Incorporating Pricing Decisions into the Stochastic Dynamic Fleet Management Problem

Huseyin Topaloglu

School of Operations Research and Industrial Engineering,
Cornell University, Ithaca, NY 14853, USA
topaloglu@orie.cornell.edu

Warren B. Powell

Department of Operations Research and Financial Engineering,
Princeton University, Princeton, NJ 08544, USA
powell@princeton.edu

February 25, 2007

Abstract

This paper shows how to coordinate the pricing and fleet management decisions of a freight carrier. We consider a setting where the carrier announces its prices at the beginning of a certain time horizon and the load arrivals over this horizon depend on the announced prices. Assuming that the vehicle fleet is managed according to a particular class of fleet management models, we present a tractable method to obtain sample path-based directional derivatives of the objective function with respect to the prices. We use this information to search for a *good* set of prices. Numerical experiments show that our approach yields high-quality solutions.

An extensive literature has evolved around the problem of managing a fleet of vehicles to serve the loads occurring at different locations in a transportation network, but relatively little attention has been directed to the problem of determining what prices to charge. Due to the competitive nature of the freight transportation industry, lower prices can increase the number of loads over different traffic lanes (origin-destination pairs), but the *correct* prices have to consider not only the profit maximization problem over each traffic lane, but also the downstream effects that arise when the vehicles become empty and have to be repositioned to other locations.

In this paper, we address the question of how to coordinate the pricing and fleet management decisions of a freight carrier. Letting $\{1, \ldots, T\}$ be the set of time periods in the planning horizon, at the beginning of time period 1, the carrier decides what prices to charge over the next T time periods. These prices vary by traffic lanes, and may or may not vary by time. The objective is to find a set of prices that maximize the total expected profit over the planning horizon. We explicitly model the random load arrivals and the price-demand interactions by letting the number of loads over each traffic lane be a random variable whose distribution depends on the price. Assuming that the carrier makes its fleet management decisions by using the stochastic fleet management model previously developed by Godfrey & Powell (2002), we provide a tractable algorithm to obtain sample path-based directional derivatives of the objective function with respect to the prices. Starting from given prices, we use this information to search for better ones.

Fleet management models have their roots in some of the earliest applications of linear programming and min-cost network flow algorithms; see Dantzig & Fulkerson (1954), Ferguson & Dantzig (1955), White & Bomberault (1969) and White (1972). These early models essentially formulate the problem over a *state-time network*, where the nodes represent the supply of vehicles at different locations at different time periods, the arcs represent the vehicle movements and the load availabilities act as upper bounds on the arcs. They are referred to as deterministic models because they assume that the load arrivals over the entire planning horizon are known in advance and they incorporate the uncertain future load arrivals only through their expected values.

A second class of fleet management models try to treat the randomness in the load arrivals explicitly by decomposing the problem into time periods and assessing the impact of the current decisions on the future through value functions. Due to the large number of decision variables and possible load realizations, classical stochastic optimization techniques are usually not practical for computing the value functions. Therefore, most of the stochastic models require approximating the value functions in a tractable manner; see Jordan & Turnquist (1983), Frantzeskakis & Powell (1990), Crainic, Gendreau & Dejax (1993), Carvalho & Powell (2000), Godfrey & Powell (2002), Kleywegt, Nori & Savelsbergh (2002), Adelman (2004) and Topaloglu & Powell (2006). Of particular interest

to us is the model proposed by Godfrey & Powell (2002), which approximates the value functions by separable, piecewise-linear, concave functions. The primary appeal of this strategy is that it decomposes the fleet management problem into a sequence of min-cost network flow problems. Recently, Topaloglu & Powell (2004) exploit this min-cost network flow structure to propose a *sensitivity analysis* algorithm that computes the profit impact of an additional load introduced into the system. In this paper, we embed their algorithm in a pricing mechanism that considers the interactions between the prices and the load arrivals. We emphasize that the papers mentioned in this paragraph do not consider pricing issues. Instead, given fixed prices and a stochastic process that characterizes the load arrivals, they focus on either making the fleet management decisions or computing the profit impact of an additional load introduced into the system.

There has not been much work on pricing decisions in the fleet management context. Several papers address the problem of determining how much the total profit would change if an additional load is introduced into the system. If the underlying model utilizes a deterministic *state-time network* representation, then this problem can be solved by using the dual variables associated with the constraints that represent load availabilities; see Powell (1985), Powell, Sheffi, Nickerson, Butterbaugh & Atherton (1988) and Powell (1989). However, determining the profit impact of an additional load becomes difficult in the presence of uncertainty, primarily because the optimal fleet management policy is not known. Topaloglu & Powell (2004) show how to compute this profit impact when the fleet management decisions are made according to a (suboptimal) policy with a certain structure.

Gorman (2001) and King & Topaloglu (2006) are two papers we are aware of that explicitly model the interactions between the prices and the load arrivals. They consider the problem of what prices to charge assuming that the number of loads over a traffic lane is a deterministic function of the price. King & Topaloglu (2006) point out that the total profit, when viewed as a function of the prices, is not concave, and one must be content with locally optimal prices. Gorman (2001) reports significant reductions in the empty repositioning miles and attributes the reductions to the fact that the *correct* prices balance the traffic over the transportation network. An important point is that these two papers and our paper assume that the price-demand functions completely characterize the reactions of the shippers to the prices and ignore the competitor response. Gorman (2001) notes that one way to partially address this shortcoming is to use a kinked price-demand function, capturing the fact that the shippers react immensely to the price increases by reducing the volumes, but they do not react too much to the price decreases since the price decreases are likely to be matched by the competitors. In addition, these two papers and our paper do not consider possible shifts in the shipper demand patterns. For example, if the vehicle movements follow cycles and the vehicle supply at a particular location tends to be high on a particular day of the week, then the carrier may reduce

the price for the loads outbound from that location on that day of the week and the shippers may try to take advantage of the reduced price. Finally, Gorman (2001) and King & Topaloglu (2006) assume that the load arrivals are deterministic functions of the prices and do not address the difficulties that arise when the load arrivals are random. When the load arrivals are deterministic, the optimal fleet management policy can be found by solving a linear program, whereas it is practically impossible to find the optimal fleet management policy when the load arrivals are random.

Related models and solution methods also appear in the revenue management literature; see Talluri & van Ryzin (2004) for a comprehensive coverage. One possible approach for revenue management problems is to establish booking limits that capture what portion of the available product inventory should be reserved for different customer classes. Recently, Karaesmen & van Ryzin (2004), Bertsimas & de Boer (2005) and van Ryzin & Vulcano (2006) propose stochastic gradient algorithms that are somewhat similar to our approach for establishing the booking limits. Stochastic gradient algorithms have been successfully applied in other settings as well. For example, Fu (1994), Glasserman & Tayur (1995) and Mahajan & van Ryzin (2001) respectively present applications in finding attractive (s, S) policies in single-echelon inventory systems, computing base-stock levels in capacitated multi-echelon inventory systems and making stocking decisions for substitutable products.

In this paper, we make the following research contributions. 1) We present a model that coordinates the pricing and fleet management decisions, assuming that the fleet management decisions are made according to the class of policies proposed by Godfrey & Powell (2002). The interaction between the pricing and fleet management decisions is rarely investigated in the fleet management literature, and when it is, the usual assumption is that the load arrivals are deterministic functions of the prices. Our paper fills this gap. 2) We develop an efficient algorithm to obtain sample path-based directional derivatives of the profit function with respect to the prices and show how to use this information to search for a *good* set of prices. 3) Through extensive numerical experiments, we show that the proposed pricing approach yields high-quality prices and performs significantly better than deterministic methods that assume that the demand is a deterministic function of the price.

The organization of the paper is as follows. In Section 1, we describe our notation and the class of fleet management policies we consider, present a model of how the load arrivals react to changes in the prices and formulate the core pricing problem. In Section 2, we show how to obtain sample path-based directional derivatives of the objective function with respect to the prices and give a pricing algorithm. This section assumes that, for a given traffic lane, the carrier can announce different prices for different time periods. Section 3 modifies the results of Section 2 to address the situation where the carrier has to announce a single price for each traffic lane. In Section 4, we present numerical experiments that test the quality of the prices obtained through our pricing approach.

1 Problem Formulation

We have a homogeneous fleet of vehicles to serve the loads occurring at different locations over a finite planning horizon. At the beginning of the planning horizon, we set the prices that will be charged over each traffic lane and at each time period. At every time period, depending on the prices charged, a random number of loads enter the system, and we have to decide which loads to cover and to which locations we should reposition the empty vehicles. We are interested in maximizing the total expected profit. We assume that advance information about the future loads is not available and the loads that cannot be covered at a given time period are served by an emergency subcontractor. These enable us to assume that the uncovered loads immediately leave the system. For notational brevity, we assume that it takes one time period to move between any pair of locations, but our analysis can be repeated verbatim under multi-period travel times by using the modeling approach in Topaloglu & Powell (2006). We define the following.

$\mathcal{T}$ = Set of time periods in the planning horizon. We have $\mathcal{T} = \{1, \ldots, T\}$ for a finite T.

$\mathcal{I}$ = Set of locations in the transportation network.

$\mathcal{L}$ = Set of traffic lanes in the transportation network. Intuitively, a traffic lane represents an *arc*, whereas a location represents a *node* in the transportation network.

o_l, d_l = Origin and destination locations for traffic lane $l \in \mathcal{L}$.

x_{lt} = Decision variable representing the number of vehicles dispatched over traffic lane $l \in \mathcal{L}$ at time period $t \in \mathcal{T}$.

A traffic lane l represents a direct movement from location o_l to d_l, which can be through empty repositioning or serving a load. We assume that $\mathcal{L} = \mathcal{E} \bigcup \mathcal{D}$, where $\mathcal{E}$ and $\mathcal{D}$ are the sets of traffic lanes that respectively correspond to empty repositioning and loaded movements. In this case, the decision variable x_{lt} captures the empty repositioning movements from location o_l to d_l when $l \in \mathcal{E}$ and it captures the loaded movements from location o_l to d_l when $l \in \mathcal{D}$. We also define the following.

c_{lt} = Cost of empty repositioning a vehicle over traffic lane $l \in \mathcal{E}$ at time period $t \in \mathcal{T}$.

p_{lt} = Decision variable representing the price charged for serving one load over traffic lane $l \in \mathcal{D}$ at time period $t \in \mathcal{T}$.

$\rho_{lt}(p_{lt})$ = Profit from serving one load over traffic lane $l \in \mathcal{D}$ at time period $t \in \mathcal{T}$ given that the price charged is p_{lt}. We assume that $\rho_{lt}(\cdot)$ is strictly increasing and differentiable, and its derivative is finite everywhere.

$D_{lt}(p_{lt}) =$ Random variable representing the number of loads that need to be served over traffic

lane $l \in \mathcal{D}$ at time period $t \in \mathcal{T}$ given that the price charged is p_{lt}.

Since the costs of the empty repositioning movements are fixed, the decision variable p_{lt} is only defined for $l \in \mathcal{D}$, $t \in \mathcal{T}$. For notational uniformity, we assume that the decision variable p_{lt} is also defined for $l \in \mathcal{E}$, $t \in \mathcal{T}$, but we fix its value at $p_{lt} = -c_{lt}$ and follow the convention that $\rho_{lt}(p_{lt}) = p_{lt}$ for all $l \in \mathcal{E}$, $t \in \mathcal{T}$. Furthermore, since the empty repositioning movements are not bounded, we let $D_{lt}(p_{lt}) = \infty$ for all $l \in \mathcal{E}$, $t \in \mathcal{T}$. In this way, we do not have to make a distinction between the lanes in $\mathcal{E}$ and $\mathcal{D}$. For example, we can succinctly write the profit at time period t as $\sum_{l \in \mathcal{L}} \rho_{lt}(p_{lt}) \, x_{lt}$. Throughout the paper, we use $\tilde{D}_{lt}(p_{lt})$ to denote a particular realization of $D_{lt}(p_{lt})$. By suppressing some of the indices in the variables above, we denote a vector composed of the elements ranging over the suppressed indices. For example, we have $x_t = \{x_{lt} : l \in \mathcal{L}\}$, $p_t = \{p_{lt} : l \in \mathcal{L}\}$, $\rho_t(p_t) = \{\rho_{lt}(p_{lt}) : l \in \mathcal{L}\}$, $\tilde{D}_t(p_t) = \{\tilde{D}_{lt}(p_{lt}) : l \in \mathcal{L}\}$ and $\tilde{D}(p) = \{\tilde{D}_{lt}(p_{lt}) : l \in \mathcal{L}, \ t \in \mathcal{T}\}$.

In the remainder of this section, we review the fleet management model of Godfrey & Powell (2002). This allows us to define the cumulative profit function of a fleet management policy and to formulate the pricing problem that we want to solve. We conclude this section by describing how we model the load random variables $\{D_{lt}(p_{lt}) : p_{lt} \in \mathbb{R}, \ l \in \mathcal{L}, \ t \in \mathcal{T}\}$.

1.1 Fleet management policies

The fleet management policies that we consider are based on the model developed by Godfrey & Powell (2002). This model formulates the fleet management problem as a dynamic program and uses tractable approximations of the value functions.

To capture the state of the system, we define

$r_{it} =$ Number of vehicles that are available at location $i \in \mathcal{I}$ at time period $t \in \mathcal{T}$.

Clearly, the vector $r_t = \{r_{it} : i \in \mathcal{I}\}$ defines the state of the vehicles at time period t. Given this state vector and the realization of the loads at time period t, the set of feasible decision vectors and the set of state vectors generated by these decisions are defined by

$$\mathcal{Y}(r_t, \tilde{D}_t(p_t)) = \Big\{ (x_t, r_{t+1}) \in \mathbb{Z}_+^{|\mathcal{L}|+|\mathcal{I}|} : \sum_{l \in \mathcal{L}, o_l = i} x_{lt} = r_{it} \qquad \text{for all } i \in \mathcal{I} \qquad (1)$$

$$\sum_{l \in \mathcal{L}, d_l = i} x_{lt} - r_{i,t+1} = 0 \qquad \text{for all } i \in \mathcal{I} \qquad (2)$$

$$x_{lt} \leq \tilde{D}_{lt}(p_{lt}) \qquad \text{for all } l \in \mathcal{L} \Big\}. \qquad (3)$$

6

Thus, $(x_t, r_{t+1}) \in \mathcal{Y}(r_t, \tilde{D}_t(p_t))$ means that the decisions x_t are feasible when the state of the system is r_t and the realization of the loads is $\tilde{D}_t(p_t)$, and applying the decisions x_t generates the state vector r_{t+1} for the next time period. Constraints (3) represent the load availability constraints and imply that each vehicle can cover at most one load at a time. Constraints (1) and (2) respectively represent the vehicle availability and system dynamics constraints, and imply that the loads have to be carried all the way to their destination locations once they are picked up.

For fixed prices $\{p_t : t \in \mathcal{T}\}$, the optimal fleet management policy π^* can be found by computing the value functions $\{V_t(\cdot) : t \in \mathcal{T}\}$ through the backward recursion

$$V_t(r_t) = \mathbb{E}\left\{ \max_{(x_t, r_{t+1}) \in \mathcal{Y}(r_t, D_t(p_t))} \rho_t(p_t) \cdot x_t + V_{t+1}(r_{t+1}) \right\}; \tag{4}$$

see Puterman (1994). For a given state vector r_t and load realizations $\tilde{D}_t(p_t)$ at time period t, this optimal policy makes the decisions by solving the problem

$$\max_{(x_t, r_{t+1}) \in \mathcal{Y}(r_t, \tilde{D}_t(p_t))} \rho_t(p_t) \cdot x_t + V_{t+1}(r_{t+1}).$$

For almost all problem instances of practical significance, computing the value functions is intractable due to the well-known curse of dimensionality. Instead, we focus on a class of policies obtained by replacing the value functions $\{V_t(\cdot) : t \in \mathcal{T}\}$ with tractable approximations, say $\{\hat{V}_t^\pi(\cdot) : t \in \mathcal{T}\}$. Consequently, each set of value function approximations $\{\hat{V}_t^\pi(\cdot) : t \in \mathcal{T}\}$ characterize a policy π, which makes the decisions by solving the problem

$$\max_{(x_t, r_{t+1}) \in \mathcal{Y}(r_t, \tilde{D}_t(p_t))} \rho_t(p_t) \cdot x_t + \hat{V}_{t+1}^\pi(r_{t+1}). \tag{5}$$

For computational tractability, we use separable value function approximations of the form

$$\hat{V}_t^\pi(r_t) = \sum_{i \in \mathcal{I}} \hat{V}_{it}^\pi(r_{it}),$$

where each $\hat{V}_{it}^\pi(\cdot)$ is a one-dimensional, piecewise-linear, concave function with points of nondifferentiability being a subset of positive integers.

Godfrey & Powell (2002) show that problem (5) is a min-cost network flow problem under this value function approximation strategy and give an iterative algorithm to obtain value function approximations that characterize a *good* fleet management policy. The question of whether this class of policies yield high-quality solutions is outside the scope of this paper and we refer the reader to Godfrey & Powell (2002) where the experimental work indicates that this class of policies perform better than standard benchmarks. In this paper, we assume that we already have a policy characterized by separable, piecewise-linear, concave value function approximations. In Sections 2 and 3, we show how to find a *good* set of prices given that the fleet management decisions are made according to

7

this policy. In Section 4, we note that whether a particular fleet management policy is *good* actually depends on the prices, and use an iterative improvement heuristic to find a *good* fleet management policy and a *good* set of prices simultaneously.

1.2 Cumulative profit function

A Markovian deterministic fleet management policy π can be characterized by a sequence of decision functions $\{X_t^\pi(\cdot,\cdot,\cdot) : t \in \mathcal{T}\}$, such that $X_t^\pi(\cdot,\cdot,\cdot)$ maps the state vector r_t, prices p_t and load realizations $\tilde{D}_t(p_t)$ for time period t to a decision vector x_t. We can also define the state transition functions $\{R_{t+1}^\pi(\cdot,\cdot,\cdot) : t \in \mathcal{T}\}$ of policy π, such that $R_{t+1}^\pi(\cdot,\cdot,\cdot)$ maps the state vector, prices and load realizations for time period t to a state vector for the next time period. The decision and state transition functions for the class of policies that we consider are

$$\left(X_t^\pi(r_t, p_t, \tilde{D}_t(p_t)), R_{t+1}^\pi(r_t, p_t, \tilde{D}_t(p_t))\right) = \underset{(x_t, r_{t+1}) \in \mathcal{Y}(r_t, \tilde{D}_t(p_t))}{\operatorname{argmax}} \rho_t(p_t) \cdot x_t + \hat{V}_{t+1}^\pi(r_{t+1}). \tag{6}$$

Throughout the paper, we restrict our attention to the prices p_t that make the optimal solution to problem (6) unique for all state vectors r_t and all load realizations $\tilde{D}_t(p_t)$. In this case, the decision and state transition functions are well-defined. The next lemma and the discussion that follows show that such an assumption is satisfied almost everywhere. The proof is deferred to the appendix.

Lemma 1 *There exist a finite number of at most $(|\mathcal{L}| - 1)$-dimensional subspaces of $\mathbb{R}^{|\mathcal{L}|}$, say $\mathcal{Q}_t^1, \ldots, \mathcal{Q}_t^N$, such that if we have $\rho_t(p_t) \in \mathbb{R}^{|\mathcal{L}|} \setminus \bigcup_{n=1}^N \mathcal{Q}_t^n$, then the optimal solution to problem (6) is unique for all state vectors r_t and all load realizations $\tilde{D}_t(p_t)$.*

Therefore, if we ignore the cases where the objective function coefficients are in the vacuous set $\bigcup_{n=1}^N \mathcal{Q}_t^n$, then problem (6) has a unique optimal solution for all state vectors r_t and all load realizations $\tilde{D}_t(p_t)$. Since the functions $\{\rho_{lt}(\cdot) : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ are strictly increasing, there is a one-to-one correspondence between the objective function coefficients and the prices. In this case, the prices that we need to ignore also lie in a vacuous set. In particular, letting

$$\mathcal{P}_t = \left\{ p_t : \rho_t(p_t) \in \mathbb{R}^{|\mathcal{L}|} \setminus \bigcup_{n=1}^N \mathcal{Q}_t^n \right\},$$

$\mathcal{P}_t$ differs from $\mathbb{R}^{|\mathcal{L}|}$ by a set of measure zero and if we have $p_t \in \mathcal{P}_t$, then the optimal solution to problem (6) is unique for all state vectors r_t and all load realizations $\tilde{D}_t(p_t)$.

For a given state vector r_t at time period t, and prices $p = \{p_{lt} : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ and load realizations $\tilde{D}(p) = \{\tilde{D}_{lt}(p_{lt}) : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ over the whole planning horizon, we write the cumulative profit function for policy π recursively as

$$F_t^\pi\left(r_t, p, \tilde{D}(p)\right) = \rho_t(p_t) \cdot X_t^\pi(r_t, p_t, \tilde{D}_t(p_t)) + F_{t+1}^\pi\left(R_{t+1}^\pi(r_t, p_t, \tilde{D}_t(p_t)), p, \tilde{D}(p)\right), \tag{7}$$

with $F^{\pi}_{T+1}(\cdot, \cdot, \cdot) = 0$. By repeated application of (7), it is easy to see that

$$F^{\pi}_1(r_1, p, \tilde{D}(p)) = \rho_1(p_1) \cdot X^{\pi}_1(r_1, p_1, \tilde{D}_1(p_1)) + \rho_2(p_2) \cdot X^{\pi}_2(R^{\pi}_2(r_1, p_1, \tilde{D}_1(p_1)), p_2, \tilde{D}_2(p_2))$$
$$+ \ldots + \rho_T(p_T) \cdot X^{\pi}_T(R^{\pi}_T(\ldots, p_{T-1}, \tilde{D}_{T-1}(p_{T-1})), p_T, \tilde{D}_T(p_T)).$$

Thus, $F^{\pi}_1(r_1, p, \tilde{D}(p))$ is the profit obtained over the whole planning horizon, if we use policy π, the initial state vector is r_1, the prices are p and the load realizations are $\tilde{D}(p)$. Letting $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2 \times \ldots \times \mathcal{P}_T$, the quantity $F^{\pi}_1(r_1, p, \tilde{D}(p))$ is well-defined whenever we have $p \in \mathcal{P}$.

1.3 Pricing problem

Letting π be a fixed fleet management policy belonging to the class of policies we consider, the pricing problem we want to solve is

$$\sup_p \; \mathbb{E}\{F^{\pi}_1(r_1, p, D(p))\}. \tag{8}$$

Noting the discussion following Lemma 1, the objective function of problem (8) is well-defined almost everywhere. One way to approach this problem is to use a method that resembles gradient search. Letting e_{lt} be the $|\mathcal{L}||\mathcal{T}|$-dimensional unit vector with a 1 in the element corresponding to $l \in \mathcal{L}$, $t \in \mathcal{T}$ and α be a scalar, one can start with the prices $\hat{p} \in \mathcal{P}$ and compute the directional derivatives

$$\lim_{\alpha \uparrow \downarrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{F^{\pi}_1(r_1, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})) - F^{\pi}_1(r_1, \hat{p}, D(\hat{p}))\Big\}, \tag{9}$$

where $\lim_{\alpha \uparrow \downarrow 0}$ denotes the left and right limits as α approaches zero. The expression above gives the change in the total expected profit of policy π with respect to a change in the price charged over lane l at time period t. This can be used as a step direction in a quasi-gradient search method.

However, computing the expression in (9) can be prohibitive due to the expectation operator. Instead, we opt for using sample path-based directional derivatives. In particular, letting $\tilde{D}(\hat{p})$ be a fixed realization of $D(\hat{p})$, we propose using the sample path-based directional derivatives

$$\lim_{\alpha \uparrow \downarrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{F^{\pi}_1(r_1, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})) - F^{\pi}_1(r_1, \hat{p}, D(\hat{p})) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p})\Big\}. \tag{10}$$

For all $\hat{p} \in \mathcal{P}$, we have

$$\mathbb{E}\bigg\{ \lim_{\alpha \uparrow \downarrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{F^{\pi}_1(r_1, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})) - F^{\pi}_1(r_1, \hat{p}, D(\hat{p})) \,\big|\, D(\hat{p})\Big\}\bigg\}$$
$$= \lim_{\alpha \uparrow \downarrow 0} \frac{1}{\alpha} \mathbb{E}\bigg\{\mathbb{E}\Big\{F^{\pi}_1(r_1, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})) - F^{\pi}_1(r_1, \hat{p}, D(\hat{p})) \,\big|\, D(\hat{p})\Big\}\bigg\}$$
$$= \lim_{\alpha \uparrow \downarrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{F^{\pi}_1(r_1, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})) - F^{\pi}_1(r_1, \hat{p}, D(\hat{p}))\Big\}, \tag{11}$$

which implies that (10) provides a valid sample path-based directional derivative at $\hat{p}$. The next lemma shows that the expression inside the limits in (10) is uniformly bounded. Therefore, exchanging the order of the expectation and the limits in (11) is justified by appealing to the bounded convergence theorem; see Royden (1988). The proof of Lemma 2 uses results that we derive in the proof of Proposition 5 and is deferred to the appendix.

Lemma 2 *Assume that $\hat{p} \in \mathcal{P}$. In this case, there exist $\epsilon > 0$ and $M < \infty$ such that if we have $\alpha \in [-\epsilon, \epsilon] \setminus \{0\}$, then*

$$\left| \frac{1}{\alpha} \mathbb{E} \left\{ F_1^\pi \left(r_1, \hat{p} + \alpha \, e_{lt}, D(\hat{p} + \alpha \, e_{lt}) \right) - F_1^\pi \left(r_1, \hat{p}, D(\hat{p}) \right) \mid D(\hat{p}) \right\} \right| \le M \tag{12}$$

holds for all realizations of $D(\hat{p})$.

Our main contribution is to show that the limits in (10) exist whenever we have $\hat{p} \in \mathcal{P}$ and to provide a tractable algorithm to compute them. This is the subject of Sections 2 and 3. We conclude this section by describing how we model the load random variables $\{D_{lt}(p_{lt}) : p_{lt} \in \mathbb{R}, \ l \in \mathcal{L}, \ t \in \mathcal{T}\}$.

1.4 Load random variables

We assume that $D_{lt}(p_{lt})$ has a Poisson distribution. Furthermore, we want the family of random variables $\{D_{lt}(p_{lt}) : p_{lt} \in \mathbb{R}\}$ to be *decreasing* in p_{lt}. In particular, we want them to satisfy

$$D_{lt}(p_{lt} + \alpha) \le D_{lt}(p_{lt}) \le D_{lt}(p_{lt} - \alpha) \tag{13}$$

with probability 1 for all $p_{lt} \in \mathbb{R}$ and $\alpha \ge 0$. To model this kind of a relationship, we let $\{N_{lt}(s) : s \in \mathbb{R}_+\}$ be a Poisson process with unit arrival rate and $\lambda_{lt}(\cdot)$ be a decreasing, differentiable, positive-valued function with a finite derivative everywhere. Then, we let $D_{lt}(p_{lt})$ be the number of arrivals of this Poisson process over the time interval $(0, \lambda_{lt}(p_{lt})]$; see Figure 1. Due to this construction, $D_{lt}(p_{lt})$ has a Poisson distribution with mean $\lambda_{lt}(p_{lt})$ and (13) is satisfied. Furthermore, for $\alpha \ge 0$, the change in the load arrivals in response to a decrease in the price is described by

$$\mathbb{P}\left\{ D_{lt}(p_{lt} - \alpha) = \tilde{D}_{lt}(p_{lt}) \mid D_{lt}(p_{lt}) = \tilde{D}_{lt}(p_{lt}) \right\} = 1 - \lambda_{lt}(p_{lt} - \alpha) + \lambda_{lt}(p_{lt}) + o(\alpha) \tag{14}$$

$$\mathbb{P}\left\{ D_{lt}(p_{lt} - \alpha) = \tilde{D}_{lt}(p_{lt}) + 1 \mid D_{lt}(p_{lt}) = \tilde{D}_{lt}(p_{lt}) \right\} = \lambda_{lt}(p_{lt} - \alpha) - \lambda_{lt}(p_{lt}) + o(\alpha) \tag{15}$$

$$\mathbb{P}\left\{ D_{lt}(p_{lt} - \alpha) > \tilde{D}_{lt}(p_{lt}) + 1 \mid D_{lt}(p_{lt}) = \tilde{D}_{lt}(p_{lt}) \right\} = o(\alpha), \tag{16}$$

where $o(\cdot)$ stands for a function $g(\cdot)$ satisfying $\lim_{\alpha \to 0} g(\alpha)/\alpha = 0$. Exact expressions for the probabilities above can be given by using the probability mass function of the Poisson distribution, but the way they are expressed will be more convenient for proving our results.

Similarly, the change in the load arrivals in response to an increase in the price can be described by using the fact that given times $0 < u \leq v$ and that $N_{lt}(v) = k$, $N_{lt}(u)$ has a binomial distribution with parameters k and u/v. For example, we have

$$\mathbb{P}\{D_{lt}(p_{lt} + \alpha) = \tilde{D}_{lt}(p_{lt}) - 1 \mid D_{lt}(p_{lt}) = \tilde{D}_{lt}(p_{lt})\}$$

$$= \tilde{D}_{lt}(p_{lt})\left[1 - \frac{\lambda_{lt}(p_{lt} + \alpha)}{\lambda_{lt}(p_{lt})}\right]\left[\frac{\lambda_{lt}(p_{lt} + \alpha)}{\lambda_{lt}(p_{lt})}\right]^{\tilde{D}_{lt}(p_{lt}) - 1}$$

for $\alpha \geq 0$. Throughout the paper, we assume that the processes $\{N_{lt}(s) : s \in \mathbb{R}_+\}$ and $\{N_{l't'}(s) : s \in \mathbb{R}_+\}$ are independent when $l \neq l'$ or $t \neq t'$ so that the load arrivals over different lanes or at different time periods are independent.

2   Characterizing the Directional Derivatives

In this section, we provide a tractable method to compute the limits in (10) and show how the information provided by them can be used to make pricing decisions.

We begin with the next lemma, which follows from the fact that problem (6) has a unique optimal solution whenever we have $p_t \in \mathcal{P}_t$. Its proof is in the appendix. In this lemma and throughout the paper, we let e_l be the $|\mathcal{L}|$-dimensional unit vector with a 1 in the element corresponding to $l \in \mathcal{L}$.

Lemma 3 *Assume that $p_t \in \mathcal{P}_t$. In this case, there exists $\epsilon > 0$ such that, for all state vectors r_t and all load realizations $\tilde{D}_t(p_t)$, if we have $\alpha \in [-\epsilon, \epsilon]$, then the optimal solution to the problem*

$$\max_{(x_t, r_{t+1}) \in \mathcal{Y}(r_t, \tilde{D}_t(p_t))} \rho_t(p_t + \alpha\, e_l) \cdot x_t + \hat{V}_{t+1}^{\pi}(r_{t+1})$$

does not depend on α.

To compute the limits in (10), we fix prices $\hat{p} \in \mathcal{P}$ and let $\tilde{D}(\hat{p})$ be a fixed realization of $D(\hat{p})$. We let $\{\hat{x}_t : t \in \mathcal{T}\}$ and $\{\hat{r}_t : t \in \mathcal{T}\}$ be the sequences of decision and state vectors visited by policy π under prices $\hat{p}$ and load realizations $\tilde{D}(\hat{p})$. Specifically, $\{\hat{x}_t : t \in \mathcal{T}\}$ and $\{\hat{r}_t : t \in \mathcal{T}\}$ are recursively computed by starting with $\hat{r}_1 = r_1$, and letting $\hat{x}_t = X_t^{\pi}(\hat{r}_t, \hat{p}, \tilde{D}_t(\hat{p}_t))$ and $\hat{r}_{t+1} = R_{t+1}^{\pi}(\hat{r}_t, \hat{p}, \tilde{D}_t(\hat{p}_t))$ for all $t \in \mathcal{T}$. The next lemma shows that, for the purpose of computing (10), we can concentrate on the system starting from time period t with the state vector $\hat{r}_t$.

Lemma 4 *For the fixed prices $\hat{p}$ and load realizations $\tilde{D}(\hat{p})$, the limits in (10) can be written as*

$$\lim_{\alpha \uparrow \downarrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{F_1^{\pi}\big(r_1, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})\big) - F_1^{\pi}\big(r_1, \hat{p}, D(\hat{p})\big) \mid D(\hat{p}) = \tilde{D}(\hat{p})\Big\}$$

$$= \lim_{\alpha \uparrow \downarrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{F_t^{\pi}\big(\hat{r}_t, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})\big) - F_t^{\pi}\big(\hat{r}_t, \hat{p}, D(\hat{p})\big) \mid D(\hat{p}) = \tilde{D}(\hat{p})\Big\}. \tag{17}$$

11

Proof The result is a consequence of the fact that the load realizations and decisions at time periods $\{1, \ldots, t-1\}$ do not depend on the prices at time period t. To formalize, we fix α and let $\tilde{D}(\hat{p} + \alpha\,e_{lt})$ be a fixed realization of $D(\hat{p} + \alpha\,e_{lt})$. We let $\{\bar{x}_t : t \in \mathcal{T}\}$ and $\{\bar{r}_t : t \in \mathcal{T}\}$ be the sequences of decision and state vectors visited by policy π under prices $\hat{p} + \alpha\,e_{lt}$ and load realizations $\tilde{D}(\hat{p} + \alpha\,e_{lt})$. Since we have $\hat{p} \in \mathcal{P}$, if α is small enough, then $\{\bar{x}_t : t \in \mathcal{T}\}$ and $\{\bar{r}_t : t \in \mathcal{T}\}$ are well-defined by Lemma 3. The prices $\hat{p}$ and $\hat{p} + \alpha\,e_{lt}$ coincide at time periods $\{1, \ldots, t-1\}$. Noting our construction in Section 1.4, this implies that the load random variables $D(\hat{p})$ and $D(\hat{p} + \alpha\,e_{lt})$ coincide at time periods $\{1, \ldots, t-1\}$ with probability 1. In this case, given that $D(\hat{p}) = \tilde{D}(\hat{p})$, we have $\bar{x}_s = \hat{x}_s$ and $\bar{r}_s = \hat{r}_s$ for all $s \in \{1, \ldots, t-1\}$. Furthermore, given that $D(\hat{p}) = \tilde{D}(\hat{p})$, we have $\bar{r}_t = R_t^\pi(\bar{r}_{t-1}, \hat{p}_{t-1}, \tilde{D}_{t-1}(\hat{p}_{t-1})) = R_t^\pi(\hat{r}_{t-1}, \hat{p}_{t-1}, \tilde{D}_{t-1}(\hat{p}_{t-1})) = \hat{r}_t$. Therefore, the result follows by noting that the conditional expectation on the right side of (17) is

$$\mathbb{E}\Big\{ \textstyle\sum_{s=1}^{t-1} \rho_s(\hat{p}_s) \cdot \bar{x}_s + F_t^\pi\big(\bar{r}_t, \hat{p} + \alpha\,e_{lt}, D(\hat{p} + \alpha\,e_{lt})\big)$$
$$- \textstyle\sum_{s=1}^{t-1} \rho_s(\hat{p}_s) \cdot \hat{x}_s - F_t^\pi\big(\hat{r}_t, \hat{p}, D(\hat{p})\big) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p}) \Big\}. \qquad \square$$

The next proposition gives our main result.

Proposition 5 *For the fixed prices $\hat{p}$ and load realizations $\tilde{D}(\hat{p})$, we have*

$$\lim_{\alpha\uparrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{ F_1^\pi\big(r_1, \hat{p} + \alpha\,e_{lt}, D(\hat{p} + \alpha\,e_{lt})\big) - F_1^\pi\big(r_1, \hat{p}, D(\hat{p})\big) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p}) \Big\}$$
$$= \dot{\rho}_{lt}(\hat{p}_{lt})\,\hat{x}_{lt} + \dot{\lambda}_{lt}(\hat{p}_{lt})\Big\{ F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) + e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big) \Big\} \qquad (18)$$

$$\lim_{\alpha\downarrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{ F_1^\pi\big(r_1, \hat{p} + \alpha\,e_{lt}, D(\hat{p} + \alpha\,e_{lt})\big) - F_1^\pi\big(r_1, \hat{p}, D(\hat{p})\big) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p}) \Big\}$$
$$= \dot{\rho}_{lt}(\hat{p}_{lt})\,\hat{x}_{lt} + \dot{\lambda}_{lt}(\hat{p}_{lt}) \frac{\tilde{D}_{lt}(\hat{p}_{lt})}{\lambda_{lt}(\hat{p}_{lt})} \Big\{ F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) - e_{lt}\big) \Big\}, \quad (19)$$

where $\dot{\rho}_{lt}(\hat{p}_{lt})$ and $\dot{\lambda}_{lt}(\hat{p}_{lt})$ are respectively the derivatives of $\rho_{lt}(\cdot)$ and $\lambda_{lt}(\cdot)$ evaluated at $\hat{p}_{lt}$. If we have $\tilde{D}_{lt}(\hat{p}_{lt}) = 0$, then $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) - e_{lt}\big)$ is not well-defined, and we let $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) - e_{lt}\big) = -\infty$ and use the convention that $0\infty = 0/0 = 0$.

Proof We only show that (18) holds. Our proof is in two parts, each of which shows that

$$\lim_{\alpha\uparrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{ F_t^\pi\big(\hat{r}_t, \hat{p} + \alpha\,e_{lt}, D(\hat{p} + \alpha\,e_{lt})\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, D(\hat{p} + \alpha\,e_{lt})\big) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p}) \Big\} = \dot{\rho}_{lt}(p_{lt})\,\hat{x}_{lt} \quad (20)$$

$$\lim_{\alpha\uparrow 0} \frac{1}{\alpha} \mathbb{E}\Big\{ F_t^\pi\big(\hat{r}_t, \hat{p}, D(\hat{p} + \alpha\,e_{lt})\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, D(\hat{p})\big) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p}) \Big\}$$
$$= \dot{\lambda}_{lt}(\hat{p}_{lt})\Big\{ F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) + e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big) \Big\}. \quad (21)$$

In this case, the results follows by Lemma 4.

Part 1. Lemma 3 implies that there exists $\epsilon > 0$ such that

$$X_t^\pi(\hat{r}_t, \hat{p}_t + \alpha\, e_l, D_t(\hat{p}_t + \alpha\, e_l)) = X_t^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\, e_l)) \tag{22}$$

$$R_{t+1}^\pi(\hat{r}_t, \hat{p}_t + \alpha\, e_l, D_t(\hat{p}_t + \alpha\, e_l)) = R_{t+1}^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\, e_l)) \tag{23}$$

for all $\alpha \in [-\epsilon, \epsilon]$ and all realizations of $D_t(\hat{p}_t + \alpha\, e_l)$. Throughout the proof, we assume that $\alpha \in [-\epsilon, \epsilon]$. Once we know the state vector at time period $t + 1$, the cumulative profit starting from this time period depends on the prices only at time periods $\{t + 1, \ldots, T\}$. Therefore, we have

$$F_{t+1}^\pi\big(R_{t+1}^\pi(\hat{r}_t, \hat{p}_t + \alpha\, e_l, D_t(\hat{p}_t + \alpha\, e_l)), \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})\big)$$

$$= F_{t+1}^\pi\big(R_{t+1}^\pi(\hat{r}_t, \hat{p}_t + \alpha\, e_l, D_t(\hat{p}_t + \alpha\, e_l)), \hat{p}, D(\hat{p} + \alpha\, e_{lt})\big)$$

$$= F_{t+1}^\pi\big(R_{t+1}^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\, e_l)), \hat{p}, D(\hat{p} + \alpha\, e_{lt})\big), \tag{24}$$

where the second equality follows from (23). Finally, we can use (7) to write

$$F_t^\pi\big(\hat{r}_t, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})\big) = \rho_t(\hat{p}_t + \alpha\, e_l) \cdot X_t^\pi(\hat{r}_t, \hat{p}_t + \alpha\, e_l, D_t(\hat{p}_t + \alpha\, e_l))$$

$$+ F_{t+1}^\pi\big(R_{t+1}^\pi(\hat{r}_t, \hat{p}_t + \alpha\, e_l, D_t(\hat{p}_t + \alpha\, e_l)), \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})\big)$$

$$F_t^\pi\big(\hat{r}_t, \hat{p}, D(\hat{p} + \alpha\, e_{lt})\big) = \rho_t(\hat{p}_t) \cdot X_t^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\, e_l))$$

$$+ F_{t+1}^\pi\big(R_{t+1}^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\, e_l)), \hat{p}, D(\hat{p} + \alpha\, e_{lt})\big).$$

Subtracting the two equalities above, (22) and (24) imply that

$$F_t^\pi\big(\hat{r}_t, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, D(\hat{p} + \alpha\, e_{lt})\big)$$

$$= \big[\rho_t(\hat{p}_t + \alpha\, e_l) - \rho_t(\hat{p}_t)\big] \cdot X_t^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\, e_l)). \tag{25}$$

Consequently, the limit in (20) is equal to

$$\lim_{\alpha\uparrow 0} \frac{1}{\alpha}\big[\rho_t(\hat{p}_t + \alpha\, e_l) - \rho_t(\hat{p}_t)\big] \cdot \lim_{\alpha\uparrow 0} \mathbb{E}\Big\{X_t^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\, e_l)) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p})\Big\}.$$

We conclude this part by showing that the second limit above is equal to $\hat{x}_t$. Since the load random variables over different lanes or at different time periods are independent, (14)-(16) imply that

$$\mathbb{P}\big\{D_t(\hat{p}_t + \alpha\, e_l) = \tilde{D}_t(\hat{p}_t) \,|\, D(\hat{p}) = \tilde{D}(\hat{p})\big\} = 1 - \lambda_{lt}(\hat{p}_{lt} + \alpha) + \lambda_{lt}(\hat{p}_{lt}) + o(\alpha) \tag{26}$$

$$\mathbb{P}\big\{D_t(\hat{p}_t + \alpha\, e_l) = \tilde{D}_t(\hat{p}_t) + e_l \,|\, D(\hat{p}) = \tilde{D}(\hat{p})\big\} = \lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt}) + o(\alpha) \tag{27}$$

for $\alpha \leq 0$. The total probability of the remaining events is $o(\alpha)$. Therefore, we have

$$\mathbb{E}\Big\{X_t^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\, e_l)) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p})\Big\} =$$

$$\big[1 - \lambda_{lt}(\hat{p}_{lt} + \alpha) + \lambda_{lt}(\hat{p}_{lt})\big] X_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t))$$

$$+ \big[\lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt})\big] X_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t) + e_l) + o(\alpha).$$

Taking the limits of both sides as $\alpha \uparrow 0$ and noting that $X_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t)) = \hat{x}_t$ establishes Part 1.

Part 2. The load random variables $D(\hat{p})$ and $D(\hat{p} + \alpha\, e_{lt})$ coincide at time periods $\{1, \ldots, t-1, t+1, \ldots, T\}$ with probability 1. Therefore, (26) and (27) imply that

$$\mathbb{P}\{D(\hat{p} + \alpha\, e_{lt}) = \tilde{D}(\hat{p}) \mid D(\hat{p}) = \tilde{D}(\hat{p})\} = 1 - \lambda_{lt}(\hat{p}_{lt} + \alpha) + \lambda_{lt}(\hat{p}_{lt}) + o(\alpha)$$

$$\mathbb{P}\{D(\hat{p} + \alpha\, e_{lt}) = \tilde{D}(\hat{p}) + e_{lt} \mid D(\hat{p}) = \tilde{D}(\hat{p})\} = \lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt}) + o(\alpha)$$

for $\alpha \leq 0$. In this case, the conditional expectation in (21) is

$$\big[1 - \lambda_{lt}(\hat{p}_{lt} + \alpha) + \lambda_{lt}(\hat{p}_{lt})\big]\Big\{F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)\Big\}$$

$$+ \big[\lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt})\big]\Big\{F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) + e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)\Big\} + o(\alpha). \quad (28)$$

The first term above is zero and Part 2 follows by dividing the expression above by α and taking the limit as $\alpha \uparrow 0$. $\qquad\square$

Since the functions $\rho_{lt}(\cdot)$ and $\lambda_{lt}(\cdot)$ are assumed to be known, $\dot{\rho}_{lt}(\hat{p}_{lt})$ and $\dot{\lambda}_{lt}(\hat{p}_{lt})$ in (18) and (19) are readily available. The values of the decision variables $\{\hat{x}_{lt} : l \in \mathcal{L}, t \in \mathcal{T}\}$ are obtained from a single simulation of policy π under prices $\hat{p}$ and load realizations $\tilde{D}(\hat{p})$. Therefore, the only computationally-problematic terms in (18) and (19) are $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)$. We propose three methods to compute them, each with a different level of tractability and accuracy.

Numerical difference-based method. We simulate policy π two more times, once under prices $\hat{p}$ and load realizations $\tilde{D}(\hat{p}) + e_{lt}$, and once under prices $\hat{p}$ and load realizations $\tilde{D}(\hat{p}) - e_{lt}$. This method is clearly exact, but computing $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)$ for all $l \in \mathcal{L}, t \in \mathcal{T}$ requires $2|\mathcal{L}||\mathcal{T}|$ additional simulations.

Value function approximation-based method. This method rests on a heuristic argument, but it is the most efficient of the three. The difference $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)$ characterizes how much the profit obtained by policy π under prices $\hat{p}$ and load realizations $\tilde{D}(\hat{p})$ changes in response to a load introduced over or removed from lane l at time period t. If $\hat{x}_{lt} < \tilde{D}_{lt}(\hat{p}_{lt})$ holds, then we have $X_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t)) = \hat{x}_{lt} = X_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t) \mp e_{lt})$, which implies that $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big) = 0$. Therefore, we assume that $\hat{x}_{lt} = \tilde{D}_{lt}(\hat{p}_{lt})$ for the rest of the discussion.

Roughly speaking, $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)$ is the profit impact of a load introduced over or removed from lane l at time period t. Letting

$$\widetilde{V}_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t)) = \max_{(x_t, r_{t+1}) \in \mathcal{Y}(\hat{r}_t, \tilde{D}_t(\hat{p}_t))} \rho_t(\hat{p}_t) \cdot x_t + \hat{V}_{t+1}^\pi(r_{t+1}) \quad (29)$$

and using e_i' to denote the $|\mathcal{I}|$-dimensional unit vector with a 1 in the element corresponding to $i \in \mathcal{I}$, the idea behind the value function approximation-based method is to use $\widetilde{V}_t^\pi(\hat{r}_t \mp e_i', \hat{p}_t, \tilde{D}_t(\hat{p}_t)) - $

14

$\widetilde{V}_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t))$ to approximate the profit impact of a vehicle introduced at or removed from location i at time period t.

Assume that we introduce an additional load over lane l at time period t. If the additional load is not covered, then its profit impact is zero. However, if the additional load is covered, then it occupies a vehicle at location o_l and makes the vehicle move to location d_l. Due to our single-period travel time assumption, this movement is completed at time period $t+1$. Consequently, we use

$$
\max\Big\{ \widetilde{V}_t^\pi(\hat{r}_t - e'_{o_l}, \hat{p}_t, \tilde{D}_t(\hat{p}_t)) - \widetilde{V}_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t)) + \rho_{lt}(\hat{p}_{lt})
$$
$$
+ \widetilde{V}_{t+1}^\pi(\hat{r}_{t+1} + e'_{d_l}, \hat{p}_{t+1}, \tilde{D}_{t+1}(\hat{p}_{t+1})) - \widetilde{V}_{t+1}^\pi(\hat{r}_{t+1}, \hat{p}_{t+1}, \tilde{D}_{t+1}(\hat{p}_{t+1})), \, 0 \Big\}
$$

as an approximation to $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) + e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)$, where the two terms in the max operator above respectively correspond to the cases where the additional load is and is not covered.

Assume that we remove a load from lane l at time period t. Due to the assumption that $\hat{x}_{lt} = \tilde{D}_{lt}(\hat{p}_{lt})$, we stop covering the removed load and this makes a vehicle available at location o_l at time period t. Consequently, we use

$$
\widetilde{V}_t^\pi(\hat{r}_t + e'_{o_l}, \hat{p}_t, \tilde{D}_t(\hat{p}_t)) - \widetilde{V}_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t)) - \rho_{lt}(\hat{p}_{lt})
$$
$$
+ \widetilde{V}_{t+1}^\pi(\hat{r}_{t+1} - e'_{d_l}, \hat{p}_{t+1}, \tilde{D}_{t+1}(\hat{p}_{t+1})) - \widetilde{V}_{t+1}^\pi(\hat{r}_{t+1}, \hat{p}_{t+1}, \tilde{D}_{t+1}(\hat{p}_{t+1}))
$$

as an approximation to $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) - e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)$. We note the similarity between the expressions above and the computation of the dual variable associated with an upper bound constraint in a min-cost network flow problem, which can be stated as $-\pi_i + c_{ij} + \pi_j$, where π_i is the dual variable associated with the flow balance constraint of node i and c_{ij} is the cost of arc (i,j).

Since problem (29) is a min-cost network flow problem, we can compute $\widetilde{V}_t^\pi(\hat{r}_t \mp e'_i, \hat{p}_t, \tilde{D}_t(\hat{p}_t)) - \widetilde{V}_t^\pi(\hat{r}_t, \hat{p}_t, \tilde{D}_t(\hat{p}_t))$ for all $i \in \mathcal{I}$ through two shortest path tree computations; see Powell (1989). Therefore, this method requires $2|\mathcal{T}|$ shortest path tree computations. Although it lacks theoretical foundation, Section 4 shows that this approach performs well in practice.

Perturbation analysis-based method. For the class of policies we consider, Topaloglu & Powell (2004) develop an algorithm that approximates $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)$. Their approach is based on tracing the impact of an additional load on the future trajectory of policy π and it uses ideas similar to those in infinitesimal perturbation analysis; see Glasserman (1991). Their algorithm is exact when $F_{t+1}^\pi(\cdot, \hat{p}, \tilde{D}(\hat{p}))$ is a separable function and yields accurate approximations even when the separability assumption does not hold. It computes $F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}\big) - F_t^\pi\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)$ for all $l \in \mathcal{L}$, $t \in \mathcal{T}$ through $2|\mathcal{I}||\mathcal{T}|$ shortest path tree computations. We refer the reader to Topaloglu & Powell (2004) for the details.

Figure 2 describes Algorithm 1, which uses (18) and (19) to find a *good* set of prices for a fixed fleet management policy π. The objective function of problem (8) is not necessarily concave. Since the left and right limits in (18) and (19) are not necessarily equal to each other, this objective function is not continuously differentiable either. Consequently, the standard results in Ermoliev (1988) and Bertsekas & Tsitsiklis (1996) that exploit the concavity or smoothness of the objective function are not applicable to show the convergence of Algorithm 1 and we do not have a convergence result for Algorithm 1.

Standard convergence results require the step size parameters in Step 6 to be positive, and to satisfy $\sum_{n=1}^{\infty} \gamma_{lt}^n = \infty$ and $\sum_{n=1}^{\infty} \left[\gamma_{lt}^n\right]^2 < \infty$ for all $l \in \mathcal{L}$, $t \in \mathcal{T}$. We experiment with numerous step size parameters, some of which do not satisfy these requirements. We describe our final choice in Section 4.3. One may consider perturbing $g_{lt}^{\pi\mp}(r_t^n, p^n, \tilde{D}^n)$ in Step 5 by a random variable that is uniformly distributed over the small interval $[-\upsilon, \upsilon]$. Noting the discussion that follows Lemma 1, since $\mathcal{P}$ differs from $\mathbb{R}^{|\mathcal{L}||\mathcal{T}|}$ by a set of measure zero, this ensures that the prices p^{n+1} lie in $\mathcal{P}$ with probability 1. Step 6 assumes that the prices are unconstrained, but lower and upper bound constraints of the form $L_{lt} \leq p_{lt} \leq U_{lt}$ can be imposed by modifying this step as

$$
p_{lt}^{n+1} = \begin{cases} \max\{L_{lt}, \, p_{lt}^n + \gamma_{lt}^n \, g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n)\} & \text{if } g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n) < 0 \text{ and} \\ & \qquad\qquad - g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n) > g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n) \\ \min\{U_{lt}, \, p_{lt}^n + \gamma_{lt}^n \, g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n)\} & \text{if } g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n) > 0 \text{ and} \\ & \qquad\qquad g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n) \geq -g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n) \\ p_{lt}^n & \text{otherwise.} \end{cases}
$$

To ensure that the prices p^{n+1} lie in $\mathcal{P}$ with probability 1, one may also consider perturbing L_{lt} and U_{lt} by random variables that are uniformly distributed over the small intervals $[0, \upsilon]$ and $[-\upsilon, 0]$. The stopping criterion in Step 7 uses the pairwise t-test to compare the performances of the prices obtained at different iterations. We describe the stopping criterion in detail in Section 4.3.

3 Charging Uniform Prices over the Planning Horizon

Up to this point, we have assumed that the prices charged over a particular lane at different time periods can be different. We now consider the situation where the carrier has to announce a single price for each lane that is valid over the whole planning horizon. Therefore, we have to have $p_{l1} = p_{l2} = \ldots = p_{lT}$ for all $l \in \mathcal{L}$. Consequently, if the price charged over lane l at time period t is adjusted, then the price charged over this lane for every time period has to be adjusted. This implies that the sample path-based directional derivatives we need are characterized by the limits

$$
\lim_{\alpha \uparrow \downarrow 0} \frac{1}{\alpha} \mathbb{E}\left\{ F_1^\pi\left(r_1, \hat{p} + \alpha \sum_{t \in \mathcal{T}} e_{lt}, D(\hat{p} + \alpha \sum_{t \in \mathcal{T}} e_{lt})\right) - F_1^\pi\left(r_1, \hat{p}, D(\hat{p})\right) \,\Big|\, D(\hat{p}) = \tilde{D}(\hat{p}) \right\}.
$$

The next proposition, shown in the appendix, is a natural extension of Proposition 5.

Proposition 6 *For the fixed prices $\hat{p}$ and load realizations $\tilde{D}(\hat{p})$, we have*

$$\lim_{\alpha \uparrow 0} \frac{1}{\alpha} \mathbb{E}\left\{F_1^\pi\left(r_1, \hat{p} + \alpha \sum_{t \in \mathcal{T}} e_{lt}, D(\hat{p} + \alpha \sum_{t \in \mathcal{T}} e_{lt})\right) - F_1^\pi\left(r_1, \hat{p}, D(\hat{p})\right) \,\Big|\, D(\hat{p}) = \tilde{D}(\hat{p})\right\}$$

$$= \sum_{t \in \mathcal{T}} \dot{\rho}_{lt}(\hat{p}_{lt})\, \hat{x}_{lt} + \sum_{t \in \mathcal{T}} \dot{\lambda}_{lt}(\hat{p}_{lt})\left\{F_t^\pi\left(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) + e_{lt}\right) - F_t^\pi\left(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\right)\right\} \quad (30)$$

$$\lim_{\alpha \downarrow 0} \frac{1}{\alpha} \mathbb{E}\left\{F_1^\pi\left(r_1, \hat{p} + \alpha \sum_{t \in \mathcal{T}} e_{lt}, D(\hat{p} + \alpha \sum_{t \in \mathcal{T}} e_{lt})\right) - F_1^\pi\left(r_1, \hat{p}, D(\hat{p})\right) \,\Big|\, D(\hat{p}) = \tilde{D}(\hat{p})\right\}$$

$$= \sum_{t \in \mathcal{T}} \dot{\rho}_{lt}(\hat{p}_{lt})\, \hat{x}_{lt} + \sum_{t \in \mathcal{T}} \dot{\lambda}_{lt}(\hat{p}_{lt}) \frac{\tilde{D}_{lt}(\hat{p}_{lt})}{\lambda_{lt}(\hat{p}_{lt})}\left\{F_t^\pi\left(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\right) - F_t^\pi\left(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) - e_{lt}\right)\right\}. \quad (31)$$

Using Proposition 6, we can modify Algorithm 1 to find a *good* set of prices that are fixed over the planning horizon. If we use $h_l^{\pi-}(r_1, \hat{p}, \tilde{D}(\hat{p}))$ and $h_l^{\pi+}(r_1, \hat{p}, \tilde{D}(\hat{p}))$ to respectively denote the right sides of (30) and (31), then all we need is to compute $h_l^{\pi-}(r_1^n, p^n, \tilde{D}^n)$ and $h_l^{\pi+}(r_1^n, p^n, \tilde{D}^n)$ for all $l \in \mathcal{L}$ in Step 5 and modify Step 6 as

$$p_{lt}^{n+1} = \begin{cases} p_{lt}^n + \gamma_l^n\, h_l^{\pi-}(r_1^n, p^n, \tilde{D}^n) & \text{if } h_l^{\pi-}(r_1^n, p^n, \tilde{D}^n) < 0 \text{ and} \\ & \qquad\qquad -h_l^{\pi-}(r_1^n, p^n, \tilde{D}^n) > h_l^{\pi+}(r_1^n, p^n, \tilde{D}^n) \\ p_{lt}^n + \gamma_l^n\, h_l^{\pi+}(r_1^n, p^n, \tilde{D}^n) & \text{if } h_l^{\pi+}(r_1^n, p^n, \tilde{D}^n) > 0 \text{ and} \\ & \qquad\qquad h_l^{\pi+}(r_1^n, p^n, \tilde{D}^n) \geq -h_l^{\pi-}(r_1^n, p^n, \tilde{D}^n) \\ p_{lt}^n & \text{otherwise,} \end{cases}$$

where γ_l^n is a step size parameter. Therefore, if we start the algorithm with prices satisfying $p_{l1}^1 = p_{l2}^1 = \ldots = p_{lT}^1$ for all $l \in \mathcal{L}$, then the prices satisfy $p_{l1}^n = p_{l2}^n = \ldots = p_{lT}^n$ for all $l \in \mathcal{L}$ and for any intermediate iteration n.

It is easy to see that Proposition 6 can be extended to cover the case where the planning horizon is partitioned into k disjoint time intervals, say $\{1, \ldots, n_1\}, \{n_1 + 1, \ldots, n_2\}, \ldots, \{n_{k-1} + 1, \ldots, T\}$, and the price for a particular lane cannot vary during a particular time interval. For example, this situation occurs when the planning horizon is multiple weeks and the carrier cannot announce different prices for different days of a particular week.

4 Computational Experiments

The main objective of this section is to show that Algorithm 1 yields high-quality prices. In Section 4.1, we begin by describing our experimental setup. Section 4.2 describes the iterative improvement heuristic mentioned in Section 1.1 that one can use to find a *good* fleet management policy and a *good* set of prices simultaneously. In Section 4.3, we dwell on the step size parameters and the stopping criteria. Section 4.4 compares the performances of the three methods that we propose to approximate $F_t^\pi\left(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}\right) - F_t^\pi\left(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\right)$ in Section 2. Sections 4.5 and 4.6 present the results of two sets of experiments. In the first set of experiments, we observe how the prices react to

changes in certain problem parameters and make sure that they comply with our expectations. The second set of experiments show that the prices obtained through our pricing approach perform well when compared with the ones obtained through benchmark methods.

4.1 Experimental setup

We use two transportation networks, both involving locations somewhat uniformly spread over the United States. The first network involves 20 locations and the second one involves 40 locations. The initial locations of the vehicles are uniformly distributed over the network. We use fleets consisting of 50, 100 and 250 vehicles. We use planning horizons consisting of 7 and 21 time periods.

We assume that the prices are set on a per-mile basis. In particular, if lane l corresponds to serving a load, then we let $\rho_{lt}(p_{lt}) = \zeta(o_l, d_l) [p_{lt} - c]$, where $\zeta(o_l, d_l)$ is the distance from location o_l to d_l, and c is the per-mile empty repositioning cost. We let $\rho_{lt}(p_{lt}) = -\zeta(o_l, d_l) c$ when lane l corresponds to empty repositioning.

We use two functional forms for $\lambda_{lt}(\cdot)$. The first one is linear and has the form

$$\lambda_{lt}(p_{lt}) = \mu_{lt} \left[-\frac{1}{q_l} \left(\frac{p_{lt}}{\varrho_l} \right) + \frac{1 + q_l}{q_l} \right], \tag{32}$$

with $q_l > 0$, $\mu_{lt} \geq 0$, $\varrho_l > 0$. In the expression above, ϱ_l stands for the prevailing price charged over lane l and μ_{lt} stands for the forecasted number of loads over lane l at time period t given that we continue charging the prevailing price ϱ_l. We note that $\lambda_{lt}(\cdot)$ satisfies $\lambda_{lt}(\varrho_l) = \mu_{lt}$, in which case if we continue charging the prevailing prices, then the expected number of loads over each lane and at each time period is equal to the forecast. We restrict p_{lt} to be in $[0, (1 + q_l)\varrho_l]$ so that $\lambda_{lt}(\cdot)$ does not take negative values. We generate the parameters $\{\varrho_l : l \in \mathcal{L}\}$ and $\{q_l : \in \mathcal{L}\}$ from the uniform distributions respectively over $[0.7, 1.7]$ and $[0.5, 1.5]$. Gorman (2001) uses linear price-demand functions similar to (32). The second functional form we use for $\lambda_{lt}(\cdot)$ is nonlinear and can be written as

$$\lambda_{lt}(p_{lt}) = \mu_{lt} \left[1 + q_l - q_l \left(\frac{p_{lt}}{\varrho_l} \right)^{k_l} \right], \tag{33}$$

with $q_l > 0$, $k_l > 1$, $\mu_{lt} \geq 0$, $\varrho_l > 0$. We note that we have $\lambda_{lt}(\varrho_l) = \mu_{lt}$ similar to (32). We restrict p_{lt} to be in $\left[0, \left(\frac{1+q_l}{q_l} \right)^{1/k_l} \varrho_l \right]$ so that $\lambda_{lt}(\cdot)$ does not take negative values. We generate the parameters $\{\varrho_l : l \in \mathcal{L}\}$, $\{q_l : \in \mathcal{L}\}$ and $\{k_l : l \in \mathcal{L}\}$ from the uniform distributions respectively over $[0.7, 1.7]$, $[0.5, 1.5]$ and $[1, 3]$.

We generate the parameters $\{\mu_{lt} : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ in (32) and (33) in such a way that we have $\sum_{l \in \mathcal{L}} \mu_{lt} = 100$ for all $t \in \mathcal{T}$. Consequently, if we continue charging the prevailing prices, then the

total expected number of loads at each time period is 100 and we obtain balanced problems when we use fleets consisting of 100 vehicles. The term balanced should be taken cautiously here, because one can clearly change the total expected number of loads by adjusting the prices.

Throughout this section, we use $(x_1, x_2, x_3, x_4) \in \{l, n\} \times \{20, 40\} \times \{7, 21\} \times \{50, 100, 250\}$ to denote the characteristics of our test problems, where the first element describes whether the functions $\{\lambda_{lt}(\cdot) : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ have linear or nonlinear form, the second element describes the number of locations in the network, the third element describes the number of time periods in the planning horizon and the fourth element describes the size of the fleet.

4.2 Iterative improvement heuristic

An important point in Algorithm 1 is that a fleet management policy, which performs well with the initial prices, may have room for improvement after the prices have been adjusted considerably. Therefore, in our numerical experiments, we iteratively attempt to improve the fleet management policy. In this way, we hope to reach a *good* fleet management policy and a *good* set of prices simultaneously. Throughout this section, we use $\Psi(\pi)$ to denote the output of Algorithm 1. Thus, $\Psi(\cdot)$ can be understood as a function that takes a fleet management policy and returns a *good* set of prices for this fleet management policy. On the other hand, in the appendix, we describe an algorithm due to Godfrey & Powell (2002) that finds a *good* fleet management policy for given prices p. We use $\Pi(p)$ to denote the output of this algorithm, so that $\Pi(\cdot)$ can be understood as a function that takes a set of prices and returns a set of value function approximations characterizing a *good* fleet management policy for these prices. Algorithm 2, which is described in Figure 3, summarizes our iterative policy improvement idea. In Step 2 of Algorithm 2, we find a *good* set of prices for the current fleet management policy, whereas we try to improve the current fleet management policy in Step 3. We dwell on the stopping criterion in Step 4 in the next section.

4.3 Step size parameters and stopping criteria

This section describes the step size parameters and the stopping criteria used in Algorithms 1 and 2.

Step size parameters. The step size parameter that we use in Algorithm 1 is of the form $\gamma_{lt}^n = \frac{a}{b+n}$, but we modify it to remedy its two shortcomings. First, depending on the values of a and b, this step size parameter may decline too quickly, causing the prices to stall early, or may decline too slowly, causing the prices to oscillate from one iteration to the next. To remedy this, we use the step size parameter $\gamma_{lt}^n = \frac{a}{b+\Phi_{lt}^n}$, where Φ_{lt}^n is the number of times that the price over lane l at time period t oscillates during the first n iterations; see Kesten (1958). That is, we have

$\Phi_{lt}^n = \sum_{k=1}^{n-2} \mathbf{1}\left([p_{lt}^{k+1} - p_{lt}^k][p_{lt}^{k+2} - p_{lt}^{k+1}] < 0\right)$ and $\mathbf{1}(\cdot)$ is the indicator function. The idea is that if the price starts oscillating from one iteration to the next, then the step size parameter should quickly be made smaller. We find that the behavior of Algorithm 1 is relatively insensitive to the choice of b and we let $b = 3$.

Second, the magnitudes of the sample path-based directional derivatives $g_{lt}^{\pi\mp}(r_t^n, p^n, \tilde{D}^n)$ and the price p_{lt}^n can be quite different. Clearly, this becomes problematic when letting $p_{lt}^{n+1} = p_{lt}^n + \frac{a}{3+\Phi_{lt}^n} g_{lt}^{\pi\mp}(r_t^n, p^n, \tilde{D}^n)$ in Step 6 of Algorithm 1. To remedy this, we choose a in such a way that the magnitudes of p_{lt}^n and $\frac{a}{3+\Phi_{lt}^n} g_{lt}^{\pi\mp}(r_t^n, p^n, \tilde{D}^n)$ are comparable. In particular, since the prices range roughly over the interval $[0, 2]$ with the midpoint 1, we let

$$a = \frac{1}{[|g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n)| + |g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n)|]/2}.$$

In this case, Step 6 of Algorithm 1 can be written as

$$p_{lt}^{n+1} = p_{lt}^n + \frac{2}{3 + \Phi_{lt}^n} \frac{g_{lt}^{\pi\mp}(r_t^n, p^n, \tilde{D}^n)}{|g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n)| + |g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n)|}$$

and our choice of a amounts to normalizing the sample path-based directional derivatives. A similar normalization idea for subgradient optimization is described in Nemhauser & Wolsey (1988).

Stopping criteria. The stopping criterion that we use for Algorithm 2 requires keeping incumbent prices and an incumbent policy, say p^ϕ and π^ϕ, throughout the algorithm. Initially, we let $p^\phi = p^1$ and $\pi^\phi = \pi^1$. Every Θ iterations of Algorithm 2, we use the pairwise t-test to check if the prices and the policy obtained at the current iteration provide a better objective value than the incumbent prices and the incumbent policy; see Law & Kelton (2000). In particular, letting p^{n+1} and π^{n+1} respectively be the prices and the policy obtained at iteration n, we generate N samples of $F_1^{\pi^{n+1}}(r_1, p^{n+1}, D(p^{n+1})) - F_1^{\pi^\phi}(r_1, p^\phi, D(p^\phi))$ and test the null hypothesis

$$\mathbb{E}\left\{F_1^{\pi^{n+1}}(r_1, p^{n+1}, D(p^{n+1}))\right\} - \mathbb{E}\left\{F_1^{\pi^\phi}(r_1, p^\phi, D(p^\phi))\right\} \leq 0$$

at 5% significance level. If we fail to reject the null hypothesis in Λ consecutive tests, then we conclude that no significant improvement has been made during the last $\Theta\Lambda$ iterations, and we stop and return the incumbent prices and the incumbent policy. If we reject the null hypothesis, then we update the incumbent prices and the incumbent policy by letting $p^\phi = p^{n+1}$ and $\pi^\phi = \pi^{n+1}$, and continue with the new incumbent prices and the new incumbent policy. For all of our test problems, choosing N around 100-150 yields reasonable Type II error probabilities.

We note that it is important to choose Θ and Λ carefully. A small value for Θ causes the pairwise t-tests to be carried out too frequently and wastes computational effort. Furthermore, the prices should be allowed to improve between the successive tests to avoid premature termination.

On the other hand, a large value for Θ also wastes computational effort by unnecessarily delaying termination. After some experimentation, we let $\Theta = 5$ and $\Lambda = 20$. To illustrate the behavior of our stopping criterion, the solid data series in Figure 4 show the estimates of $\mathbb{E}\big\{F_1^{\pi^n}\big(r_1, p^n, D(p^n)\big)\big\}$ for test problem $(n, 20, 7, 50)$ as a function of the iteration counter n in Algorithm 2. The circles show the estimates of $\mathbb{E}\big\{F_1^{\pi^\phi}\big(r_1, p^\phi, D(p^\phi)\big)\big\}$ for the iterations at which the incumbent prices and the incumbent policy are updated. For given prices p and policy π, we estimate $\mathbb{E}\big\{F_1^{\pi}\big(r_1, p, D(p)\big)\big\}$ by simulating the behavior of policy π under prices p and under different realizations of $D(p)$. The square shows the iteration at which our stopping criterion suggests termination.

We emphasize that our stopping criterion does not guarantee the termination of Algorithm 2 within a finite number of iterations. To address this shortcoming, we put a prespecified limit on the total number of iterations in Algorithm 2. Such hard iteration limits are not ideal, but they are often used due to the lack of good stopping criteria for stochastic quasi-gradient algorithms. We use a similar stopping criterion for Algorithm 1 and also impose a hard iteration limit. The only difference is that since Algorithm 1 assumes that the policy is fixed, we only keep incumbent prices.

It is important to note that when we use Algorithm 1 as a subroutine in Step 2 of Algorithm 2, it is not necessary or clear that Algorithm 1 should be run until termination. In particular, we may consider running Algorithm 1 only for a few iterations rather than running it until our stopping criterion suggests termination or we reach the hard iteration limit. In this way, we simply try to improve the prices in Step 2 of Algorithm 2 rather than try to find the *best* prices that can be obtained by Algorithm 1. For all of our test problems, running Algorithm 1 only for a few iterations in Step 2 of Algorithm 2 provides essentially the same results as running it until termination. The computational results in the subsequent sections are for the case where we run Algorithm 1 until termination in Step 2 of Algorithm 2.

Noting the discussion at the end of Section 2, to ensure that the prices that we obtain at each iteration of Algorithm 1 lie in $\mathcal{P}$, we also tried perturbing the sample path-based directional derivatives by random variables that are uniformly distributed over the small interval $[-\upsilon, \upsilon]$. Since the magnitudes of the sample path-based directional derivatives can be as large as several hundreds and the prices range roughly over the interval $[0, 2]$, we let $\upsilon = 10^{-4}$. For all of our test problems, using these perturbations provides essentially the same results as completely omitting them. Given the finite granularity of computer arithmetic, one can admittedly argue about the plausibility of these perturbations. Nevertheless, it is comforting to see that our pricing approach is robust to small perturbations of the sample path-based directional derivatives. The computational results in the subsequent sections are for the case where we do not explicitly perturb the sample path-based directional derivatives.

4.4 Computing the profit impact of an additional load

In Section 2, we propose three methods that can be used to approximate $F_t^{\pi}\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}\big) - F_t^{\pi}\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big)$. The numerical difference-based method requires $2|\mathcal{L}||\mathcal{T}|$ additional simulations, where each additional simulation requires solving $|\mathcal{T}|$ min-cost network flow problems. Since the number of lanes is on the order of hundreds for our test problems, this method is clearly not practical. To see the tradeoff between the other two methods, we apply Algorithm 2 twice to six test problems, once using the value function approximation-based method and once using the perturbation analysis-based method. We let (p^v, π^v) and (p^a, π^a) respectively be the pairs of the prices and the policy obtained by Algorithm 2 by using the value function approximation-based method and the perturbation analysis-based method. The first two columns in Table 1 show the estimates of $\mathbb{E}\big\{F_1^{\pi^v}\big(r_1, p^v, D(p^v)\big)\big\}$ and the CPU seconds per iteration for Algorithm 2 when it uses the value function approximation-based method, whereas the last two columns show the estimates of $\mathbb{E}\big\{F_1^{\pi^a}\big(r_1, p^a, D(p^a)\big)\big\}$ and the CPU seconds per iteration for Algorithm 2 when it uses the perturbation analysis-based method. Table 1 indicates that the performances of the two methods are essentially the same, but the perturbation analysis-based method takes considerably more time. Nevertheless, since it has some theoretical foundation, we use the perturbation analysis-based method in our computational experiments.

4.5 General behavior

In this section, we observe our pricing approach from a qualitative viewpoint and ensure that its behavior complies with our expectations. In particular, we show that the prices have a tendency to decline as the vehicles become more abundant and to become more uniform as the differences in the regional market conditions diminish. We also demonstrate that when we apply our pricing approach starting from different initial prices, the performances of the final prices we obtain are similar.

Price reactions to increasing vehicle availability. To show that the prices obtained through our pricing approach decline as the fleet size increases, we apply Algorithm 2 to problems with varying fleet sizes. Table 2 shows the summary statistics for the prices obtained by Algorithm 2. Letting $\{p_{lt}^o : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ be the prices obtained by this algorithm, the entries of the table are the mean, and the 20-th and 80-th percentiles of $\{p_{lt}^o : l \in \mathcal{L}, \ t \in \mathcal{T}\}$. That is, for example, the mean is computed as $\frac{1}{|\mathcal{T}||\mathcal{L}|} \sum_{t \in \mathcal{T}} \sum_{l \in \mathcal{L}} p_{lt}^o$ and the 20-th percentile is computed as the 20-th percentile of the data $\{p_{lt}^o : l \in \mathcal{L}, \ t \in \mathcal{T}\}$. Table 2 confirms our expectation that the prices should decline as the fleet size increases.

Price reactions to diminishing differences in regional market conditions. We now consider problems where we have $\lambda_{lt}(\cdot) \approx \lambda_{l't'}(\cdot)$ and $\rho_{lt}(\cdot) \approx \rho_{l't'}(\cdot)$ for all $l, l' \in \mathcal{L}, \ t, t' \in \mathcal{T}$. In this case,

since the load realizations and the per-load profits react to price changes in the same manner for all $l \in \mathcal{L}$, $t \in \mathcal{T}$, we should expect the prices charged over different lanes and at different time periods to be similar. Table 3 shows the standard deviation, and the 20-th and 80-th percentiles of the prices obtained through our pricing approach when applied to problems with these characteristics. For every test problem, the initial prices over different lanes and at different time periods range over the interval $[0, 2.40]$, with a standard deviation of 0.73, and 20-th and 80-th percentiles of respectively 0.48 and 1.69. Table 3 indicates that the 20-th and 80-th percentiles of the final prices are close to each other and the standard deviations are quite small, which imply that the final prices over different lanes and at different time periods are close to each other.

Price reactions to different initial prices. Our pricing approach does not have a convergence guarantee. Furthermore, the objective function of problem (8) is not necessarily concave. Therefore, if we apply our pricing approach starting from different initial prices, then the performances of the final prices we obtain can be different. To ensure that this is not a major issue, we apply our pricing approach starting from three different initial prices, say $\tilde{p}^1, \hat{p}^1$ and $\bar{p}^1$. The coefficients of correlation among $\{\tilde{p}_{lt}^1 : l \in \mathcal{L}, \ t \in \mathcal{T}\}$, $\{\hat{p}_{lt}^1 : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ and $\{\bar{p}_{lt}^1 : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ are -0.96, -0.36 and 0.36. We let $(\tilde{p}^o, \tilde{\pi}^o)$, $(\hat{p}^o, \hat{\pi}^o)$ and $(\bar{p}^o, \bar{\pi}^o)$ respectively be the pairs of the prices and the policy obtained by Algorithm 2 starting from the initial prices $\tilde{p}^1$, $\hat{p}^1$ and $\bar{p}^1$. Table 4 shows the estimates of $\mathbb{E}\big\{F_1^{\tilde{\pi}^o}\big(r_1, \tilde{p}^o, D(\tilde{p}^o)\big)\big\}$, $\mathbb{E}\big\{F_1^{\hat{\pi}^o}\big(r_1, \hat{p}^o, D(\hat{p}^o)\big)\big\}$ and $\mathbb{E}\big\{F_1^{\bar{\pi}^o}\big(r_1, \bar{p}^o, D(\bar{p}^o)\big)\big\}$, and indicates that the performances of the final prices that are obtained by starting from different initial prices are close to each other. Furthermore, Figure 5 gives plots of $\{(\hat{p}_{lt}^1, \bar{p}_{lt}^1) : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ and $\{(\hat{p}_{lt}^o, \bar{p}_{lt}^o) : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ for test problem $(l, 20, 21, 250)$, and indicates that $\hat{p}^o$ and $\bar{p}^o$ are in close agreement, although $\hat{p}^1$ and $\bar{p}^1$ are not.

4.6 Solution quality

This section tests the quality of the prices obtained by Algorithm 2. We present two sets of experiments. The first set of experiments compare the performance of our pricing approach with that of a deterministic nonlinear program. In the second set of experiments, we assume that the price over every lane is fixed, except for three lanes. We carry out an exhaustive numerical search to find the best prices to charge over these three lanes.

Comparisons against deterministic nonlinear program. When dealing with a problem that involves uncertainty, a common engineering practice is to assume that all random variables take on their expected values and to formulate a deterministic approximation to the original problem. Noting

that $\mathbb{E}\left\{D_{lt}(p_{lt})\right\} = \lambda_{lt}(p_{lt})$, a deterministic pricing model can be formulated as

$$\max \quad \sum_{t\in\mathcal{T}}\sum_{l\in\mathcal{L}} \rho_{lt}(p_{lt})\, x_{lt} \tag{34}$$

$$\text{subject to} \quad \sum_{l\in\mathcal{L},o_l=i} x_{l1} = r_{i1} \qquad \text{for all } i \in \mathcal{I}$$

$$\sum_{l\in\mathcal{L},d_l=i} x_{l,t-1} - \sum_{l\in\mathcal{L},o_l=i} x_{lt} = 0 \qquad \text{for all } i \in \mathcal{I},\ t = 2,\ldots,T$$

$$x_{lt} \leq \lambda_{lt}(p_{lt}) \qquad \text{for all } l \in \mathcal{L},\ t \in \mathcal{T}$$

$$p_{lt},\, x_{lt} \geq 0 \qquad \text{for all } l \in \mathcal{L},\ t \in \mathcal{T}.$$

The objective function accounts for the total profit over the whole planning horizon. The first two sets of constraints are the flow balance constraints similar to (1) and (2), whereas the third set of constraints are the demand availability constraints similar to (3). One can consider imposing integrality constraints on the variables $\{x_{lt} : l \in \mathcal{L},\ t \in \mathcal{T}\}$, but this would make the problem intractable. Due to our choice of the parameters in (32) and (33), $\lambda_{lt}(\cdot)$ is concave, and hence, the feasible region of the problem above is convex. However, since the objective function is not necessarily concave, we solve problem (34) by starting a nonlinear programming package from 30 different initial solutions. We let $p^d = \{p_{lt}^d : l \in \mathcal{L},\ t \in \mathcal{T}\}$ be the prices obtained through this deterministic benchmark method.

Letting ζ^d be the optimal objective value of problem (34) and $\{V_t^d(\cdot) : t \in \mathcal{T}\}$ be the value functions obtained by applying the backward recursion in (4) under the prices p^d, it is well-known that we have $\mathbb{E}\left\{F_1^\pi\left(r_1, p^d, D(p^d)\right)\right\} \leq V_1^d(r_1) \leq \zeta^d$ for any fleet management policy π; see Birge & Louveaux (1997). However, there does not exist a relationship between ζ^d and $\sup_p \mathbb{E}\left\{F_1^\pi\left(r_1, p, D(p)\right)\right\}$, which is the optimal objective value of the problem we want to solve. Consequently, we use problem (34) only to obtain the prices p^d.

To test the quality of the prices p^d, we find a *good* policy π^d for these prices and evaluate the performance of the policy π^d under the prices p^d. That is, we set $\pi^d = \Pi(p^d)$ and estimate $\mathbb{E}\left\{F_1^{\pi^d}\left(r_1, p^d, D(p^d)\right)\right\}$. Table 5 shows our findings. Letting (p^o, π^o) be the prices and the policy obtained by Algorithm 2, the first and second columns of this table show the estimates of $\mathbb{E}\left\{F_1^{\pi^o}\left(r_1, p^o, D(p^o)\right)\right\}$ and $\mathbb{E}\left\{F_1^{\pi^d}\left(r_1, p^d, D(p^d)\right)\right\}$. The third and fourth columns show the number of iterations and the CPU seconds required for Algorithm 2 to terminate. Finally, the fifth column shows what percent of the CPU seconds is spent on checking the stopping criteria. The results indicate that the prices obtained by Algorithm 2 perform significantly better than the ones obtained by problem (34). The number of iterations required for termination seems to increase slightly as the number of locations increases. A significant portion of the CPU seconds is spent on checking the stopping criteria. This is a well-known difficulty for stochastic gradient algorithms. In particular,

24

checking the quality of a solution may sometimes take longer than obtaining the solution in the first place; see Higle & Zhao (2004).

Comparisons against exhaustive numerical search. In this set of experiments, we choose three lanes, say l_1, l_2 and l_3, and assume that the prices charged over only these lanes are decision variables. The prices charged over the remaining lanes are fixed at predetermined values, say $\tilde{p} = \{\tilde{p}_{lt} : l \in \mathcal{L},\ t \in \mathcal{T}\}$. We also assume that the prices charged over these lanes cannot vary over the planning horizon, so that the number of price decision variables is only three.

For a given mesh size $\alpha > 0$, our exhaustive numerical search tests the performances of the prices

$$\tilde{p} + \alpha \left[k_1 \sum_{t \in \mathcal{T}} e_{l_1 t} + k_2 \sum_{t \in \mathcal{T}} e_{l_2 t} + k_3 \sum_{t \in \mathcal{T}} e_{l_3 t} \right]$$

for different values of $(k_1, k_2, k_3) \in \mathbb{Z}^3$. We test every value of (k_1, k_2, k_3) in $\mathcal{Z}_1 \times \mathcal{Z}_2 \times \mathcal{Z}_3$, where

$$\mathcal{Z}_i = \{ k_i \in \mathbb{Z} : \tilde{p}_{l_i t} + \alpha\, k_i \geq 0,\ \lambda_{l_i t}(\tilde{p}_{l_i t} + \alpha\, k_i) \geq 0 \ \text{ for all } t \in \mathcal{T} \}.$$

Letting $\Delta_l = \sum_{t \in \mathcal{T}} e_{lt}$ for notational brevity, to assess the quality of the prices $\tilde{p} + \alpha\,[k_1\,\Delta_{l_1} + k_2\,\Delta_{l_2} + k_3\,\Delta_{l_3}]$, we first find a *good* policy $\pi(k_1, k_2, k_3)$ by setting $\pi(k_1, k_2, k_3) = \Pi(\tilde{p} + \alpha\,[k_1\,\Delta_{l_1} + k_2\,\Delta_{l_2} + k_3\,\Delta_{l_3}])$. Then, we evaluate the performance of this policy by estimating

$$\mathbb{E}\left\{ F_1^{\pi(k_1, k_2, k_3)}\big(r_1, \tilde{p} + \alpha\,[k_1\,\Delta_{l_1} + k_2\,\Delta_{l_2} + k_3\,\Delta_{l_3}], D(\tilde{p} + \alpha\,[k_1\,\Delta_{l_1} + k_2\,\Delta_{l_2} + k_3\,\Delta_{l_3}])\big) \right\}.$$

By testing every value of (k_1, k_2, k_3) in $\mathcal{Z}_1 \times \mathcal{Z}_2 \times \mathcal{Z}_3$, we find (k_1^*, k_2^*, k_3^*) that maximizes the quantity above. This is the maximum profit that can be attained by the numerical search. Letting (p^o, π^o) be the prices and the policy obtained by Algorithm 2, Table 6 shows the ratio of the estimate of $\mathbb{E}\{ F_1^{\pi^o}(r_1, p^o, D(p^o)) \}$ to the maximum profit that can be attained by the numerical search. The results indicate that the prices obtained through our pricing approach perform as well as the ones obtained through the numerical search.

5 Conclusions

In this paper, we showed how to coordinate the pricing and fleet management decisions of a freight carrier assuming that a particular type of model is responsible from making the fleet management decisions. We developed an algorithm to obtain sample path-based directional derivatives of the objective function with respect to the prices and used this information to search for a *good* set of prices. Numerical experiments showed that the proposed pricing approach yields high-quality prices.

Our main results, Propositions 5 and 6, only assume that the underlying fleet management policy is Markovian, and the results of the decision and state transition functions do not change

with infinitesimal perturbations in the prices. Therefore, (18), (19), (30) and (31) continue to hold for an arbitrary fleet management policy as long as it satisfies these assumptions. Nevertheless, computing $F_t^\pi(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) \mp e_{lt}) - F_t^\pi(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}))$ may be difficult for arbitrary fleet management policies, whereas this is relatively easy for the class of policies proposed by Godfrey & Powell (2002).

6 Acknowledgements

A Omitted Proofs

The next result is useful for proving Lemma 1.

Lemma 7 *For fixed $W \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, consider the linear programming problem $\max\{q^T y : W y \leq b, y \geq 0\}$ with a bounded and nonempty set of feasible solutions. In this case, there exist a finite number of at most $(n-1)$-dimensional subspaces of $\mathbb{R}^n$ such that if the linear programming problem has multiple optimal solutions for some $q \in \mathbb{R}^n$, then q must lie in one of these subspaces.*

Proof of Lemma 7 Defining the necessary slack decision variables, and augmenting the constraint matrix and the objective function coefficients appropriately, we write the problem $\max\{q^T y : W y \leq b, y \geq 0\}$ as $\max\{c^T x : A x = b, x \geq 0\}$, where $A \in \mathbb{R}^{m \times (n+m)}$ and $c \in \mathbb{R}^{n+m}$. Assume that the latter problem has multiple optimal solutions for some $c \in \mathbb{R}^{n+m}$. In this case, we can find a basis B with the following properties. 1) The basic feasible solution corresponding to the basis B is an optimal solution to the problem $\max\{c^T x : A x = b, x \geq 0\}$. 2) If we let x^* be the basic feasible solution corresponding to the basis B and partition x^* as $[x_B^*, x_N^*]$, where x_B^* and x_N^* respectively correspond to the basic and nonbasic parts, then the reduced cost corresponding to one of the elements of x_N^* is zero. 3) This element of x_N^* corresponds to one of the original decision variables y (rather than one of the slack decision variables).

We partition the matrix A as $[B, N]$, where B is the basis above and N is the remaining portion. We also partition the vector c as $[c_B, c_N]$, where c_B and c_N respectively correspond to x_B^* and x_N^*. The second property above implies that one element of the vector $(B^{-1}N)^T c_B - c_N$ is equal to zero. The third property above implies that this element corresponds to one of the original decision

variables. Therefore, one element of the vector q can be written as a linear combination of the other elements. The result follows by noting that there are only a finite number of possible bases. $\qquad\square$

Proof of Lemma 1 To prove the result, we show that there exist a finite number of at most $(|\mathcal{L}| - 1)$-dimensional subspaces of $\mathbb{R}^{|\mathcal{L}|}$ such that if problem (6) has multiple optimal solutions for some state vector r_t and load realizations $\tilde{D}_t(p_t)$, then $\rho_t(p_t)$ must lie in one of these subspaces.

Since the number of vehicles is finite, there are only a finite number of possible state vectors. Furthermore, letting K be the number of vehicles, it is enough to concentrate on the load realizations $\tilde{D}_t(p_t)$ that satisfy $\tilde{D}_{lt}(p_{lt}) \leq K$ for all $l \in \mathcal{L}$. Whenever we have $\tilde{D}_{lt}(p_{lt}) > K$ for some $l \in \mathcal{L}$, we can assume that $\tilde{D}_{lt}(p_{lt}) = K$ without changing the optimal solution to problem (6). Therefore, there are only a finite number of possible load realizations that are of interest.

For fixed state vector r_t and load realizations $\tilde{D}_t(p_t)$, we assume that problem (6) has multiple optimal solutions. Since problem (6) is a min-cost network flow problem as mentioned in Section 1.1, Lemma 7, with little extra work, implies that $\rho_t(p_t)$ must lie in one of the finite number of at most $(|\mathcal{L}| - 1)$-dimensional subspaces of $\mathbb{R}^{|\mathcal{L}|}$. The result follows by noting that there are only a finite number of state vectors and load realizations that are of interest. $\qquad\square$

Since the proofs of Lemmas 1 and 3 are related, we prefer to prove Lemma 3 before Lemma 2.

Proof of Lemma 3 We fix state vector r_t and load realizations $\tilde{D}_t(p_t)$. Since $p_t \in \mathcal{P}_t$, the optimal solution to problem (6) is unique. Furthermore, since the number of vehicles is finite and we have the integrality constraints, there are only a finite number of feasible solutions to problem (6). Therefore, the objective value obtained by the optimal solution to problem (6) is larger than the objective value obtained by all other feasible solutions by a strictly positive quantity. Since problem (6) is a min-cost network flow problem and $\rho_t(p_t)$ is a continuous function of p_t, the optimal objective value of problem (6) is a continuous function of the prices. All of these observations imply that we can change the prices by a small amount $\epsilon(p_t, r_t, \tilde{D}_t(p_t)) > 0$ without changing the optimal solution to problem (6). We use the arguments p_t, r_t and $\tilde{D}_t(p_t)$ to emphasize the dependence of $\epsilon(p_t, r_t, \tilde{D}_t(p_t))$ on these quantities. The result follows by noting that there are only a finite number of state vectors and load realizations that are of interest. $\qquad\square$

Proof of Lemma 2 We let $\epsilon > 0$ be such that (22) and (23) hold for all $\alpha \in [-\epsilon, \epsilon]$ and all realizations of $D_t(\hat{p}_t + \alpha\, e_l)$. For $\alpha \in [-\epsilon, 0)$, (25) implies that

$$\left| \frac{1}{\alpha} \left\{ F_t^{\pi}\left(\hat{r}_t, \hat{p} + \alpha\, e_{lt}, D(\hat{p} + \alpha\, e_{lt})\right) - F_t^{\pi}\left(\hat{r}_t, \hat{p}, D(\hat{p} + \alpha\, e_{lt})\right) \right\} \right| \leq \left| \frac{K}{\alpha} \left[\rho_{lt}(\hat{p}_{lt} + \alpha) - \rho_{lt}(\hat{p}_{lt}) \right] \right|$$

for all realizations of $D(\hat{p} + \alpha\, e_{lt})$, where K is the number of vehicles. Letting $g_1(\alpha) = \left| \frac{K}{\alpha} \left[\rho_{lt}(\hat{p}_{lt} + \alpha) - \rho_{lt}(\hat{p}_{lt}) \right] \right|$ and using the fact that $\rho_{lt}(\cdot)$ is an increasing function, we have $\lim_{\alpha \uparrow 0} g_1(\alpha) = K\, \dot{\rho}_{lt}(\hat{p}_{lt})$,

which is finite. Furthermore, $g_1(-\epsilon)$ is also finite and $g_1(\cdot)$ is a continuous function over $[-\epsilon, 0)$. By using the last three facts, it is easy to see that there exists M_1 such that $g_1(\alpha) \leq M_1 < \infty$ for all $\alpha \in [-\epsilon, 0)$. This implies that we have

$$\left| \frac{1}{\alpha} \mathbb{E}\left\{ F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha e_{lt}, D(\hat{p} + \alpha\, e_{lt})\right) - F_t^\pi\left(\hat{r}_t, \hat{p}, D(\hat{p} + \alpha\, e_{lt})\right) \,\big|\, D(\hat{p})\right\} \right| \leq M_1 \tag{35}$$

for all $\alpha \in [-\epsilon, 0)$ and all realizations of $D(\hat{p})$.

Letting M_2 be such that $|\rho_{lt}(\hat{p}_{lt})| \leq M_2 < \infty$ for all $l \in \mathcal{L}$, $t \in \mathcal{T}$, the cumulative profit over the whole planning horizon under prices $\hat{p}$ is bounded by $K\, M_2\, |\mathcal{T}|$. Therefore, we have $\left| F_t^\pi\left(\hat{r}_t, \hat{p}, D(\hat{p}) + n\, e_{lt}\right) \right| \leq K\, M_2\, |\mathcal{T}|$ for all $n = 0, 1, \ldots$. In this case, the same argument used to obtain (28) yields

$$\left| \frac{1}{\alpha} \mathbb{E}\left\{ F_t^\pi\left(\hat{r}_t, \hat{p}, D(\hat{p} + \alpha\, e_{lt})\right) - F_t^\pi\left(\hat{r}_t, \hat{p}, D(\hat{p})\right) \,\big|\, D(\hat{p})\right\} \right|$$

$$= \left| \frac{1}{\alpha} \left[\lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt}) \right] \left\{ F_t^\pi\left(\hat{r}_t, \hat{p}, D(\hat{p}) + e_{lt}\right) - F_t^\pi\left(\hat{r}_t, \hat{p}, D(\hat{p})\right) \right\} + \frac{o(\alpha)}{\alpha} \right|.$$

$$\leq \left| \frac{2\, K\, M_2\, |\mathcal{T}|}{\alpha} \left[\lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt}) \right] \right| + \left| \frac{g_2(\alpha)}{\alpha} \right|,$$

where $g_2(\cdot)$ is a finite-valued, continuous function that satisfies $\lim_{\alpha \uparrow 0} g_2(\alpha)/\alpha = 0$. Letting $g_3(\alpha) = \left| \frac{2KM_2|\mathcal{T}|}{\alpha} \left[\lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt}) \right] \right| + \left| \frac{g_2(\alpha)}{\alpha} \right|$ and using the fact that $\lambda_{lt}(\cdot)$ is a decreasing function, we have $\lim_{\alpha \uparrow 0} g_3(\alpha) = -2\, K\, M_2\, |\mathcal{T}|\, \dot{\lambda}_{lt}(\hat{p}_{lt})$, which is finite. Furthermore, $g_3(-\epsilon)$ is also finite and $g_3(\cdot)$ is a continuous function over $[-\epsilon, 0)$. Therefore, there exists M_3 such that $g_3(\alpha) \leq M_3 < \infty$ for all $\alpha \in [-\epsilon, 0)$. Noting (35), this implies that if we let $M = M_1 + M_3$, then (12) holds for all $\alpha \in [-\epsilon, 0)$ and all realizations of $D(\hat{p})$.

Using the same argument, we can find $M_4 < \infty$ such that if we let $M = M_4$, then (12) holds for all $\alpha \in (0, \epsilon]$ and all realizations of $D(\hat{p})$. Consequently, letting $M = \max\{M_1 + M_3, M_4\}$ suffices. $\square$

Proof of Proposition 6 We only show that (30) holds. For notational brevity, we let $\Delta_{lt} = \sum_{s=t}^T e_{ls}$, with $\Delta_{l,T+1} = 0$. The conditional expectation on the left side of (30) can be written as

$$\mathbb{E}\left\{ F_1^\pi\left(r_1, \hat{p} + \alpha\, \Delta_{l1}, D(\hat{p} + \alpha\, \Delta_{l1})\right) - F_1^\pi\left(r_1, \hat{p}, D(\hat{p})\right) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p})\right\}$$

$$= \sum_{t \in \mathcal{T}} \mathbb{E}\left\{ F_1^\pi\left(r_1, \hat{p} + \alpha\, \Delta_{lt}, D(\hat{p} + \alpha\, \Delta_{lt})\right) - F_1^\pi\left(r_1, \hat{p} + \alpha\, \Delta_{l,t+1}, D(\hat{p} + \alpha\, \Delta_{l,t+1})\right) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p})\right\}.$$

For each one of the terms in the summation on the right side above, Lemma 8 below shows that

$$\mathbb{E}\left\{ F_1^\pi\left(r_1, \hat{p} + \alpha\, \Delta_{lt}, D(\hat{p} + \alpha\, \Delta_{lt})\right) - F_1^\pi\left(r_1, \hat{p} + \alpha\, \Delta_{l,t+1}, D(\hat{p} + \alpha\, \Delta_{l,t+1})\right) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p})\right\}$$

$$= \mathbb{E}\left\{ F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\, \Delta_{lt}, D(\hat{p} + \alpha\, \Delta_{lt})\right) - F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\, \Delta_{l,t+1}, D(\hat{p} + \alpha\, \Delta_{l,t+1})\right) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p})\right\}. \tag{36}$$

We write the expression on the right side above as

$$\mathbb{E}\left\{ F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\, \Delta_{lt}, D(\hat{p} + \alpha\, \Delta_{lt})\right) - F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\, \Delta_{l,t+1}, D(\hat{p} + \alpha\, \Delta_{lt})\right) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p})\right\}$$

$$+ \mathbb{E}\left\{ F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\, \Delta_{l,t+1}, D(\hat{p} + \alpha\, \Delta_{lt})\right) - F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\, \Delta_{l,t+1}, D(\hat{p} + \alpha\, \Delta_{l,t+1})\right) \,\big|\, D(\hat{p}) = \tilde{D}(\hat{p})\right\}.$$

For the first term above, Lemma 9 below shows that

$$\lim_{\alpha\uparrow 0}\frac{1}{\alpha}\mathbb{E}\Big\{F_t^\pi\big(\hat{r}_t,\hat{p}+\alpha\,\Delta_{lt},D(\hat{p}+\alpha\,\Delta_{lt})\big)$$
$$-F_t^\pi\big(\hat{r}_t,\hat{p}+\alpha\,\Delta_{l,t+1},D(\hat{p}+\alpha\,\Delta_{lt})\big)\,\big|\,D(\hat{p})=\tilde{D}(\hat{p})\Big\}=\dot{\rho}_{lt}(\hat{p}_{lt})\,\hat{x}_{lt},\quad(37)$$

whereas, for the second term, Lemma 10 below shows that

$$\lim_{\alpha\uparrow 0}\frac{1}{\alpha}\mathbb{E}\Big\{F_t^\pi\big(\hat{r}_t,\hat{p}+\alpha\,\Delta_{l,t+1},D(\hat{p}+\alpha\,\Delta_{lt})\big)$$
$$-F_t^\pi\big(\hat{r}_t,\hat{p}+\alpha\,\Delta_{l,t+1},D(\hat{p}+\alpha\,\Delta_{l,t+1})\big)\,\big|\,D(\hat{p})=\tilde{D}(\hat{p})\Big\}$$
$$=\dot{\lambda}_{lt}(\hat{p}_{lt})\Big\{F_t^\pi\big(\hat{r}_t,\hat{p},\tilde{D}(\hat{p})+e_{lt}\big)-F_t^\pi\big(\hat{r}_t,\hat{p},\tilde{D}(\hat{p})\big)\Big\}.\quad(38)$$

This establishes the result. □

Lemma 8 *For $\hat{p}\in\mathcal{P}$, the equality in (36) holds.*

Proof of Lemma 8 We fix α, and let $\tilde{D}(\hat{p}+\alpha\,\Delta_{lt})$ and $\tilde{D}(\hat{p}+\alpha\,\Delta_{l,t+1})$ respectively be fixed realizations of $D(\hat{p}+\alpha\,\Delta_{lt})$ and $D(\hat{p}+\alpha\,\Delta_{l,t+1})$. We let $\{\bar{x}_t : t\in\mathcal{T}\}$ and $\{\bar{r}_t : t\in\mathcal{T}\}$ be the sequences of decision and state vectors visited by policy π under prices $\hat{p}+\alpha\,\Delta_{lt}$ and load realizations $\tilde{D}(\hat{p}+\alpha\,\Delta_{lt})$. We let $\{\tilde{x}_t : t\in\mathcal{T}\}$ and $\{\tilde{r}_t : t\in\mathcal{T}\}$ be the sequences of decision and state vectors visited by policy π under prices $\hat{p}+\alpha\,\Delta_{l,t+1}$ and load realizations $\tilde{D}(\hat{p}+\alpha\,\Delta_{l,t+1})$. The prices $\hat{p}$, $\hat{p}+\alpha\,\Delta_{lt}$ and $\hat{p}+\alpha\,\Delta_{l,t+1}$ coincide at time periods $\{1,\dots,t-1\}$. In this case, given that $D(\hat{p})=\tilde{D}(\hat{p})$, we can follow the same argument as in the proof of Lemma 4 to obtain $\hat{x}_s=\bar{x}_s=\tilde{x}_s$ for all $s\in\{1,\dots,t-1\}$ and $\hat{r}_s=\bar{r}_s=\tilde{r}_s$ for all $s\in\{1,\dots,t\}$. Therefore, the result follows by noting that the conditional expectation on the left side of (36) is equal to

$$\mathbb{E}\Big\{\textstyle\sum_{s=1}^{t-1}\rho_s(\hat{p}_s)\cdot\bar{x}_s+F_t^\pi\big(\bar{r}_t,\hat{p}+\alpha\,\Delta_{lt},D(\hat{p}+\alpha\,\Delta_{lt})\big)$$
$$-\textstyle\sum_{s=1}^{t-1}\rho_s(\hat{p}_s)\cdot\tilde{x}_s-F_t^\pi\big(\tilde{r}_t,\hat{p}+\alpha\,\Delta_{l,t+1},D(\hat{p}+\alpha\,\Delta_{l,t+1})\big)\,\big|\,D(\hat{p})=\tilde{D}(\hat{p})\Big\}.\ \square$$

Lemma 9 *For $\hat{p}\in\mathcal{P}$, the equality in (37) holds.*

Proof of Lemma 9 The proof follows from the same argument as in Part 1 of Proposition 5. We let α be small enough so that (22) and (23) hold for all realizations of $D_t(\hat{p}_t+\alpha\,e_l)$. Since the cumulative profit starting from time period $t+1$ depends on the prices at time periods $\{t+1,\dots,T\}$, we have

$$F_{t+1}^\pi\big(R_{t+1}^\pi(\hat{r}_t,\hat{p}_t+\alpha\,e_l,D_t(\hat{p}_t+\alpha\,e_l)),\hat{p}+\alpha\,\Delta_{lt},D(\hat{p}+\alpha\,\Delta_{lt})\big)$$
$$=F_{t+1}^\pi\big(R_{t+1}^\pi(\hat{r}_t,\hat{p}_t+\alpha\,e_l,D_t(\hat{p}_t+\alpha\,e_l)),\hat{p}+\alpha\,\Delta_{l,t+1},D(\hat{p}+\alpha\,\Delta_{lt})\big)$$
$$=F_{t+1}^\pi\big(R_{t+1}^\pi(\hat{r}_t,\hat{p}_t,D_t(\hat{p}_t+\alpha\,e_l)),\hat{p}+\alpha\,\Delta_{l,t+1},D(\hat{p}+\alpha\,\Delta_{lt})\big),\quad(39)$$

29

where we use (23) in the second equality. Using (7) to expand the terms in the conditional expectation in (37) and applying (39) in the resulting expression, we obtain

$$F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\,\Delta_{lt}, D(\hat{p} + \alpha\,\Delta_{lt})\right) - F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, D(\hat{p} + \alpha\,\Delta_{lt})\right)$$

$$= \rho_t(\hat{p}_t + \alpha\,e_l) \cdot X_t^\pi(\hat{r}_t, \hat{p}_t + \alpha\,e_l, D_t(\hat{p}_t + \alpha\,e_l)) - \rho_t(\hat{p}_t) \cdot X_t^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\,e_l))$$

$$= \left[\,\rho_t(\hat{p}_t + \alpha\,e_l) - \rho_t(\hat{p}_t)\,\right] \cdot X_t^\pi(\hat{r}_t, \hat{p}_t, D_t(\hat{p}_t + \alpha\,e_l)),$$

where we use (22) in the second equality. The result follows from the same argument as in Part 1 of Proposition 5. $\qquad\square$

Lemma 10 *For $\hat{p} \in \mathcal{P}$, the equality in (38) holds.*

Proof of Lemma 10 The prices $\hat{p} + \alpha\,\Delta_{lt}$ and $\hat{p} + \alpha\,\Delta_{l,t+1}$ coincide at time periods $\{1,\ldots,t-1,t+1,\ldots,T\}$. Therefore, the load random variables $D(\hat{p} + \alpha\,\Delta_{lt})$ and $D(\hat{p} + \alpha\,\Delta_{l,t+1})$ coincide at time periods $\{1,\ldots,t-1,t+1,\ldots,T\}$ with probability 1. Also, the load random variables over different lanes or at different time periods are independent. These observations and (14) imply that

$$\mathbb{P}\big\{D(\hat{p} + \alpha\,\Delta_{lt}) = D(\hat{p} + \alpha\,\Delta_{l,t+1})\,\big|\,D(\hat{p}) = \tilde{D}(\hat{p})\big\} = 1 - \lambda_{lt}(\hat{p}_{lt} + \alpha) + \lambda_{lt}(\hat{p}_{lt}) + o(\alpha) \qquad (40)$$

for $\alpha \le 0$. Similarly, the prices $\hat{p}$ and $\hat{p} + \alpha\,\Delta_{l,t+1}$ coincide at time periods $\{1,\ldots,t\}$. Therefore, the load random variables $D(\hat{p})$ and $D(\hat{p} + \alpha\,\Delta_{l,t+1})$ coincide at time periods $\{1,\ldots,t\}$ with probability 1. In this case, (14) and (15) imply that

$$\mathbb{P}\big\{D(\hat{p} + \alpha\,\Delta_{lt}) = D(\hat{p} + \alpha\,\Delta_{l,t+1}) + e_{lt},\ D(\hat{p} + \alpha\,\Delta_{l,t+1}) = D(\hat{p})\,\big|\,D(\hat{p}) = \tilde{D}(\hat{p})\big\}$$

$$= \big[\lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt})\big]\left(\prod_{s=t+1}^T \big[1 - \lambda_{ls}(\hat{p}_{ls} + \alpha) + \lambda_{ls}(\hat{p}_{ls})\big]\right) + o(\alpha) \quad (41)$$

for $\alpha \le 0$. The total probability of the events not covered by (40) or (41) is $o(\alpha)$, because

$$1 - \left\{1 - \lambda_{lt}(\hat{p}_{lt} + \alpha) + \lambda_{lt}(\hat{p}_{lt}) + \big[\lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt})\big]\left(\prod_{s=t+1}^T \big[1 - \lambda_{ls}(\hat{p}_{ls} + \alpha) + \lambda_{ls}(\hat{p}_{ls})\big]\right)\right\}$$

is $o(\alpha)$. Using (40) and (41), the conditional expectation in (38) can be written as

$$\mathbb{E}\Big\{F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, D(\hat{p} + \alpha\,\Delta_{lt})\right) - F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, D(\hat{p} + \alpha\,\Delta_{l,t+1})\right)\,\big|\,D(\hat{p}) = \tilde{D}(\hat{p})\Big\}$$

$$= \big[1 - \lambda_{lt}(\hat{p}_{lt} + \alpha) + \lambda_{lt}(\hat{p}_{lt})\big]$$

$$\times \mathbb{E}\Big\{F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, D(\hat{p} + \alpha\,\Delta_{lt})\right) - F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, D(\hat{p} + \alpha\,\Delta_{l,t+1})\right)$$

$$\big|\,D(\hat{p}) = \tilde{D}(\hat{p}),\ D(\hat{p} + \alpha\,\Delta_{lt}) = D(\hat{p} + \alpha\,\Delta_{l,t+1})\Big\}$$

$$+ \big[\lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt})\big]\left(\prod_{s=t+1}^T \big[1 - \lambda_{ls}(\hat{p}_{ls} + \alpha) + \lambda_{ls}(\hat{p}_{ls})\big]\right)$$

$$\times \Big\{F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, \tilde{D}(\hat{p}) + e_{lt}\right) - F_t^\pi\left(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, \tilde{D}(\hat{p})\right)\Big\} + o(\alpha).$$

The conditional expectation on the right side above is zero because of the condition $D(\hat{p} + \alpha\,\Delta_{lt}) = D(\hat{p} + \alpha\,\Delta_{l,t+1})$. Therefore, the limit in (38) is equal to

$$\lim_{\alpha\uparrow 0}\left\{\frac{1}{\alpha}\left[\lambda_{lt}(\hat{p}_{lt} + \alpha) - \lambda_{lt}(\hat{p}_{lt})\right]\left(\prod_{s=t+1}^{T}\left[1 - \lambda_{ls}(\hat{p}_{ls} + \alpha) + \lambda_{ls}(\hat{p}_{ls})\right]\right)\right\}$$

$$\times \lim_{\alpha\uparrow 0}\left\{F_t^{\pi}\big(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, \tilde{D}(\hat{p}) + e_{lt}\big) - F_t^{\pi}\big(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, \tilde{D}(\hat{p})\big)\right\}$$

$$= \dot{\lambda}_{lt}(\hat{p}_{lt})\lim_{\alpha\uparrow 0}\left\{F_t^{\pi}\big(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, \tilde{D}(\hat{p}) + e_{lt}\big) - F_t^{\pi}\big(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, \tilde{D}(\hat{p})\big)\right\}.$$

Since we have $\hat{p} \in \mathcal{P}$, the results of the decision and state transition functions do not change with infinitesimal perturbations in the prices by Lemma 3, which implies that

$$\lim_{\alpha\uparrow 0}\left\{F_t^{\pi}\big(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, \tilde{D}(\hat{p}) + e_{lt}\big) - F_t^{\pi}\big(\hat{r}_t, \hat{p} + \alpha\,\Delta_{l,t+1}, \tilde{D}(\hat{p})\big)\right\}$$

$$= F_t^{\pi}\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p}) + e_{lt}\big) - F_t^{\pi}\big(\hat{r}_t, \hat{p}, \tilde{D}(\hat{p})\big). \qquad \square$$

B   Obtaining a *Good* Fleet Management Policy for Given Prices

Throughout the paper, we use $\Pi(p)$ to denote the output of an algorithm that finds a *good* fleet management policy for given prices p. Godfrey & Powell (2002) give an iterative, sampling-based algorithm to obtain a set of value function approximations $\{\hat{V}_t^{\pi}(\cdot) : t \in \mathcal{T}\}$ that characterize a *good* fleet management policy π. In this section, we present a variant of their algorithm, which differs in the value function updating procedure employed in Step 4 below. This new updating procedure is due to Powell, Ruszczynski & Topaloglu (2004).

Step 1   Initialize the value function approximations: Set iteration counter $n = 1$. Pick an initial set of value function approximations $\{\hat{V}_t^1(\cdot) : t \in \mathcal{T}\}$ such that each $\hat{V}_t^1(\cdot)$ is of the form $\hat{V}_t^1(r_t) = \sum_{i\in\mathcal{I}}\hat{V}_{it}^1(r_{it})$, where each $\hat{V}_{it}^1(\cdot)$ is a piecewise-linear, concave function with points of nondifferentiability being a subset of positive integers. Letting K be the number of vehicles, the relevant domain of $\hat{V}_{it}^1(\cdot)$ is $\{0, 1, \ldots, K\}$. So, $\hat{V}_{it}^1(\cdot)$ can be characterized by a sequence of numbers $\{v_{it}^1(k) : k = 1, \ldots, K\}$, where $v_{it}^1(k) = \hat{V}_{it}^1(k) - \hat{V}_{it}^1(k-1)$. The intercept of $\hat{V}_{it}^1(\cdot)$ is not crucial since the value function approximations are embedded in optimization problems.

Step 2   Sample load realizations under given prices: Let $\tilde{D}^n = \{\tilde{D}_{lt}^n : l \in \mathcal{L},\ t \in \mathcal{T}\}$ be a sample of $D(p) = \{D_{lt}(p_{lt}) : l \in \mathcal{L},\ t \in \mathcal{T}\}$.

Step 3   Simulate the behavior of the current policy under load realizations $\tilde{D}^n$: Set $r_1^n = r_1$, $t = 1$.

Step 3.a   Let

$$\widetilde{V}_t^n(r_t^n, p_t, \tilde{D}_t^n) = \max_{(x_t, r_{t+1})\in\mathcal{Y}(r_t^n, \tilde{D}_t^n)} \rho_t(p_t)\cdot x_t + \hat{V}_{t+1}^n(r_{t+1})$$

$$\widetilde{V}_t^n(r_t^n \mp e_i', p_t, \tilde{D}_t^n) = \max_{(x_t, r_{t+1})\in\mathcal{Y}(r_t^n\mp e_i', \tilde{D}_t^n)} \rho_t(p_t)\cdot x_t + \hat{V}_{t+1}^n(r_{t+1}).$$

Let (x_t^n, r_{t+1}^n) be the optimal solution to the first problem above. Also, let

$$\vartheta_t^n(\mp e_i') = \widetilde{V}_t^n(r_t^n \mp e_i', p_t, \tilde{D}_t^n) - \widetilde{V}_t^n(r_t^n, p_t, \tilde{D}_t^n).$$

Step 3.b   Increment t by 1. If $t \leq T$, then go to Step 3.a.

Step 4   Update the value function approximations: For all $i \in \mathcal{I}$, $t \in \mathcal{T}$ and $k = 1, \ldots, K$, set

$$z_{it}^{n+1}(k) = \begin{cases} [1 - \gamma_{it}^n] v_{it}^n(k) + \gamma_{it}^n \left(-\vartheta_t^n(-e_i')\right) & \text{if } k = r_{it}^n \\ [1 - \gamma_{it}^n] v_{it}^n(k) + \gamma_{it}^n \vartheta_t^n(e_i') & \text{if } k = r_{it}^n + 1 \\ v_{it}^n(k) & \text{otherwise,} \end{cases}$$

where γ_{it}^n is a step size parameter. Letting v_{it}^{n+1} be the vector $\{v_{it}^{n+1}(k) : k = 1, \ldots, K\}$ for all $i \in \mathcal{I}$ and $t \in \mathcal{T}$, set

$$\begin{aligned} v_{it}^{n+1} = \operatorname*{argmin} \quad & \sum_{k=1}^{K} \left(w(k) - z_{it}^{n+1}(k)\right)^2 \\ \text{subject to} \quad & w(k-1) \geq w(k) \qquad\qquad\qquad k = 2, \ldots, K. \end{aligned} \qquad (42)$$

For all $i \in \mathcal{I}$ and $t \in \mathcal{T}$, let $\hat{V}_{it}^{n+1}(\cdot)$ be the piecewise-linear, concave function whose slope over $(k-1, k)$ is given by $v_{it}^{n+1}(k)$. (Concavity of $V_{it}^{n+1}(\cdot)$ is ensured by constraints (42).) For all $t \in \mathcal{T}$, let $\hat{V}_t^{n+1}(r_t) = \sum_{i \in \mathcal{I}} \hat{V}_{it}^{n+1}(r_{it})$.

Step 5   If $n = M$, where M is a predetermined iteration number, then stop and return the fleet management policy characterized by the value function approximations $\{\hat{V}_t^{n+1}(\cdot) : t \in \mathcal{T}\}$. Otherwise, increment n by 1 and go to Step 2.

References

Adelman, D. (2004), Price-directed control of a closed logistics queueing network, Technical report, The University of Chicago, Graduate School of Business.

Bertsekas, D. P. & Tsitsiklis, J. N. (1996), *Neuro-Dynamic Programming*, Athena Scientific, Belmont, MA.

Bertsimas, D. & de Boer, S. (2005), 'Simulation-based booking limits for airline revenue management', *Operations Research* **53**(1), 90–106.

Birge, J. R. & Louveaux, F. (1997), *Introduction to Stochastic Programming*, Springer-Verlag, New York.

Carvalho, T. A. & Powell, W. B. (2000), 'A multiplier adjustment method for dynamic resource allocation problems', *Transportation Science* **34**, 150–164.

Crainic, T., Gendreau, M. & Dejax, P. (1993), 'Dynamic and stochastic models for the allocation of empty containers', *Operations Research* **41**, 102–126.

Dantzig, G. & Fulkerson, D. (1954), 'Minimizing the number of tankers to meet a fixed schedule', *Naval Research Logistics Quarterly* **1**, 217–222.

Ermoliev, Y. (1988), Stochastic quasigradient methods, *in* Y. Ermoliev & R. Wets, eds, '*Numerical Techniques for Stochastic Optimization*', Springer-Verlag, Berlin.

Ferguson, A. & Dantzig, G. B. (1955), 'The problem of routing aircraft - A mathematical solution', *Aeronautical Engineering Review* **14**, 51–55.

Frantzeskakis, L. & Powell, W. B. (1990), 'A successive linear approximation procedure for stochastic dynamic vehicle allocation problems', *Transportation Science* **24**(1), 40–57.

Fu, M. (1994), 'Sample path derivatives for (s, S) inventory systems', *Operations Research* **42**(2), 351–363.

Glasserman, P. (1991), *Gradient Estimation via Perturbation Analysis*, Kluwer Academic Publishers, Norwell, Massachusetts.

Glasserman, P. & Tayur, S. (1995), 'Sensitivity analysis for base-stock levels in multiechelon production-inventory systems', *Management Science* **41**(2), 263–281.

Godfrey, G. A. & Powell, W. B. (2002), 'An adaptive, dynamic programming algorithm for stochastic resource allocation problems I: Single period travel times', *Transportation Science* **36**(1), 21–39.

Gorman, M. (2001), 'Intermodal pricing model creates a network pricing perspective at BNSF', *Interfaces* **31**(4), 37–49.

Higle, J. L. & Zhao, L. (2004), Adaptive and nonadaptive samples in solving stochastic linear programs: A computational investigation,, Technical report, The University of Arizona, Department of Systems and Industrial Engineering.

Jordan, W. & Turnquist, M. (1983), 'A stochastic dynamic network model for railroad car distribution', *Transportation Science* **17**, 123–145.

Karaesmen, I. & van Ryzin, G. (2004), 'Overbooking with substitutable inventory classes', *Operations Research* **52**(1), 83–104.

Kesten, H. (1958), 'Accelerated stochastic approximation', *The Annals of Mathematical Statistics* **29**(4), 41–59.

King, G. J. & Topaloglu, H. (2006), 'Incorporating the pricing decisions into dynamic fleet management models', *Journal of The Operational Research Society* (to appear).

Kleywegt, A. J., Nori, V. S. & Savelsbergh, M. W. P. (2002), 'The stochastic inventory routing problem with direct deliveries', *Transportation Science* **36**(1), 94–118.

Law, A. L. & Kelton, W. D. (2000), *Simulation Modeling and Analysis*, McGraw-Hill, Boston, MA.

Mahajan, S. & van Ryzin, G. (2001), 'Stocking retail assortments under dynamic customer substitution', *Operations Research* **49**(3), 334–351.

Nemhauser, G. & Wolsey, L. (1988), *Integer and Combinatorial Optimization*, John Wiley & Sons, Inc., Chichester.

Powell, W. B. (1985), 'Marginal cost pricing of truckload services: A comparison of two approaches', *Transportation Research B* **19B**(5), 433–445.

Powell, W. B. (1989), 'A review of sensitivity results for linear networks and a new approximation to reduce the effects of degeneracy', *Transportation Science* **23**(4), 231–243.

Powell, W. B., Ruszczynski, A. & Topaloglu, H. (2004), 'Learning algorithms for separable approximations of stochastic optimization problems', *Mathematics of Operations Research* **29**(4), 814–836.

Powell, W. B., Sheffi, Y., Nickerson, K., Butterbaugh, K. & Atherton, S. (1988), 'Maximizing profits for North American Van Lines' truckload division: A new framework for pricing and operations', *Interfaces* **18**, 21–41.

Puterman, M. L. (1994), *Markov Decision Processes*, John Wiley and Sons, Inc., New York.

Royden, H. L. (1988), *Real Analysis*, Prentice Hall, Englewood Cliffs, New Jersey.

Talluri, K. T. & van Ryzin, G. J. (2004), *The Theory and Practice of Revenue Management*, Kluver Academic Publishers.

Topaloglu, H. & Powell, W. B. (2004), 'Sensitivity analysis of a dynamic fleet management model using approximate dynamic programming', *Operations Research* (to appear). Available at `http://legacy.orie.cornell.edu/∼huseyin/publications/publications.html`.

Topaloglu, H. & Powell, W. B. (2006), 'Dynamic programming approximations for stochastic, time-staged integer multicommodity flow problems', *INFORMS Journal on Computing* **18**(1), 31–42.

van Ryzin, G. & Vulcano, G. (2006), 'Simulation-based optimization of virtual nesting controls for network revenue management', *Operations Research* (to appear).

White, W. (1972), 'Dynamic transshipment networks: An algorithm and its application to the distribution of empty containers', *Networks* **2**(3), 211–236.

White, W. & Bomberault, A. (1969), 'A network algorithm for empty freight car allocation', *IBM Systems Journal* **8**(2), 147–171.
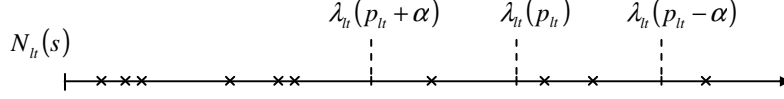
Figure 1: Construction of $\{D_{lt}(p_{lt}) : p_{lt} \in \mathbb{R}\}$ through $\{N_{lt}(s) : s \in \mathbb{R}_+\}$ and $\lambda_{lt}(\cdot)$. The sign $\times$ denotes an arrival of $\{N_{lt}(s) : s \in \mathbb{R}_+\}$. We have $D_{lt}(p_{lt}) = 7$ and $D_{lt}(p_{lt} + \alpha) = 6$ for the sample path above.

Step 1 Initialization: Set iteration counter $n = 1$. Pick initial prices $p^1 = \{p_{lt}^1 : l \in \mathcal{L}, \ t \in \mathcal{T}\}$.

Step 2 Sample the load realizations: Let $\tilde{D}^n = \{\tilde{D}_{lt}^n : l \in \mathcal{L}, \ t \in \mathcal{T}\}$ be a sample of $D(p^n) = \{D_{lt}(p_{lt}^n) : l \in \mathcal{L}, \ t \in \mathcal{T}\}$.

Step 3 Simulate the behavior of policy π under prices p^n and load realizations $\tilde{D}^n$: Set $r_1^n = r_1$, $t = 1$.

    Step 3.a Set $(x_t^n, r_{t+1}^n) = \underset{(x_t, r_{t+1}) \in \mathcal{Y}(r_t^n, \tilde{D}_t^n)}{\mathrm{argmax}} \ \rho_t(p_t^n) \cdot x_t + \hat{V}_{t+1}^\pi(r_{t+1})$.

    Step 3.b Increment t by 1. If $t \leq T$, then go to Step 3.a.

Step 4 For all $l \in \mathcal{L}, \ t \in \mathcal{T}$, approximate $F_t^\pi\big(r_t^n, p^n, \tilde{D}^n \mp e_{lt}\big) - F_t^\pi\big(r_t^n, p^n, \tilde{D}^n\big)$ by using one of the three methods described in Section 2.

Step 5 For all $l \in \mathcal{L}, \ t \in \mathcal{T}$, set

$$g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n) = \dot{\rho}_{lt}(p_{lt}^n) \, x_{lt}^n + \dot{\lambda}_{lt}(p_{lt}^n) \left\{ F_t^\pi\big(r_t^n, p^n, \tilde{D}^n + e_{lt}\big) - F_t^\pi\big(r_t^n, p^n, \tilde{D}^n\big) \right\}$$

$$g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n) = \dot{\rho}_{lt}(p_{lt}^n) \, x_{lt}^n + \dot{\lambda}_{lt}(p_{lt}^n) \frac{\tilde{D}_{lt}^n}{\lambda_{lt}(p_{lt}^n)} \left\{ F_t^\pi\big(r_t^n, p^n, \tilde{D}^n\big) - F_t^\pi\big(r_t^n, p^n, \tilde{D}^n - e_{lt}\big) \right\}.$$

Step 6 Update the prices: For all $l \in \mathcal{L}, \ t \in \mathcal{T}$, set

$$p_{lt}^{n+1} = \begin{cases} p_{lt}^n + \gamma_{lt}^n \, g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n) & \text{if } g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n) < 0 \text{ and} \\ & \qquad -g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n) > g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n) \\ p_{lt}^n + \gamma_{lt}^n \, g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n) & \text{if } g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n) > 0 \text{ and} \\ & \qquad g_{lt}^{\pi+}(r_t^n, p^n, \tilde{D}^n) \geq -g_{lt}^{\pi-}(r_t^n, p^n, \tilde{D}^n) \\ p_{lt}^n & \text{otherwise,} \end{cases}$$

    where γ_{lt}^n is a step size parameter.

Step 7 If the stopping criterion is not met, then increase n by 1 and go to Step 2.

Figure 2: Description of Algorithm 1.

Step 1 Initialization: Set policy improvement iteration counter $n = 1$. Pick initial prices p^1. Pick an initial *good* fleet management policy π^1 by setting $\pi^1 = \Pi(p^1)$.

Step 2 Find a *good* set of prices for the current fleet management policy by using Algorithm 1: Set $p^{n+1} = \Psi(\pi^n)$.

Step 3 Improve the current fleet management policy: Set $\pi^{n+1} = \Pi(p^{n+1})$.

Step 4 If the stopping criterion is not met, then increase n by 1 and go to Step 2.

Figure 3: Description of Algorithm 2.

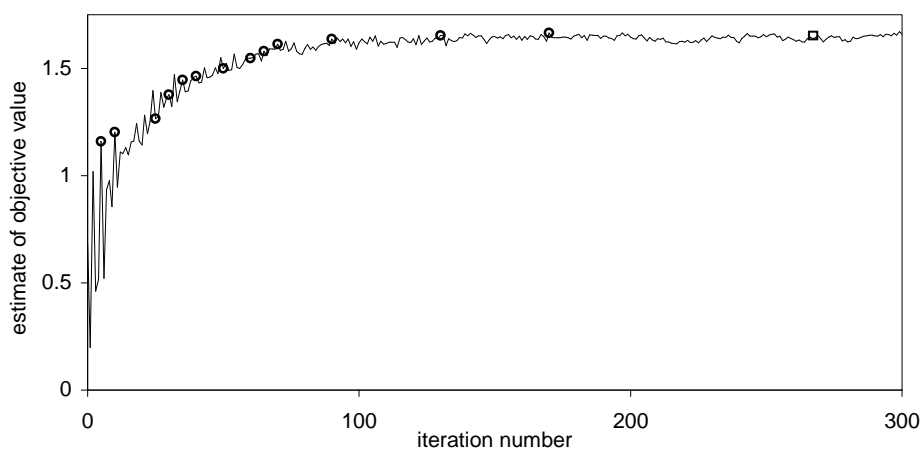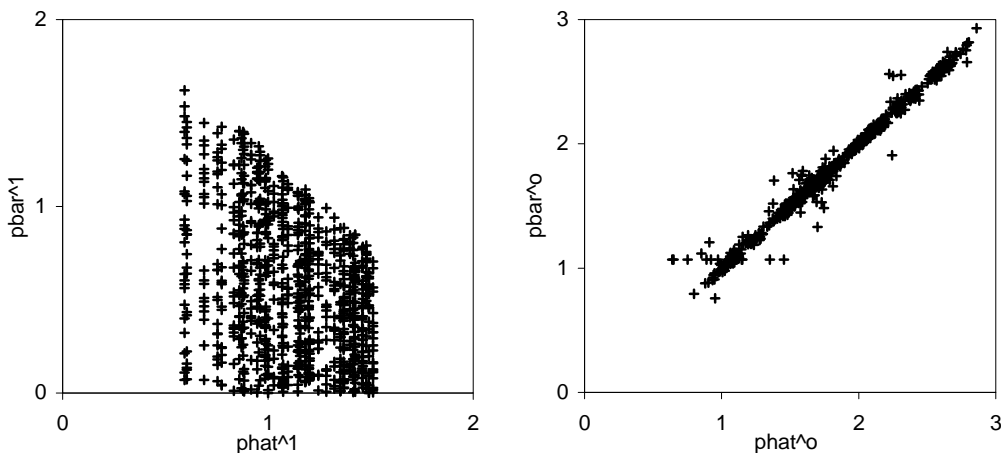Figure 4: Performance of the stopping criterion for test problem $(n, 20, 7, 50)$.

Figure 5: Comparison of $\{(\hat{p}_{lt}^1, \bar{p}_{lt}^1) : l \in \mathcal{L},\ t \in \mathcal{T}\}$ and $\{(\hat{p}_{lt}^o, \bar{p}_{lt}^o) : l \in \mathcal{L},\ t \in \mathcal{T}\}$ for test problem $(l, 20, 21, 250)$.

Problem	$\mathcal{F}(\pi^v, p^v)$	CPU^v	$\mathcal{F}(\pi^a, p^a)$	CPU^a
$(l, 40, 21, 50)$	6.94	32	6.94	94
$(l, 40, 21, 100)$	9.67	44	9.67	117
$(l, 40, 21, 250)$	11.86	52	11.87	142
$(n, 40, 21, 50)$	3.49	32	3.47	88
$(n, 40, 21, 100)$	4.77	39	4.75	106
$(n, 40, 21, 250)$	6.00	46	5.99	121

Table 1: Comparison of the value function approximation-based and perturbation analysis-based methods. We let $\mathcal{F}(\pi, p) = \mathbb{E}\{F_1^\pi(r_1, p, D(p))\}$ in the table heading. All profits are in $\$100,000$.

Problem	Mean	20 perc.	80 perc.
$(l, 20, 21, 50)$	1.84	1.36	2.26
$(l, 20, 21, 100)$	1.78	1.38	2.14
$(l, 20, 21, 250)$	1.72	1.38	2.05
$(n, 20, 21, 50)$	1.40	1.04	1.75
$(n, 20, 21, 100)$	1.37	1.01	1.70
$(n, 20, 21, 250)$	1.36	1.01	1.69
$(l, 20, 7, 50)$	1.85	1.42	2.23
$(l, 20, 7, 100)$	1.80	1.46	2.15
$(l, 20, 7, 250)$	1.75	1.43	2.06
$(n, 20, 7, 50)$	1.40	1.06	1.76
$(n, 20, 7, 100)$	1.37	1.01	1.71
$(n, 20, 7, 250)$	1.34	1.00	1.66

Problem	Mean	20 perc.	80 perc.
$(l, 40, 21, 50)$	2.03	1.51	2.43
$(l, 40, 21, 100)$	1.95	1.51	2.31
$(l, 40, 21, 250)$	1.90	1.51	2.22
$(n, 40, 21, 50)$	1.60	1.20	1.96
$(n, 40, 21, 100)$	1.59	1.23	1.91
$(n, 40, 21, 250)$	1.55	1.18	1.86
$(l, 40, 7, 50)$	2.01	1.51	2.40
$(l, 40, 7, 100)$	1.96	1.56	2.30
$(l, 40, 7, 250)$	1.90	1.52	2.23
$(n, 40, 7, 50)$	1.56	1.19	1.94
$(n, 40, 7, 100)$	1.52	1.13	1.85
$(n, 40, 7, 250)$	1.48	1.10	1.84

Table 2: Prices obtained by Algorithm 2 tend to decrease as the fleet size increases.

Problem	St. dv.	20 perc.	80 perc.
$(l, 20, 21, 50)$	0.12	1.93	2.09
$(l, 20, 21, 100)$	0.13	1.67	1.85
$(l, 20, 21, 250)$	0.16	1.22	1.39
$(n, 20, 21, 50)$	0.07	1.48	1.58
$(n, 20, 21, 100)$	0.06	1.38	1.47
$(n, 20, 21, 250)$	0.08	1.12	1.20
$(l, 20, 7, 50)$	0.13	1.89	2.07
$(l, 20, 7, 100)$	0.13	1.66	1.83
$(l, 20, 7, 250)$	0.14	1.22	1.38
$(n, 20, 7, 50)$	0.06	1.49	1.56
$(n, 20, 7, 100)$	0.06	1.38	1.45
$(n, 20, 7, 250)$	0.07	1.10	1.18

Problem	St. dv.	20 perc.	80 perc.
$(l, 40, 21, 50)$	0.16	1.80	2.03
$(l, 40, 21, 100)$	0.17	1.63	1.78
$(l, 40, 21, 250)$	0.27	1.26	1.48
$(n, 40, 21, 50)$	0.09	1.40	1.52
$(n, 40, 21, 100)$	0.10	1.32	1.41
$(n, 40, 21, 250)$	0.13	1.09	1.23
$(l, 40, 7, 50)$	0.18	1.78	2.01
$(l, 40, 7, 100)$	0.17	1.63	1.81
$(l, 40, 7, 250)$	0.24	1.26	1.44
$(n, 40, 7, 50)$	0.10	1.39	1.52
$(n, 40, 7, 100)$	0.09	1.30	1.41
$(n, 40, 7, 250)$	0.13	1.09	1.22

Table 3: When $\lambda_{lt}(\cdot) \approx \lambda_{l't'}(\cdot)$ and $\rho_{lt}(\cdot) \approx \rho_{l't'}(\cdot)$ for all $l, l' \in \mathcal{L}$, $t, t' \in \mathcal{T}$, the prices over different lanes and at different time periods are close to each other.

Problem	$\mathcal{F}(\tilde{\pi}^o, \tilde{p}^o)$	$\mathcal{F}(\hat{\pi}^o, \hat{p}^o)$	$\mathcal{F}(\bar{\pi}^o, \bar{p}^o)$	Problem	$\mathcal{F}(\tilde{\pi}^o, \tilde{p}^o)$	$\mathcal{F}(\hat{\pi}^o, \hat{p}^o)$	$\mathcal{F}(\bar{\pi}^o, \bar{p}^o)$
$(l, 20, 21, 50)$	9.55	9.53	9.51	$(l, 40, 21, 50)$	6.94	6.95	6.94
$(l, 20, 21, 100)$	12.87	12.84	12.83	$(l, 40, 21, 100)$	9.67	9.65	9.63
$(l, 20, 21, 250)$	15.07	14.99	15.02	$(l, 40, 21, 250)$	11.87	11.87	11.87
$(n, 20, 21, 50)$	4.74	4.73	4.74	$(n, 40, 21, 50)$	3.47	3.46	3.46
$(n, 20, 21, 100)$	5.77	5.70	5.72	$(n, 40, 21, 100)$	4.75	4.75	4.75
$(n, 20, 21, 250)$	6.75	6.75	6.78	$(n, 40, 21, 250)$	5.99	5.96	5.97
$(l, 20, 7, 50)$	3.26	3.24	3.23	$(l, 40, 7, 50)$	2.28	2.27	2.28
$(l, 20, 7, 100)$	4.43	4.42	4.41	$(l, 40, 7, 100)$	3.31	3.30	3.31
$(l, 20, 7, 250)$	5.69	5.69	5.66	$(l, 40, 7, 250)$	4.49	4.49	4.48
$(n, 20, 7, 50)$	1.68	1.67	1.68	$(n, 40, 7, 50)$	1.17	1.16	1.17
$(n, 20, 7, 100)$	2.17	2.16	2.16	$(n, 40, 7, 100)$	1.79	1.79	1.80
$(n, 20, 7, 250)$	2.88	2.88	2.89	$(n, 40, 7, 250)$	2.49	2.48	2.49

Table 4: Even though we start Algorithm 2 with different initial prices, the performances of the final prices are in close agreement.

Problem	$\mathcal{F}(\pi^o, p^o)$	$\mathcal{F}(\pi^b, p^b)$	No. itns.	CPU	% CPU on stopping
$(l, 20, 21, 50)$	9.55	9.18	260	11,332	20
$(l, 20, 21, 100)$	12.87	12.60	245	13,355	22
$(l, 20, 21, 250)$	15.07	12.43	245	15,089	23
$(n, 20, 21, 50)$	4.74	4.30	270	11,109	20
$(n, 20, 21, 100)$	5.77	4.81	265	13,474	21
$(n, 20, 21, 250)$	6.75	5.42	270	14,650	23
$(l, 20, 7, 50)$	3.26	3.12	270	3,557	20
$(l, 20, 7, 100)$	4.43	4.33	255	3,853	20
$(l, 20, 7, 250)$	5.69	5.51	255	3,959	22
$(n, 20, 7, 50)$	1.68	1.58	270	3,265	19
$(n, 20, 7, 100)$	2.17	1.91	265	3,469	20
$(n, 20, 7, 250)$	2.88	2.28	200	2,859	22
$(l, 40, 21, 50)$	6.94	6.58	235	25,126	14
$(l, 40, 21, 100)$	9.67	9.02	230	30,659	15
$(l, 40, 21, 250)$	11.87	11.03	425	70,975	15
$(n, 40, 21, 50)$	3.47	3.16	440	44,917	14
$(n, 40, 21, 100)$	4.75	3.85	425	53,891	16
$(n, 40, 21, 250)$	5.99	4.56	425	60,721	16
$(l, 40, 7, 50)$	2.28	2.16	200	6,387	13
$(l, 40, 7, 100)$	3.31	3.21	235	8,886	14
$(l, 40, 7, 250)$	4.49	4.37	245	10,655	15
$(n, 40, 7, 50)$	1.17	1.05	445	13,180	13
$(n, 40, 7, 100)$	1.79	1.64	415	14,314	13
$(n, 40, 7, 250)$	2.49	2.03	175	6,583	14

Table 5: Comparison of the prices obtained by Algorithm 2 and problem (34).

Problem	Profit ratio
$(l, 40, 21, 50)$	0.99
$(l, 40, 21, 100)$	1.00
$(l, 40, 21, 250)$	1.00
$(n, 40, 21, 50)$	1.00
$(n, 40, 21, 100)$	0.99
$(n, 40, 21, 250)$	1.00

Table 6: Comparison of the prices obtained by Algorithm 2 and the exhaustive numerical search.