,t},\dots,\alpha_{i,t},\dots,\epsilon^{\alpha}_{\bar{s},t}).\end{aligned} \tag{B.7}$$

Define the scalar function $f_\theta:A\mapsto\mathbb{R}$, where A contains all real numbers no smaller than $\beta_{i,t}$ and $\theta:=\{\{\epsilon^{\alpha}_{j,t},\epsilon^{\beta}_{j,t},\epsilon^{\alpha\beta}_{j,t}\}_{j\neq i},\beta_{i,t}\}$, such that

$$\begin{aligned}f_\theta(\alpha_{i,t})=&\ \phi(\epsilon^{\beta}_{1,t},\dots,\beta_{i,t},\dots,\epsilon^{\beta}_{\bar{s},t})\\&-\phi(\epsilon^{\alpha\beta}_{1,t},\dots,\alpha_{i,t},\dots,\epsilon^{\alpha\beta}_{\bar{s},t})\\&+\phi(\epsilon^{\alpha}_{1,t},\dots,\alpha_{i,t},\dots,\epsilon^{\alpha}_{\bar{s},t}).\end{aligned} \tag{B.8}$$

Consider the derivative of f_θ, given by

$$\frac{\mathrm{d}f_\theta}{\mathrm{d}\alpha_{i,t}}(\alpha_{i,t})=-\frac{\partial\phi}{\partial\alpha_{i,t}}(z^{\alpha\beta})+\frac{\partial\phi}{\partial\alpha_{i,t}}(z^{\alpha}), \tag{B.9}$$

where we have defined $z^{\alpha\beta}:=[\epsilon^{\alpha\beta}_{1,t},\dots,\alpha_{i,t},\dots,\epsilon^{\alpha\beta}_{\bar{s},t}]$ as well as $z^{\alpha}:=[\epsilon^{\alpha}_{1,t},\dots,\alpha_{i,t},\dots,\epsilon^{\alpha}_{\bar{s},t}]$. We compute the derivative of ϕ from the first equation in (B.6) to arrive at

$$\begin{aligned}\frac{\mathrm{d}f_\theta}{\mathrm{d}\alpha_{i,t}}(\alpha_{i,t})&=\lambda\beta_d^{-1}\frac{\mathrm{d}W}{\mathrm{d}y}(y^{\alpha\beta})\frac{\partial y^{\alpha\beta}}{\partial\alpha_{i,t}}(z^{\alpha\beta})\\&\quad-\lambda\beta_d^{-1}\frac{\mathrm{d}W}{\mathrm{d}y}(y^{\alpha})\frac{\partial y^{\alpha}}{\partial\alpha_{i,t}}(z^{\alpha}),\end{aligned} \tag{B.10}$$

where $y^{\alpha\beta}:=\sum_s\psi_s(z^{\alpha\beta}_s)$ and $y^{\alpha}:=\sum_s\psi_s(z^{\alpha}_s)$. Substituting for the derivatives of $y^{\alpha\beta}$ and y^{α} and simplifying yields

$$\frac{\mathrm{d}f_\theta}{\mathrm{d}\alpha_{i,t}}(\alpha_{i,t})=\lambda\psi_i(\alpha_{i,t})\left[\frac{\mathrm{d}W}{\mathrm{d}z}(z^{\alpha\beta})-\frac{\mathrm{d}W}{\mathrm{d}z}(z^{\alpha})\right]\geq0. \tag{B.11}$$

The inequality follows from noting that $z^{\alpha}\geq z^{\alpha\beta}$ and that W has a negative second derivative over its domain. We conclude that f_θ is non-decreasing in $\alpha_{i,t}$, which means that f_θ is non-increasing by decreasing $\alpha_{i,t}$ to its minimum value $\alpha_{i,t}=\beta_{i,t}$. Repeating this minimisation for all $i\in S$, we obtain the following bound:

$$\begin{aligned}&\phi(\gamma_t(x+1_{s'}))-\phi(\max\{\gamma_t(x+1_s),\gamma_t(x+1_{s'})\})+\phi(\gamma_t(x+1_s))\\&\geq\phi(\min\{\gamma_t(x+1_s),\gamma_t(x+1_{s'})\})-\phi(\min\{\gamma_t(x+1_s),\gamma_t(x+1_{s'})\})\\&\quad+\phi(\min\{\gamma_t(x+1_s),\gamma_t(x+1_{s'})\})\\&=\phi(\min\{\gamma_t(x+1_s),\gamma_t(x+1_{s'})\}),\end{aligned} \tag{B.12}$$

as required.
Case II: Suppose that the roles of s and s' are now reversed, i.e. $\gamma_{i,t}(x+1_{s'})\geq\gamma_{i,t}(x+1_s)$. Via symmetric arguments, we reach the same conclusion as in Case I.
As both cases reach the same conclusion and collectively exhaust all possibilities, this concludes our proof and shows that (B.4) holds.

□

B2. Proof of Proposition 4(i)

We start by reformulating the DP in (3) in terms of ϕ. Fix any $t\in T\setminus\{1\}\cup\{\bar{t}+1\}$. The unique optimisers of the unconstrained optimisation problem at time $t-1$, denoted by $d^*_s(x)$ for all $(x,s)\in X\times S$, is given by Yang and Strauss (2017) based on the development of Dong et al. (2009) as

$$d^*_s(x)=\gamma_{s,t}(x)-r-\beta_d^{-1}h_t(x) \tag{B.13}$$

for all $s\in S$, where $h_t(x)$ is the unique solution of

$$(h_t(x)-1)\exp(h_t(x))=\sum_s\exp(\beta_c+\beta_s+\beta_d(\gamma_{s,t}(x)-r)). \tag{B.14}$$

We rewrite (B.14) equivalently as

$$\begin{aligned}&\iff(h_t(x)-1)\exp(h_t(x)-1)=\sum_s\exp(\beta_c+\beta_s+\beta_d(\gamma_{s,t}(x)\\&\qquad\qquad-r)-1)\\&\iff(h_t(x)-1)\exp(h_t(x)-1)=\sum_s\psi_s(\gamma_{s,t}(x)),\end{aligned} \tag{B.15}$$

where we have used ψ_s from (B.2). By the definition of W, we obtain

$$h_t(x)=1+W\left(\sum_s\psi_s(\gamma_{s,t}(x))\right). \tag{B.16}$$

Now, we can substitute (B.16) into (B.13):

$$d^*_s(x)=\gamma_{s,t}(x)-r-\beta_d^{-1}\left[1+W\left(\sum_s\psi_s(\gamma_{s,t}(x))\right)\right]. \tag{B.17}$$

Finally we can substitute (B.17) into the unconstrained version of (3) to obtain

$$\begin{aligned}\mathcal{T}V_t(x) &= \sum_s p_s(d^*(x))\left\{r + \gamma_{s,t}(x) - r - \beta_d^{-1}\left[1 + W\left(\sum_{s'}\psi_s(\gamma_{s',t}(x))\right)\right]\right. \\ &\quad \left. - \gamma_{s,t}(x)\right\} + V_t(x) \\ &= \sum_s p_s(d^*(x))\left\{-\beta_d^{-1}\left[1 + W\left(\sum_{s'}\psi_s(\gamma_{s',t}(x))\right)\right]\right\} + V_t(x), \quad \text{(B.18)}\end{aligned}$$

where we have defined $d^*(x) := [d_1^*(x), d_2^*(x), \ldots, d_{\bar{s}}^*(x)]^\intercal$. We now substitute the customer choice model p evaluated at the optimiser $d^*(x)$ into (B.18):

$$\begin{aligned}\mathcal{T}V_t(x) = &\sum_s \frac{\lambda \exp(\beta_c + \beta_s + \beta_d d_s^*(x))}{\sum_{s''}\exp(\beta_c + \beta_{s''} + \beta_d d_{s''}^*(x)) + 1} \\ &\times\left\{-\beta_d^{-1}\left[1 + W\left(\sum_{s'}\psi_s(\gamma_{s',t}(x))\right)\right]\right\} + V_t(x). \quad \text{(B.19)}\end{aligned}$$

Note that – using the definitions of d_s^* and h_t – the following relationship holds:

$$\begin{aligned}&\sum_s \exp(\beta_c + \beta_s + \beta_d d_s^*(x)) \\ &= \sum_s \exp\left(\beta_c + \beta_s + \beta_d\left(\gamma_{s,t}(x) - r - \beta_d^{-1}h_t(x)\right)\right) \\ &= \sum_s \exp\left(\beta_c + \beta_s + \beta_d(\gamma_{s,t}(x) - r)\right)\exp(-h_t(x)) \\ &= (h_t(x) - 1)\exp(h_t(x))\exp(-h_t(x)) \\ &= h_t(x) - 1, \quad \text{(B.20)}\end{aligned}$$

where the third equality follows from (B.14). Substituting (B.20) into (B.19), we obtain

$$\mathcal{T}V_t(x) = \lambda\frac{h_t(x) - 1}{h_t(x)}\left\{-\beta_d^{-1}\left[1 + W\left(\sum_s \psi_s(\gamma_{s,t}(x))\right)\right]\right\} + V_t(x). \quad \text{(B.21)}$$

Substituting for $h_t(x)$ using (B.16) yields the desired expression:

$$\begin{aligned}\mathcal{T}V_t(x) &= \frac{\lambda W(\sum_s \psi_s(\gamma_{s,t}(x)))}{1 + W(\sum_s \psi_s(\gamma_{s,t}(x)))}\left\{-\beta_d^{-1}\left[1 + W\left(\sum_s \psi_s(\gamma_{s,t}(x))\right)\right]\right\} \\ &\quad + V_t(x) \\ &= -\lambda\beta_d^{-1}W\left(\sum_s \psi_s(\gamma_{s,t}(x))\right) + V_t(x) \\ &= \phi(\gamma_t(x)) + V_t(x), \quad \text{(B.22)}\end{aligned}$$

where the last inequality follows from the definition of ϕ in (B.3). By Assumption 3, the terminal condition of the DP satisfies the property of increasing opportunity costs.

We can now prove Proposition 4(i) by showing that a monotonic mapping exists between $\gamma_{s,t-1}(x + 1_{s'}) - \gamma_{s,t-1}(x)$ and $\gamma_{s,t}(x + 1_{s'}) - \gamma_{s,t}(x)$ for all $(x, s, s', t) \in X \times F(x) \times F(x) \times (T \cup \{\bar{t} + 1\} \setminus \{1\})$.

To this end, fix any $(x, s, t) \in X \times F(x) \times (T \cup \{\bar{t} + 1\} \setminus \{1\})$. By using the definition of opportunity costs and (B.22) we can write the opportunity cost in state x with respect to an arbitrary slot $s \in S$ at time $t - 1$ as

$$\gamma_{s,t-1}(x) = \gamma_{s,t}(x) + \phi(\gamma_t(x)) - \phi(\gamma_t(x + 1_s)). \quad \text{(B.23)}$$

Fix any $s' \in F(x)$, such that $s' \neq s$. To prove the theorem, we require

$$\gamma_{s,t-1}(x + 1_{s'}) - \gamma_{s,t-1}(x) > 0 \quad \text{(B.24)}$$

for all $t \in T \cup \{\bar{t} + 1\} \setminus \{1\}$. Substitute (B.23) into the left-hand side of (B.24) to obtain

$$\begin{aligned}&\gamma_{s,t-1}(x + 1_{s'}) - \gamma_{s,t-1}(x) \\ &= \gamma_{s,t}(x + 1_{s'}) - \gamma_{s,t}(x) + \phi(\gamma_t(x + 1_{s'})) \\ &\quad -\phi(\gamma_t(x + 1_s + 1_{s'})) - \phi(\gamma_t(x)) + \phi(\gamma_t(x + 1_s)). \quad \text{(B.25)}\end{aligned}$$

First note that $\gamma_t(x + 1_s + 1_{s'}) \geq \gamma_t(x + 1_s)$ and $\gamma_t(x + 1_s + 1_{s'}) \geq \gamma_t(x + 1_{s'})$. As $\phi(\gamma_t(x + 1_s + 1_{s'}))$ is decreasing in its argument by Lemma 6(i) and since it is subtracted on the right-hand side of the above equation, we can create the following lower bound:

$$\begin{aligned}&\gamma_{s,t-1}(x + 1_{s'}) - \gamma_{s,t-1}(x) \\ &\geq \gamma_{s,t}(x + 1_{s'}) - \gamma_{s,t}(x) + \phi(\gamma_t(x + 1_{s'})) \\ &\quad -\phi(\max\{\gamma_t(x + 1_s), \gamma_t(x + 1_{s'})\}) - \phi(\gamma_t(x)) + \phi(\gamma_t(x + 1_s)). \\ &\quad \text{(B.26)}\end{aligned}$$

Using Lemma 6(ii), we bound (B.26) from below by

$$\begin{aligned}\gamma_{s,t-1}(x + 1_{s'}) - \gamma_{s,t-1}(x) \geq\ &\phi(\min\{\gamma_t(x + 1_s), \gamma_t(x + 1_{s'})\}) \\ &- \phi(\gamma_t(x)) \\ &+ \gamma_{s,t}(x + 1_{s'}) - \gamma_{s,t}(x), \quad \text{(B.27)}\end{aligned}$$

where the minimum is taken element-wise. Since ϕ is both positive by definition, we construct a lower bound on (B.27) by dropping the $\phi(\min\{\gamma_t(x + 1_s), \gamma_t(x + 1_{s'})\})$ term on the right-hand side. Furthermore, since ϕ is decreasing in its argument by Lemma 6(i), we create another lower bound on (B.27) by setting $\gamma_t(x) = 0$. Hence, we obtain

$$\begin{aligned}&\gamma_{s,t-1}(x + 1_{s'}) - \gamma_{s,t-1}(x) \\ &> -\phi(0) + \gamma_{s,t}(x + 1_{s'}) - \gamma_{s,t}(x) \\ &= \lambda\beta_d^{-1}W\left(\sum_s \psi_s(0)\right) + \gamma_{s,t}(x + 1_{s'}) - \gamma_{s,t}(x). \quad \text{(B.28)}\end{aligned}$$

Since only β_d^{-1} is negative, $\lambda\beta_d^{-1}W(\sum_s \psi_s(0)) < 0$, independent of the choice of (s, s', t, x). Therefore, the above inequality describes a monotonic mapping from $\gamma_{s,t}(x + 1_s') - \gamma_{s,t}(x)$ to $\gamma_{s,t-1}(x + 1_s') - \gamma_{s,t-1}(x)$ for all $(s, s', t - 1, x) \in F(x) \times F(x) \times T \times X$, such that $s \neq s'$. The bound on $\gamma_{s,t}(x + 1_s') - \gamma_{s,t}(x)$ will therefore decrease as t decreases. Hence, $\gamma_{s,t}(x + 1_s') - \gamma_{s,t}(x)$ will be minimal at $t = 1$. Using the monotonicity of this mapping, we can find a λ for which $\gamma_{s,1}(x + 1_s') - \gamma_{s,1}(x) > 0$ by repetitively applying the above equation starting from the terminal condition at $t = \bar{t} + 1$ to obtain

$$\begin{aligned}\gamma_{s,1}(x + 1_{s'}) - \gamma_{s,1}(x) \geq\ &\bar{t}\lambda\beta_d^{-1}W\left(\sum_s \psi_s(0)\right) + \gamma_{s,\bar{t}+1}(x + 1_{s'}) \\ &- \gamma_{s,\bar{t}+1}(x), \quad \text{(B.29)}\end{aligned}$$

where the right-hand side is positive if

$$\begin{aligned}&0 < \bar{t}\lambda\beta_d^{-1}W\left(\sum_s \psi_s(0)\right) + \gamma_{s,\bar{t}+1}(x + 1_{s'}) - \gamma_{s,\bar{t}+1}(x) \\ &\iff \lambda < -\beta_d\frac{\gamma_{s,\bar{t}+1}(x + 1_{s'}) - \gamma_{s,\bar{t}+1}(x)}{\bar{t}W(\sum_s \psi_s(0))} \quad \text{(B.30)}\end{aligned}$$

for all $(x, s, s') \in X \times F(x) \times F(x)$, such that $s \neq s'$. The right-hand side of the above expression is strictly positive, because opportunity costs are increasing at the terminal condition by Assumption 3 and therefore, the numerator of the fraction is positive, W only takes positive values, $\bar{t} > 0$ and $\beta_d < 0$. As both X and $F(x)$ are finite sets, a small enough $\lambda > 0$ can be found that satisfies all inequalities described by (B.30). Therefore, a $\lambda > 0$ exists, such that $\gamma_{s,t}(x + 1_{s'}) > \gamma_{s,t}(x)$ for all $(x, s, s', t) \in X \times F(x) \times F(x) \times T$, such that $s \neq s'$.

B3. Proof of Proposition 4(ii)

Fix any $(x, t) \in X \times T$. Define $u := \gamma_t(x)$ and consider the function $w : \mathbb{R}_+^{\bar{s}} \mapsto \mathbb{R}^{\bar{s}}$ defined by

$$w(u) := u - r - \beta_d^{-1}\left[1 + W\left(\sum_s \psi_s(u_s)\right)\right]. \quad \text{(B.31)}$$

In comparison with (B.17), notice that $w(u)$ is mathematically identical to $d^*(x) = [d_1^*(x), d_2^*(x), \ldots, d_{\bar{s}}^*(x)]$. Furthermore, W and ψ_s for all $s \in S$ are invertible functions. In particular, their inverses in the domain of interest are

$$x = W(x)\exp(W(x)) \text{ and}$$
$$u = (\ln(\psi_s(u)) - \beta_c - \beta_s + 1)\beta_d^{-1} + r \quad \text{(B.32)}$$

for all $s \in S$. Therefore, w is a composition of invertible functions and hence, the mapping between w and u is bijective. Therefore, the mapping between $d^*(x)$ and $\gamma_t(x)$, equivalent to the mapping between w and u, is bijective.

If we constrain $d^*(x)$ to $D \subset \mathbb{R}^{\bar{s}}$, we can conclude that there still exists a bijective mapping between D and (an unknown) $\Gamma \subset \mathbb{R}_+^{\bar{s}}$ corresponding to the range of values that $\gamma_t(x)$ can take in this constrained scenario. Conversely, this means that no matter which $d^*(x) \in D$ maximises the *constrained* stage optimisation problem, there exists $\hat{\gamma}_t(x) \in \Gamma$, which is linked to the same $d^*(x)$ in the *unconstrained* problem. In other words, there exists $\hat{\gamma}_t(x) \in \Gamma$, which produces the same $V_{t-1}(x)$ in the *unconstrained* problem as $\gamma_t(x)$ does in the *constrained* problem. Due to this equivalence, the following statement is a necessary and sufficient condition for Proposition 4(i) to hold: Evaluating the *unconstrained* problem at $\hat{\gamma}_t(x)$, i.e. $V_{t-1}(x) = \phi(\hat{\gamma}_t(x)) + V_t(x)$ for all $x \in X$, there exists a sufficiently small $\lambda > 0$ that yields a value function at time $t-1$, whose opportunity cost is increasing in x. From Proposition 4(i), we know that this statement holds true for all opportunity costs $\hat{\gamma}_t(x) \in \mathbb{R}_+^{\bar{s}}$ under Assumption 3. Since $\hat{\gamma}_t(x) \in \Gamma \subset \mathbb{R}_+^{\bar{s}}$, there exists a sufficiently small $\lambda > 0$ such that the opportunity cost will also be increasing in the *constrained* DP.

B4. Completing the proof of Lemma 3(ii)

From Definition 2(b), g_t is concave extensible in (x, p) if

$$g_t(x, p) \geq \sum_{q \in Q} \mu_q g_t\left(q^{(x)}, q^{(p)}\right) \quad \text{(B.33)}$$

for all $(x, p) \in X \times P$ and for all enclosing sets Q, such that $[x, p]^\intercal = \sum_{q \in Q} \mu_q \left[q^{(x)}, q^{(p)}\right]^\intercal$. We show that this inequality holds by starting from the right-hand side and substituting for $g_t(q^{(x)}, q^{(p)})$:

$$\begin{aligned}\sum_{q \in Q} \mu_q g_t(q^{(x)}, q^{(p)}) &= \sum_{q \in Q} \mu_q \left[\sum_s q_s^{(p)}\left\{V_t(q^{(x)} + 1_s) - V_t(q^{(x)})\right\} + V_t(q^{(x)})\right] \\ &= \sum_{q \in Q} \mu_q \left[\sum_s q_s^{(p)} V_t(q^{(x)} + 1_s) + \left(1 - \sum_s q_s^{(p)}\right) V_t(q^{(x)})\right], \quad \text{(B.34)}\end{aligned}$$

where we note that each of the summed terms in square brackets is a convex combination on the set $A(q^{(x)}) := \{q^{(x)} \cup \{q^{(x)} + 1_s\}_{s \in F(q^{(x)})}\}$, where the set of indices s is $F(q^{(x)})$ instead of S, since $q_s^{(p)} = 0$ for all $s \notin F(q^{(x)})$, i.e. assigning zero transition probability to infeasible slots. To show that (B.33) holds, we now derive the supporting result that there exists a small enough $\lambda > 0$, such that the concave closure $\tilde{V}_t(y)$, evaluated at any point $y \in X$, is a hyperplane on the set $A(y)$, which will make it possible to simplify (B.34) further.

Fix any $y \in X$ and consider the unit hypercube in the positive direction of y, which we define as $B(y) := \left\{z \mid y \leq z \leq y + \sum_{s \in F(y)} 1_s\right\}$. We will show that only points in $A(y)$ form the concave closure around y by demonstrating that for all $y' \in B(y) \setminus A(y)$, the line segment between $(y, V_t(y))$ and $(y', V_t(y'))$ lies below a second line segment between two other points $(z, V_t(z))$ and $(z', V_t(z'))$ for some $(z, z') \in B(y) \times B(y)$, i.e. we will show that there exists some $(z, z') \in B(y) \times B(y)$, such that

$$\alpha V_t(y) + (1 - \alpha)V_t(y') < \beta V_t(z) + (1 - \beta)V_t(z'), \quad \text{(B.35)}$$

where $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$, such that $\alpha y + (1 - \alpha)y' = \beta z + (1 - \beta)z'$ for all $y' \in B(y) \setminus A(y)$. This means that the line segments cannot be part of the concave closure $\tilde{V}_t$. We show this result by induction on $|F(y)|$. Consider the base case when $|F(y)| = 1$, then (B.35) is trivially satisfied as there is only a single element $s \in F(y)$, meaning that the set $B(y) \setminus A(y) = \emptyset$. Let $n := |F(y)|$. Suppose by means of an induction hypothesis that (B.35) holds for all cardinalities of $F(y)$ up to and including $n - 1$. Then the only line segment that we need to consider is the one connecting y and $y' = y + \sum_{s \in F(y)} 1_s$, because otherwise, we are in a lower-dimensional case, for which (B.35) holds by the induction hypothesis. For this choice of (y, y'), we can find a quadruple (z, z', α, β) that satisfies (B.35) by repetitively invoking Proposition 4(ii) as follows: There exists a sufficiently small $\lambda > 0$, such that

$$\begin{aligned}&\gamma_1(y) \\ &< \gamma_1(y + 1_2) \\ &< \gamma_1(y + 1_2 + 1_3) \\ &\;\vdots \\ &< \gamma_1\left(y + \textstyle\sum_{s \in \sigma} 1_s\right),\end{aligned} \quad \text{(B.36)}$$

where we have defined $\sigma := F(y) \setminus \{1\}$, which – strictly speaking – depends on y and $s = 1$, but we neglect this to ease notation. We can expand the first and last line of the above chained inequality by using the definition of opportunity costs:

$$\begin{aligned}&\gamma_1(y) < \gamma_1\left(y + \sum_{s \in \sigma} 1_s\right) \\ \iff &V_t(y) - V_t(y + 1_1) < V_t\left(y + \sum_{s \in \sigma} 1_s\right) - V_t\left(y + \sum_{s \in F(y)} 1_s\right) \\ \iff &V_t(y) - V_t(y + 1_1) < V_t\left(y + \sum_{s \in \sigma} 1_s\right) - V_t(y'),\end{aligned} \quad \text{(B.37)}$$

where we have used $y' = y + \sum_{s \in F(y)} 1_s$ as previously. Rearranging and multiplying both sides by $1/2$ yields

$$\iff \frac{1}{2}V_t(y) + \frac{1}{2}V_t(y') < \frac{1}{2}V_t\left(y + \sum_{s \in \sigma} 1_s\right) + \frac{1}{2}V_t(y + 1_1). \quad \text{(B.38)}$$

Comparing with (B.35), we see that $\alpha = \beta = 1/2$, $z = y + \sum_{s \in \sigma} 1_s$ and $z' = y + 1_1$. To illustrate this for the case when $n = 3$, we plot the points $y, y', y + \sum_{s \in \sigma} 1_s$ and $y + 1_1$ as well as the line segments between the former and latter pair of points in Fig. B.4 below.

Since we have found a quadruple (z, z', α, β) that satisfies (B.35), we conclude that no $y' \in B(y) \setminus A(y)$ can be part of the concave closure $\tilde{V}_t$ for all $y \in B(y)$ and for all n. Therefore, only points in $A(y)$ can be part of the concave closure $\tilde{V}_t$ for all $y \in B(y)$ and, in fact, they all are, since the $|F(y)| + 1$ points in $A(y)$ uniquely define the support vectors of a hyperplane through $\{(y, V_t(y))\}_{y \in A(y)}$, i.e. each $y \in A(y)$ giving rise to one linearly independent equality constraint, such that all $|F(y)|$ gradients and the offset of the hyperplane are uniquely defined. This holds true for unit hypercubes $B(y)$ for all $y \in X$ and hence, there exists a small enough $\lambda > 0$, such that $\tilde{V}_t$ is a hyperplane on the set $A(y)$ for all $y \in X$.

Returning to (B.34), we notice that the expression in square brackets is a convex combination on the set $A(q^{(x)})$, which means that $\tilde{V}_t$ is a hyperplane on this set. Hence, we can rewrite (B.34) with equality as

$$\sum_{q \in Q} \mu_q g_t(q^{(x)}, q^{(p)}) = \sum_{q \in Q} \mu_q \tilde{V}_t\left(\sum_s q_s^{(p)} 1_s + q^{(x)}\right)$$

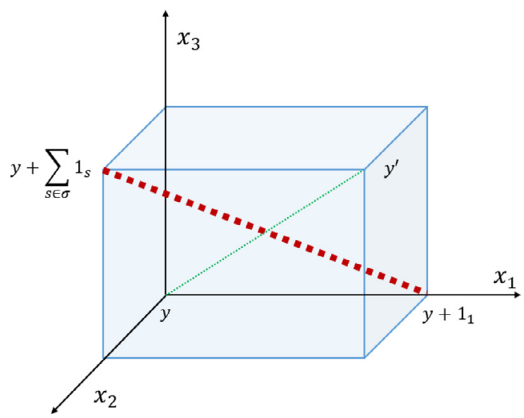


Fig. B.4. Green, dotted line segment between y and y' and red, dotted line segment between $y + \sum_{s\in\sigma} 1_s$ and $y + 1_1$. By (B.38) we have that at the intersection of the two lines, the interpolation of V_t between the two extreme points of the red line lies above the interpolation of V_t between the two extreme points of the green line. The light blue, solid line box spans all the points in $B(y)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$\begin{aligned}&\leq \tilde{V}_t\left(\sum_{q\in Q}\mu_q\left[\sum_s q_s^{(p)}1_s + q^{(x)}\right]\right)\\ &=\tilde{V}_t\left(\sum_s 1_s\sum_{q\in Q}\mu_q q_s^{(p)} + x\right)\\ &=\tilde{V}_t\left(\sum_s 1_s p_s + x\right)\\ &=\sum_s p_s\{V_t(x+1_s) - V_t(x)\} + V_t(x)\\ &=g_t(x,p),\end{aligned} \tag{B.39}$$

where the second-last equality follows from the observation that the convex combination of V_t is evaluated on a set $A(x)$, for which $\tilde{V}_t$ is again a hyperplane, and the inequality is obtained by noticing that $\tilde{V}_t$ is concave in x and therefore, Jensen's inequality holds. From the above derivation, we conclude that $g_t(x,p) \geq \sum_{q\in Q}\mu_q g_t(q^{(x)}, q^{(p)})$, as required. Therefore, there exists a sufficiently small $\lambda > 0$, for which g_t is concave extensible in (x,p) if V_t is concave extensible in x. □

Appendix C. Proof of Lemma 5

Fix any $t \in T$. We prove the result by first showing that opportunity costs γ_t are non-decreasing in x and then showing that optimal prices exhibit the desired monotonicity property with respect to opportunity costs.

By Proposition 4(i), the inequality $\gamma_{s,t}(x+1_{s'}) > \gamma_{s,t}(x)$, for all $(x,s,s') \in X\times F(x)\times F(x)$, such that $s \neq s'$, holds in the unconstrained DP. We now show that this inequality also holds for the case when $s = s'$ with non-strict inequality. By Theorem 2, V_t is concave extensible in x and hence

$$2V_t(x+1_s) \geq V_t(x) + V_t(x+1_s+1_s) \tag{C.1a}$$

$$\begin{aligned}\iff V_t(x+1_s) - V_t(x+1_s+1_s) \geq V_t(x) - V_t(x+1_s)\\ \iff \gamma_{s,t}(x+1_s) \geq \gamma_{s,t}(x),\end{aligned} \tag{C.1b}$$

for all $(x,s) \in X\times F(x)$, where (C.1)(a) follows directly from the definition of concavity and the fact that V_t and its concave closure $\tilde{V}_t$ agree on $\{x, x+1_s, x+1_s+1_s\}$. Note that (C.1)(b) follows from the definition of $\gamma_{s,t}$. Taking both cases together, we conclude that $\gamma_{s,t}(x+1_{s'}) \geq \gamma_{s,t}(x)$, for all $(x,s,s') \in X\times F(x)\times F(x)$. Hence opportunity costs are monotonically non-decreasing in x.

We now show that the desired monotonicity property holds with respect to opportunity costs, which in turn implies that the property holds with respect to x. Recall from (B.17) that

$$d_s^*(x) = \gamma_{s,t}(x) - r - \beta_d^{-1}\left[1 + W\left(\sum_{s'}\psi_{s'}(\gamma_{s',t}(x))\right)\right], \tag{C.2}$$

where W and ψ_s, for all $s \in S$ are defined in B.1. Let $u := \gamma_t(x)$. By (B.6), we have that

$$\frac{\partial d_s^*}{\partial u_{s'}} = \begin{cases}1 - \frac{W(\sum_s\psi_s(u_s))}{1+W(\sum_s\psi_s(u_s))}\frac{\psi_s(u_s)}{\sum_s\psi_s(u_s)}, & \text{for } s = s'\\ -\frac{W(\sum_s\psi_s(u_s))}{1+W(\sum_s\psi_s(u_s))}\frac{\psi_{s'}(u_{s'})}{\sum_{s'}\psi_s(u_s)}, & \text{for } s \neq s'.\end{cases} \tag{C.3}$$

The value of the product of the fractions lies in (0,1), since the value of W is non-negative and the value of ψ_s is positive. Hence, we have

$$\frac{\partial d_s^*}{\partial u_s} > 0 \quad \text{and} \quad \frac{\partial d_s^*}{\partial u_{s'}} < 0, \text{ for all } s' \in S\setminus\{s\}. \tag{C.4}$$

Noting that $u = \gamma_t(x)$ and that $\gamma_t(x)$ is monotonically non-decreasing in x yields the desired property. □

References

Akçay, Y., Natarajan, H. P., & Xu, S. H. (2010). Joint dynamic pricing of multiple perishable products under consumer choice. *Management Science, 56*(8), 1345–1361. https://doi.org/10.1287/mnsc.1100.1178.

Asdemir, K., Jacob, V. S., & Krishnan, R. (2009). Dynamic pricing of multiple home delivery options. *European Journal of Operational Research, 196*(1), 246–257. https://doi.org/10.1016/j.ejor.2008.03.005.

Bertsekas, D. P. (2012). *Dynamic Programming and Optimal Control, Vol. II* (4th). Athena Scientific.

Bertsimas, D., & Weismantel, R. (2005). *Optimization over Integers*.

Boyd, S., & Vandenberghe, L. (2004). *Convex Optimization*. New York, NY, USA: Cambridge University Press.

Campbell, A. M., & Savelsbergh, M. (2006). Incentive schemes for attended home delivery services. *Transportation Science, 40*(3), 327–341. https://doi.org/10.1287/trsc.1050.0136.

Chen, M., & Chen, Z.-L. (2015). Recent developments in dynamic pricing research: Multiple products, competition, and limited demand information. *Production and Operations Management, 24*(5), 704–731. https://doi.org/10.1111/poms.12295.

Dong, L., Kouvelis, P., & Tian, Z. (2009). Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management, 11*(2), 317–339. https://doi.org/10.1287/msom.1080.0221.

de Farias, D. P., & Roy, B. V. (2003). The linear programming approach to approximate dynamic programming. *Operations Research, 51*(6), 850–865. https://doi.org/10.1287/opre.51.6.850.24925.

Food Marketing Institute (2018). Digital shopper. https://www.fmi.org/digital-shopper/. Accessed 4 October 2019.

Gallego, G., & van Ryzin, G. (1994). Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science, 40*(8), 999–1020. https://doi.org/10.1287/mnsc.40.8.999.

Hannah, L. A., & Dunson, D. B. (2011). Bayesian nonparametric multivariate convex regression. https://arxiv.org/abs/1109.0322.

Hannah, L. A., & Dunson, D. B. (2013). Multivariate convex regression with adaptive partitioning. *J. Mach. Learn. Res., 14*(1), 3261–3294.

Kim, J., Lee, J., Vandenberghe, L., & Yang, C.-K. K. (2004). Techniques for improving the accuracy of geometric-programming based analog circuit design optimization. In *Ieee/acm international conference on computer aided design, 2004. iccad-2004.* (pp. 863–870). https://doi.org/10.1109/ICCAD.2004.1382695.

Lebedev, D., Goulart, P., & Margellos, K. (2019). A concave value function extension for the dynamic programming approach to revenue management in attended home delivery. In *18th european control conference, ECC 2019, naples, italy, june 25-28, 2019* (pp. 999–1004). https://doi.org/10.23919/ECC.2019.8795697.

Lebedev, D., Goulart, P., & Margellos, K. (2020a). Gradient-bounded dynamic programming for submodular and concave extensible value functions with probabilistic performance guarantees. *Technical Report*. Submitted for peer review at the Automatica journal. Available: https://arxiv.org/pdf/2006.02910.pdf.

Lebedev, D., Goulart, P., & Margellos, K. (2020b). Gradient-bounded dynamic programming with submodular and concave extensible value functions. In *21st ifac world congress 2020*. Available: https://arxiv.org/pdf/2005.11213.pdf.

Lebedev, D., Margellos, K., & Goulart, P. (2020c). Approximate dynamic programming for delivery time slot pricing: a sensitivity analysis. *Technical Report*. https://arxiv.org/pdf/2008.00780.pdf.

Magnani, A., & Boyd, S. P. (2009). Convex piecewise-linear fitting. *Optimization and Engineering, 10*(1), 1–17. https://doi.org/10.1007/s11081-008-9045-3.

Meissner, J., & Strauss, A. (2012). Network revenue management with inventory-sensitive bid prices and customer choice. *European Journal of Operational Research, 216*(2), 459–468. https://doi.org/10.1016/j.ejor.2011.06.033.

Murota, K., & Shioura, A. (2001). Relationship of m-/l-convex functions with discrete convex functions by miller and favatitardella. *Discrete Applied Mathematics, 115*(1), 151–176. https://doi.org/10.1016/S0166-218X(01)00222-0. First Japanese-Hungarian Symposium for Discrete Mathematics and its Applications.

Pereira, M. V. F., & Pinto, L. M. V. G. (1991). Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming, 52*(1), 359–375. https://doi.org/10.1007/BF01582895.

Pitchbook (2017). Are meal-kit delivery companies a threat or an opportunity? https://pitchbook.com/news/articles/are-meal-kit-delivery-companies-a-threat-or-an-opportunity. Accessed 4 October 2019.

Powell, W. B. (2007). *Approximate dynamic programming: solving the curses of dimensionality (Wiley series in probability and statistics)*. New York, NY, USA: Wiley-Interscience.

Rockafellar, R. T., & Wets, R. J.-B. (1998). *Variational analysis*. Heidelberg, Berlin, New York: Springer Verlag.

Sauré, A., Patrick, J., Tyldesley, S., & Puterman, M. L. (2012). Dynamic multi-appointment patient scheduling for radiation therapy. *European Journal of Operational Research, 223*(2), 573–584. https://doi.org/10.1016/j.ejor.2012.06.046.

Shapiro, A. (2011). Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research, 209*(1), 63–72. https://doi.org/10.1016/j.ejor.2010.08.007.

Suh, M., & Aydin, G. (2011). Dynamic pricing of substitutable products with limited inventories under logit demand. *IIE Transactions, 43*, 323–331. https://doi.org/10.1080/0740817X.2010.521803.

Toth, P., & Vigo, D. (2014). *Vehicle Routing*. Philadelphia, PA: Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611973594.

Yang, X., & Strauss, A. K. (2017). An approximate dynamic programming approach to attended home delivery management. *European Journal of Operational Research, 263*(3), 935–945. https://doi.org/10.1016/j.ejor.2017.06.034.

Yang, X., Strauss, A. K., Currie, C. S. M., & Eglese, R. (2016). Choice-based demand management and vehicle routing in e-fulfillment. *Transportation Science, 50*(2), 473–488. https://doi.org/10.1287/trsc.2014.0549.

Zhang, D., & Adelman, D. (2009). An approximate dynamic programming approach to network revenue management with customer choice. *Transportation Science, 43*(3), 381–394. https://doi.org/10.1287/trsc.1090.0262.

Zhang, S., & Sun, X. A. (2019). Stochastic Dual Dynamic Programming for Multistage Stochastic Mixed-Integer Nonlinear Optimization. *Technical Report*. https://arxiv.org/pdf/1912.13278.pdf.

Zou, J., Ahmed, S., & Sun, X. A. (2019). Stochastic dual dynamic integer programming. *Mathematical Programming, 175*(1), 461–502. https://doi.org/10.1007/s10107-018-1249-5.