

A DYNAMIC NEWSBOY MODEL FOR OPTIMAL RESERVE MANAGEMENT IN ELECTRICITY MARKETS

IN-KOO CHO AND SEAN P. MEYN

ABSTRACT. This paper examines a dynamic version of the newsboy problem in which a decision maker must maintain service capacity from several sources to meet demand for a perishable good, subject to the cost of providing sufficient capacity, and penalties for not meeting demand. The focus application is the real-time operation of an electric power network in which there are multiple sources of power that are distinguished by their cost, as well as their responsiveness in terms of ‘ramp rate’.

A complete characterization of the optimal outcome is obtained when normalized demand is modeled as Brownian motion. The optimal policy is *affine*: It is characterized by affine switching curves in the multidimensional state space. The optimal affine parameters are functions of variability in demand, production variables, and the cost of insufficient capacity.

KEYWORDS: inventory theory; newsboy model; optimal control; networks; electricity markets; reliability.

Acknowledgements Mike Chen allowed us to use the numerical results described in Section 2.4 which are taken from his thesis [12]. We are grateful to Hungpo Chao, Peter Cramton, Ramesh Johari, and Robert Wilson for helpful conversations.

Financial support from the National Science Foundation (SES-0004315, ECS-0217836, and ECS-0523620) is gratefully acknowledged. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

1. INTRODUCTION

Power networks throughout the world are managed through a sequence of scheduling decisions. In various regions of the United States this task is performed by the Independent Service Operator (ISO), which serves as an impartial planner to maintain the reliability of the system. In the first stage, generation units are committed to serve the predicted load for the next day. Next, to ensure that power is available despite unforeseen surges in demand, or breakdown of generation units, real-time scheduling decisions are made on time-scales ranging from a few minutes to one hour

This paper concerns the latter real-time operation of a power network. Although this amounts to a relatively small portion of the energy used in a typical region — in some cases less than 10% of the total energy — real-time scheduling is critical for the reliability of the grid. It is clear that the importance of real-time scheduling will be even greater with increased penetration of volatile resources such as wind, tidal, or solar power [20].

We argue that the scheduling decision problems arising in the electric power industry is the largest, and arguably the best known example of a *dynamic newsboy problem*. The

general newsboy problem concerns managing inventory for perishable goods, subject to cost for production and storage, and uncertainty in sales due to uncertainty in demand. There may also be penalties for not meeting demand. In the dynamic version of this problem in the case of electric power, the system operator is continuously facing the challenge of meeting rapidly changing demand through an array of generators that can ramp up rather slowly due to constraints on generation as well as the complex dynamics of the power grid [39, 19]. Uncertainty arises from unpredictable demand, volatile supply, and failure of generation resources.

Electricity cannot be stored economically in large quantities, yet the cost of not meeting demand is astronomical. The massive black-out seen in the Northeast on August 14, 2003, which cost \$4-6 billion dollars according to the US Department of Energy, reveals the tremendous cost of service disruption [37].

Although the focus application here is electric power, we note that the dynamic newsboy problem has application in many other settings. Examples are

Workforce management: In a large organization such as a hospital or a call center one must maintain a large workforce to ensure effective delivery of services (see [41] for a recent academic treatment, or IBM’s website [38].) To increase capacity of service one must bring new employees to work. Proper talent is usually identified and trained to be placed in position, which takes time. Alternatively, the organization can hire an employee through a temporary service agency which can offer talent at short notice, but at a higher price than the “usual” hiring process. In an example such as a hospital where the cost of not meeting demand has high social cost, it is crucial to secure a reliable channel of workers through these means, and through on-call staff.

Fashion manufacture and retail. In supplying a seasonal fashion product, the retailer maintains a small inventory which is available at short notice, while maintaining a contract with the supplier for deliverables in case of an unexpected surge in demand. It is more economical to have such a contract than to maintain inventory, but it takes some time to deliver the product from the producer to the retailer if demand increases unexpectedly [18].

We examine an idealized model in which a firm produces and delivers a good to the consumer in continuous time. The good is perishable, in that it must be consumed immediately. This is the case in electric power or temporary services, and to a lesser extent fashionable clothing. The firm has access to a number of sources of service that can produce the good. The social value is realized as the good is delivered and consumed. On the other hand, if some demand is not met, the social cost is proportional to the size of the excess demand.

It is assumed that the mean demand is met by *primary service* (the cheapest source of service capacity) through a prior contract. On-going demand can be met through primary service, as well as $K \geq 1$ sources of *ancillary service*. In the case of electric power, primary service will come from coal or nuclear power generators. Gas turbine generators are an example of more expensive, yet more responsive sources of power providing ancillary service.

Other constraints and costs assumed in the model are,

- **CONSTRAINED PRODUCTION AND FREE DISPOSAL:** Service capacity at time t from primary and ancillary services are denoted $\{G^p(t), G^{a_1}(t), \dots, G^{a_K}(t)\}$. Capacity is rate

constrained: $G^p(t)$ can increase at maximal rate ζ^{p+} , and $G^{a_k}(t)$ can increase at maximal rate $\zeta^{a_k+} < \infty$. Increasing capacity takes time, but the decision maker can freely (and instantaneously) dispose of excess capacity if desired.

- **CONSTANT MARGINAL CAPACITY COST:** c^p is the cost per unit capacity for primary service, and c^{a_k} the cost of building one additional unit of capacity from the k th source of ancillary service. The cost parameters are ordered, $c^p < c^{a_1} < \dots < c^{a_K}$.
- **CONSTANT MARGINAL VALUE OF CONSUMPTION:** v is the value per unit capacity for overall generation. At time t the instantaneous value of available power is $v \min(D(t), G^p(t) + G^a(t))$, where $D(t)$ denotes demand, and $G^a(t) := \sum_i G^{a_i}(t)$.
- **CONSTANT MARGINAL DIS-UTILITY OF SHORTAGE:** The excess capacity at time t (i.e., the *reserves*), is denoted

$$(1.1) \quad R(t) = G^p(t) + G^a(t) - D(t), \quad t \geq 0.$$

The marginal cost of excess demand is denoted $c^{bo} > 0$, so that the social cost of excess demand is given by $c^{bo} \max(-R(t), 0)$. The larger the shortage, the greater the damage to society. It is assumed that $c^{bo} > c^{a_K} > \dots > c^{a_1} > c^p$.

In practice, in particular in the case of power generation, supply is also subject to a finite downward ramping rate. However, the downward ramping rate is assumed to be considerably higher than the upward ramping rate — We regard the assumption of free disposal as an idealization of a fast downward ramping rate.

The basic control problem amounts to scheduling these resources to minimize average or discounted cost. When normalized demand is modeled as a driftless Brownian motion the system is viewed as a $(K + 1)$ -dimensional constrained diffusion model. The optimal policy is constructed explicitly through a detailed analysis of the dynamic programming equations for the multidimensional model, and is found to be of the precise affine form introduced in [14] (see also [46, p. 194].) The optimal solution clearly reveals how volatility of demand and the production technology of service influences the size, the composition, and the dynamics of optimal service capacity.

1.1. Background. At the close of the 1950s, Herbert Scarf obtained the optimal policy for a single period newsboy problem and showed that it is of a threshold form [57, 58], following previous research on inventory models by Arrow et. al. [2] and by Bellman ([8] and [7, Chapter 5].) Scarf points out in [57] that the conclusion that the solution is defined by a threshold follows from the convexity of the value function with respect to the decision variables. These structural issues are also developed in [7], and in dozens of papers published over the past fifty years (e.g. [8, 63, 31, 32, 1, 60, 61].)

Following these results there was an intense research program concerning the control of one-dimensional inventory models, e.g. [36, 68, 56, 62, 5, 21, 54]. More recently there have been efforts in various directions to develop hedging (or safety stocks) in multidimensional inventory models to improve performance [34, 27, 28, 60], make the system more responsive, [30]), or to obtain approximate optimality [42, 43, 33, 6, 48, 49, 50, 14, 64].

In the single-period newsboy problem, with a single source of service, the determination of the optimal service capacity based on forecast demand can be computed through a static calculus exercise. The resulting threshold is dependent upon cost, penalties, and the distribution of demand (see e.g. the early work of H. Scarf [57].) Wein in [65] considers

the special case in which $K = 1$, and uses a similar calculation to obtain a formula for optimal reserves in the diffusion model.

The results on optimal control obtained here are most closely related to results reported in [14] (and developed further in [51]). It is shown that for a large class of multiclass, multi-dimensional network models, an optimal policy can be approximated in “workload space” by a generalized threshold policy. It is called an *affine policy* since it is constructed as an affine translation of the optimal policy for a fluid model. In particular, [14, Theorem 4.4] establishes affine approximations under the discounted cost criterion, and [14, Proposition 4.5 and Theorem 4.7] establishes similar results under the average cost criterion for a diffusion model. The method of proof reduces the optimization problem to a static optimization calculation based on a one-dimensional reflected Brownian motion (see also the discussion surrounding the height process (3.35) below.) Consequently, the formula for the optimal affine parameter obtained in [14, Proposition 4.5 and Theorem 4.7] coincides with the formula presented in [65, Proposition 3], and is similar to the threshold values given in Section 2.2.

Some of the structural results reported in Section 2.2 are generalized in [13] to a more complex network setting based on an aggregate relaxation, similar to the workload relaxations employed in the analysis of queueing models [42, 48, 51]. The results of the present paper and [15] were presented in part in [16], and are surveyed in the SIAM news article [55].

In all of this prior work, in the case of models of dimension greater than one, only *approximations* of the optimal policy are obtained. The main contribution of this paper is to obtain a closed-form expression for the optimal policy, and show that it is exactly of the affine form that is used as an approximation in prior work.

The remainder of the paper is organized as follows. Section 2 describes the diffusion model in which normalized demand is a Brownian motion. The modeling assumptions and main results of this paper are contained in Section 2: Results for the case of a single source of ancillary service are contained in Section 2.2, and these results are generalized to the general model in Section 2.3. The proofs of the main results are contained in Section 3 along with some extensions. Section 4 concludes the paper.

2. DYNAMIC NEWSBOY MODEL AND MAIN RESULTS

This section summarizes the main results of this paper. A diffusion model is considered in the simplest case in which there is a single customer (also referred to as the *consumer*) that is served by primary and ancillary services. For the moment it is assumed that the consumer can access only a single source of ancillary service. It is assumed that the two sources of service are owned by the same firm, simply called the supplier.

The analysis is extended to multiple sources of service in Section 2.3.

2.1. Diffusion Model. Service capacity at time t from primary and ancillary services are denoted $\{G^p(t), G^a(t)\}$. Recall that, in application to electric power, we are considering real-time operations, so that the bulk of generation is scheduled in advance. The quantity $G^p(t)$ is the deviation in supply from this day-ahead scheduling, and hence it is not sign-constrained. The demand $D(t)$ is in fact the deviation in demand from forecast, so that

it also can take on positive or negative values. However, ancillary service is not scheduled in advance, so we impose the constraint that $G^a(t) \geq 0$ for all t .

Reserve at time t is defined by $R(t) = G^p(t) + G^a(t) - D(t)$ as expressed in (1.1). The event $R(t) < 0$ is interpreted as the failure of reliable services. In the application to electric power this represents *black-out* since the demand for power exceeds supply.

We sometimes refer to $G^p(t) + G^a(t)$ as the on-line capacity, since the supplier can offer primary and ancillary services $G^p(t)$ and $G^a(t)$ instantaneously at time t . Capacity is subject to ramping constraints: For finite, positive constants ζ^{p+}, ζ^{a+} ,

$$\frac{G^p(t') - G^p(t)}{t' - t} \leq \zeta^{p+} \quad \text{and} \quad \frac{G^a(t') - G^a(t)}{t' - t} \leq \zeta^{a+} \quad \text{for all } t' > t \geq 0.$$

We assume the free disposal of capacity, which means that $G^p(t)$ and $G^a(t)$ can decrease infinitely quickly. The ramping constraints can be equivalently expressed through the equations,

$$(2.2) \quad G^p(t) = G^p(0) - I^p(t) + \zeta^{p+}t, \quad G^a(t) = G^a(0) - I^a(t) + \zeta^{a+}t, \quad t \geq 0,$$

where the idleness processes $\{I^p, I^a\}$ are non-decreasing. It is assumed that $D(0)$ is given as an initial condition, and primary service is initialized using the definition (1.1),

$$G^p(0) = R(0) + D(0) - G^a(0).$$

Throughout most of the paper it is assumed that $D(0) = 0$.

It is assumed throughout the paper that \mathbf{D} is a driftless Brownian motion, with instantaneous variance denoted $\sigma_D^2 > 0$. The model (1.1) is then called the *controlled Brownian motion* (CBM) model, with two-dimensional state process $\mathbf{X} := (\mathbf{R}, \mathbf{G}^a)^T$. A Gaussian model for demand might be justified by considering a Central Limit Theorem scaling of a large number of individual demand processes. Rather than attempting to justify a limiting model, here we choose a Gaussian demand model for the purposes of control design.

Under this assumption, the state process \mathbf{X} evolves according to the Itô equation,

$$(2.3) \quad d\mathbf{X} = \delta_X - B d\mathbf{I}(t) - dD_X(t), \quad t \geq 0,$$

where $X(0) = (r, g^a)^T \in \mathbf{X} = \mathbb{R} \times \mathbb{R}_+$ is given, $\delta_X = (\zeta^{p+} + \zeta^{a+}, \zeta^{a+})^T$, $D_X(t) = (D(t), 0)^T$, and the 2×2 matrix B is defined by,

$$(2.4) \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

It is assumed that that the process $\mathbf{I} = (I^p, I^a)^T$ appearing in (2.2) and (2.3) is adapted to \mathbf{D} , and that the resulting state process \mathbf{X} is constrained to the state space $\mathbf{X} = \mathbb{R} \times \mathbb{R}_+$. A process \mathbf{I} satisfying these constraints is called admissible.

In what follows we restrict to *stationary Markov policies* defined as a family of admissible idleness processes $\{I^x\}$, parameterized by the initial condition $x \in \mathbf{X}$, with the defining property that the controlled process \mathbf{X} is a strong Markov process on \mathbf{X} .

One example of a stationary Markov policy is the affine policy that we define next. For $x \in \mathbb{R}$ we denote,

$$(2.5) \quad x_+ = \max(x, 0), \quad x_- = \max(-x, 0) = (-x)_+.$$

Definition 2.1. *An affine policy for the CBM model (2.3) is based on a pair of thresholds (\bar{r}^p, \bar{r}^a) , satisfying $\bar{r}^p > \bar{r}^a > 0$:*

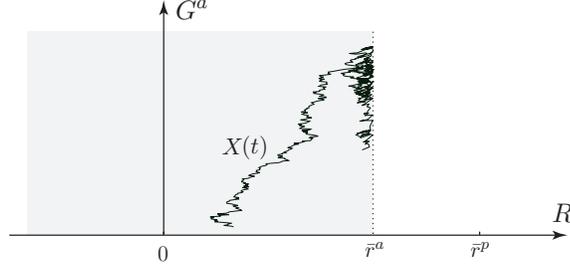


Figure 1: Trajectory of the two-dimensional model under an affine policy

(i) For the initial condition $X(0) = x = (r, g^a)^T \in \mathcal{X}$,

$$(2.6) \quad X(0+) = \begin{cases} \begin{pmatrix} r \\ g^a \end{pmatrix} - \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix}, & g^a \geq r - \bar{r}^p \\ \begin{pmatrix} \bar{r}^p \\ 0 \end{pmatrix}, & g^a \leq r - \bar{r}^p \\ \begin{pmatrix} r \\ g^a \end{pmatrix}, & \text{else,} \end{cases}$$

where $\beta := \min((r - \bar{r}^a)_+, g^a) \geq 0$.

(ii) For $t > 0$ the state process \mathbf{X} is restricted to the smaller state space given by $\mathcal{R}(\bar{r}) := \text{closure}(\mathcal{R}^p \cup \mathcal{R}^a)$, where

$$(2.7) \quad \mathcal{R}^a = \{x \in \mathcal{X} : x_1 < \bar{r}^a, x_2 \geq 0\}, \quad \mathcal{R}^p = \mathcal{R}^a \cup \{x \in \mathcal{X} : x_1 < \bar{r}^p, x_2 = 0\}.$$

(iii) For any $t > 0$, if $R(t) < \bar{r}^p$, then $\frac{d}{dt}G^p(t) = \zeta^{p+}$, and if $R(t) < \bar{r}^a$ then $\frac{d}{dt}G^a(t) = \zeta^{a+}$. Consequently, the following boundary constraints hold with probability one,

$$(2.8) \quad \int_0^\infty \mathbb{I}\{X(t) \in \mathcal{R}^p\} dI^p(t) = \int_0^\infty \mathbb{I}\{X(t) \in \mathcal{R}^a\} dI^a(t) = 0.$$

□

A sketch of a typical sample path of \mathbf{X} under an affine policy is shown in Figure 1. From the initial condition shown, the process has mean drift δ_X , up until the first time that $R(t)$ reaches the threshold \bar{r}^a . The subsequent downward motion shown is a consequence of reflection at the boundary of \mathcal{R}^a . Since $R(t)$ remains near \bar{r}^a yet primary service is ramped up at maximum rate ζ^{p+} , it follows that ancillary service capacity has long-run average drift of $-\zeta^{p+}$ up until the first time that $G^a(t)$ reaches zero. For the *fluid model* in which $\sigma_D^2 = 0$ we have $\frac{d}{dt}G^a(t) = \zeta^{a+}$ when $R(t) < \bar{r}^a$; $\frac{d}{dt}G^a(t) = -\zeta^{p+}$ whenever $G^a(t) > 0$ and $R(t) = \bar{r}^a$.

2.2. Optimization. Recall that $\{c^p, c^a\}$ denote the cost for maintaining one additional unit of capacity for primary and ancillary services. It is assumed that the marginal cost of production is higher for ancillary service,

$$c^p < c^a.$$

Welfare functions for the supplier and consumer are defined respectively by,

$$(2.9) \quad \begin{aligned} \mathcal{W}_S(t) &:= (p^p - c^p)G^p(t) + (p^a - c^a)G^a(t) \\ \mathcal{W}_D(t) &:= v \min(D(t), G^p(t) + G^a(t)) - (p^p G^p(t) + p^a G^a(t) + c^{bo} R_-(t)). \end{aligned}$$

Recall that $R_-(t) = \max(-R(t), 0)$ (see (2.5)).

The supplier is paid for the “on-line” capacity rather than the services delivered. For example, in an application to power the generator may have to burn coal in order to maintain a certain level of on-line capacity. On the other hand, the consumer obtains surplus only from the power delivered.

The welfare function for the consumer can be simplified,

$$(2.10) \quad \mathcal{W}_D(t) = vD(t) - (p^p G^p(t) + p^a G^a(t) + (c^{bo} + v)R_-(t)),$$

where we have used the identity,

$$\min(D(t), G^p(t) + G^a(t)) = \min(D(t), R(t) + D(t)) = D(t) - R_-(t).$$

The *social surplus* at time t is given by,

$$\mathcal{W}(t) = \mathcal{W}_S(t) + \mathcal{W}_D(t).$$

From the identity $G^p(t) = R(t) + D(t) - G^a(t)$, the social surplus is equivalently expressed,

$$\begin{aligned} \mathcal{W}(t) &= vD(t) - [c^p G^p(t) + c^a G^a(t) + (c^{bo} + v)R_-(t)] \\ &= (v - c^p)D(t) - [c^p R(t) + (c^a - c^p)G^a(t) + (c^{bo} + v)R_-(t)] \\ &= (v - c^p)D(t) - \mathcal{C}(t), \end{aligned}$$

where the *cost* defined by,

$$(2.11) \quad \mathcal{C}(t) := c(X(t)) := c^p R(t) + (c^a - c^p)G^a(t) + (c^{bo} + v)R_-(t).$$

The mean demand takes on a constant value $\mathbb{E}[D(t)] = \mathbb{E}[D(0)]$. Hence, for a given initial condition $D(0) = d$,

$$(2.12) \quad \mathbb{E}[\mathcal{W}(t)] = (v - c^p)d - \mathbb{E}[\mathcal{C}(t)], \quad t \geq 0.$$

In our optimization calculations we will consider the minimization of mean cost rather than maximization of mean welfare; this is justified by (2.12) since demand is assumed to be exogenous. We consider the two standard optimization criteria,

$$(2.13) \quad \text{Average cost:} \quad \eta := \limsup_{T \rightarrow \infty} \mathbb{E}_x \left[\frac{1}{T} \int_0^T c(X(t)) dt \right]$$

$$(2.14) \quad \text{Discounted cost:} \quad J_\gamma(x) := \mathbb{E}_x \left[\int_0^\infty e^{-\gamma t} c(X(t)) dt \right],$$

where $\gamma > 0$ is the discount parameter, and $x \in \mathbf{X}$ is the initial condition of \mathbf{X} . Our goal is to minimize the given criterion over all stationary policies.

The steady-state cost can be computed based on the following result of [13].

Theorem 2.2. *For any affine policy, the Markov process \mathbf{X} is exponentially ergodic [22]. The unique stationary distribution π on \mathbf{X} satisfies,*

(i) The first marginal of π is given by the distribution function,

$$\mathbf{P}_\pi\{R(t) \leq r\} = \begin{cases} e^{-\theta_p(\bar{r}^p - r)} & \bar{r}^a \leq r \leq \bar{r}^p \\ e^{-\theta_p(\bar{r}^p - \bar{r}^a) - \theta_a(\bar{r}^a - r)} & r \leq \bar{r}^a, \end{cases}$$

where

$$(2.15) \quad \theta_a = 2 \frac{\zeta^{p+} + \zeta^{a+}}{\sigma_D^2}, \quad \theta_p = 2 \frac{\zeta^{p+}}{\sigma_D^2}.$$

(ii) The steady state mean of the cost $c: \mathcal{X} \rightarrow \mathbb{R}_+$ defined in (2.11) is explicitly computable,

$$(2.16) \quad \eta(\bar{r}) := \pi(c) = \theta_a^{-1} \left(\frac{\zeta^{a+}}{\zeta^{p+}} c^a + e^{-\theta_a \bar{r}^a} (c^{bo} + v) \right) e^{-\theta_p(\bar{r}^p - \bar{r}^a)} + (\bar{r}^p - \theta_p^{-1}) c^p.$$

□

Consequently, the steady-state mean welfare functions of the supplier and consumer are computable when prices are fixed:

Corollary 2.3. *For any affine policy the steady-state mean social surplus is finite. When $D(0) = 0$, we have*

$$\lim_{t \rightarrow \infty} \mathbf{E}[\mathcal{W}(t)] = -\eta(\bar{r}),$$

where η is given in (2.16), and the convergence is exponentially fast. Moreover, if the prices (p^p, p^a) are fixed, then the individual welfare functions have finite steady-state means,

$$\lim_{t \rightarrow \infty} \mathbf{E}[\mathcal{W}_S(t)] = \theta_a^{-1} \frac{\zeta^{a+}}{\zeta^{p+}} (p^a - c^a) e^{-\theta_p(\bar{r}^p - \bar{r}^a)} + (\bar{r}^p - \theta_p^{-1}) (p^p - c^p)$$

$$\lim_{t \rightarrow \infty} \mathbf{E}[\mathcal{W}_D(t)] = - \left[\theta_a^{-1} \left(\frac{\zeta^{a+}}{\zeta^{p+}} p^a + e^{-\theta_a \bar{r}^a} (c^{bo} + v) \right) e^{-\theta_p(\bar{r}^p - \bar{r}^a)} + (\bar{r}^p - \theta_p^{-1}) p^p \right].$$

When $D(t)$ has zero-mean and the prices are fixed with $p^p \leq p^a \leq c^{bo} + v$, then necessarily $\mathbf{E}[\mathcal{W}_D(t)] < 0$: Exactly as in the derivation of (2.12) we have,

$$\mathbf{E}[\mathcal{W}_D(t)] = -\mathbf{E}[p^p R(t) + (p^a - p^p) G^a(t) + (c^{bo} + v) R_-(t)], \quad t \geq 0.$$

This is the inevitable ‘cost of variability’, i.e. risk, as seen by the consumer. Since we have assumed that demand is normalized, the consumer sees additional value arising from the contract purchase of mean-demand at time 0–.

An alternative is that the consumer forgoes the real-time market, trusting the ‘long term contract’ already secured to meet mean-demand, so that $\mathbf{G}^p = \mathbf{G}^a \equiv 0$. Based on (2.9) this leads to $\mathcal{W}_S(t) = 0$ and

$$(2.17) \quad \mathcal{W}_D(t) = -(v D_-(t) + c^{bo} D_+(t)).$$

If $D(0) = 0$ this then gives,

$$(2.18) \quad \mathbf{E}[\mathcal{W}_D(t)] = -\frac{1}{2}(c^{bo} + v) \mathbf{E}[|D(t)|] = -\frac{1}{2}(c^{bo} + v) \sqrt{t} \mathbf{E}[|D(1)|].$$

In conclusion, although the residual mean welfare seen by the consumer is always negative, there remains much benefit to engage the supplier for services if the prices are not too high. As clearly shown in (2.18), the alternative ‘open-loop’ strategy is not sustainable.

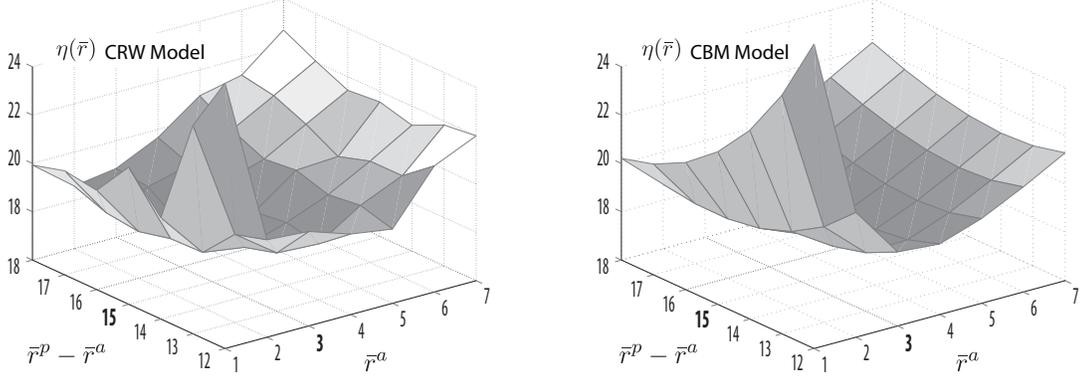


Figure 2: Shown at left is the average cost obtained using simulation for the network considered in Section 2.4 with Bernoulli demand using various affine parameters. At right is the average cost obtained from Theorem 2.2 (ii) for the CBM model.

The following two theorems describe the optimal policy for the social planner under the two cost criteria (2.13), (2.14). The proof of Theorem 2.4 is contained in Section 3.3. It is remarkable that the optimal policy is computable for this multi-dimensional model, and that the optimal solution is of this simple affine form.

Theorem 2.4. *The average-cost optimal policy (over all stationary policies) is affine, with specific parameter values given by,*

$$(2.19) \quad \bar{r}^{a*} = \theta_a^{-1} \ln \left(\frac{c^{bo} + v}{c^a} \right) \quad \bar{r}^{p*} = \bar{r}^{a*} + \theta_p^{-1} \ln \left(\frac{c^a}{c^p} \right).$$

□

From Theorem 2.2 (ii) it can be shown that the average cost is *convex* as a function of (\bar{r}^p, \bar{r}^a) , with unique minimum given in (2.19). This is plainly illustrated in the plot shown at right in Figure 2.

The proof of Theorem 2.5 is similar to the proof of Theorem 2.4. Computation of the parameters in (2.20) is provided in [51, Theorem 7.4.7] as part of the formulation of the “diffusion heuristic” (see also [12]).

Theorem 2.5. *The optimal policy under the discounted-cost control criterion is affine for a unique pair of parameters given by*

$$(2.20) \quad \bar{r}_\gamma^{a*} = \theta_a^{-1} \ln \left(\frac{c^{bo} + v}{c^a} \right) \quad \bar{r}_\gamma^{p*} = \bar{r}_\gamma^{a*} + \theta_p^{-1} \ln \left(\frac{c^a}{c^p} \right),$$

where the parameters θ_p, θ_a are each positive roots of a quadratic equation of the form

$$(2.21) \quad \frac{1}{2} \sigma_D^2 \theta^2 - \zeta^+ \theta - \gamma = 0.$$

In the case of θ_p we have $\zeta^+ = \zeta^{p+}$, and for θ_a we have $\zeta^+ = \zeta^{a+} + \zeta^{p+}$. Furthermore $\bar{r}_\gamma^{p*} \rightarrow \bar{r}^{p*}$ and $\bar{r}_\gamma^{a*} \rightarrow \bar{r}^{a*}$ as $\gamma \rightarrow 0$, where \bar{r}^{p*} and \bar{r}^{a*} are given in (2.19). □

2.3. Multiple Levels of Ancillary Service. The extension of the model to multiple sources of ancillary service contains no surprises: Suppose that there are K classes of ancillary service, with reserve capacities at time t denoted $\{G^{a_1}(t), \dots, G^{a_K}(t)\}$. The reserve remains defined as (1.1) with $G^a(t) := \sum_i G^{a_i}(t)$. The associated cost parameters and ramping rate constraints are denoted $\{c^{a_i}, \zeta^{a_i+} : 1 \leq i \leq K\}$ where

$$\frac{G^{a_i}(t') - G^{a_i}(t)}{t' - t} \leq \zeta^{a_i+} \quad \text{for all } t' > t.$$

It is assumed that the cost parameters are strictly increasing in the index i , with $c^{a_K} < c^{b_0}$.

The state process for control is $X(t) := (R(t), G^{a_1}(t), \dots, G^{a_K}(t))^T$, which is constrained to $\mathbf{X} := \mathbb{R} \times \mathbb{R}_+^K$. An affine policy is defined using the natural extension of the previous definition: For given parameters $\{\bar{r}^p > \bar{r}^{a_1} > \dots > \bar{r}^{a_K}\}$ we denote

$$\mathcal{R}^{a_i} := \{x = (r, g^{a_1}, \dots, g^{a_K}) \in \mathbb{R} \times \mathbb{R}_+^m : r < \bar{r}^{a_i}, g^{a_j} = 0 \text{ for } j > i\}.$$

The affine policy is defined so that $X(t) \in \text{closure}(\mathcal{R}^{a_i})$ whenever $G^{a_i}(t) > 0$ and moreover, for each i ,

$$\int_0^\infty \mathbb{I}\{X(t) \in \mathcal{R}^{a_i}\} dI^{a_i}(t) = 0.$$

That is, $G^{a_i}(t)$ ramps up at maximal rate when $X(t) \in \mathcal{R}^{a_i}$.

The cost function for the centralized planner is given by,

$$c(x) := c^p r + \sum_{i=1}^K (c^{a_i} - c^p) g^{a_i} + (c^{b_0} + v) r_-.$$

We present an extension of Theorem 2.4 for the model with K levels of ancillary service. It is found that the average cost optimal policy is again affine. The analogous result in the case of discounted cost is also valid.

In an optimal solution, capacity can sometimes be sought from a supplier with a high cost, especially if the supplier's capacity has a high ramping rate. It can be shown that the introduction of a new generator will strictly reduce the value of \bar{r}^{p*} obtained in (2.22) whenever $\zeta^{a_j+} > 0$.

Theorem 2.6. *The average-cost optimal policy for \mathbf{X} is affine, with specific parameter values given in the following modification of (2.19),*

$$\bar{r}^{a_i*} = \bar{r}^{a_{i+1}*} + \frac{1}{2} \frac{\sigma_D^2}{\zeta_i^+} \ln \left(\frac{c^{a_{i+1}}}{c^{a_i}} \right), \quad 1 \leq i \leq K, \quad (2.22)$$

$$\bar{r}^{p*} = \bar{r}^{a_1*} + \frac{1}{2} \frac{\sigma_D^2}{\zeta^{p+}} \ln \left(\frac{c^{a_1}}{c^p} \right),$$

where $\zeta_i^+ := \zeta^{p+} + \sum_{j \leq i} \zeta^{a_j+}$, and we denote $c^{K+1} := c^{b_0}$ and $\bar{r}^{a_{K+1}*} := 0$. \square

2.4. Numerical Examples. To conclude this section we present some numerical results based on simulation and dynamic programming experiments. These plots are taken from [12] where the reader can find further numerical results.

Simulation and optimization are performed for a two-dimensional *controlled random-walk* (CRW) model which evolves in discrete time. The forecasted excess capacity at time $t \geq 1$ is again defined by (1.1), where $D(t)$ is the demand at time t , and $(G^p(t), G^a(t))$

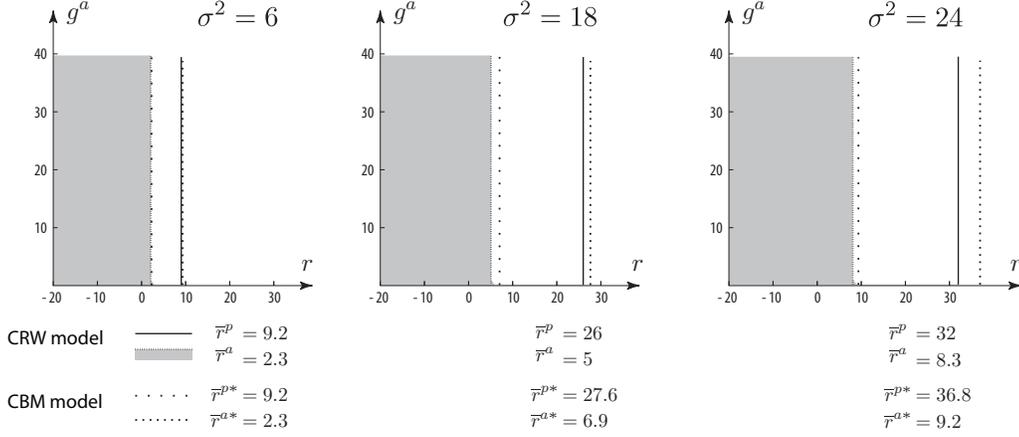


Figure 3: Optimal policies for the controlled random-walk model described in Section 2.4 obtained using value iteration. The grey region indicates those states for which ancillary service ramps up at maximum rate in the random-walk model, and the constant \bar{r}^p is the value such that primary ramps up at maximum rate when $R(t) < \bar{r}^p$. The optimal policy for the CBM model closely approximates the optimal policy for the discrete-time model in each case.

are current capacity levels from primary and ancillary service. It is assumed that \mathbf{D} is a random walk

$$D(t) = \sum_{s=1}^t \mathcal{E}(s), \quad t = 1, 2, \dots,$$

where the increment process \mathcal{E} is i.i.d., with zero mean and bounded support. The state process \mathbf{X} is constrained to the state space \mathbf{X} , and obeys the recursion,

$$(2.23) \quad X(t+1) = X(t) + BU(t) - \mathcal{E}_X(t+1), \quad t = 0, 1, \dots,$$

where the two dimensional process \mathcal{E}_X is defined as $\mathcal{E}_X(t) := (\mathcal{E}(t), 0)^T$. It is assumed that $U(t) \in \mathbf{U}(X(t))$ for all $t \in \mathbb{Z}_+$, where

$$\mathbf{U} := \{u = (u^p, u^a)^T \in \mathbb{R}^2 : -\infty \leq u^p \leq \zeta^{p+}, -\infty \leq u^a \leq \zeta^{a+}\},$$

$$\mathbf{U}(x) := \{u \in \mathbf{U} : x + Bu \in \mathbf{X}\}, \quad x \in \mathbf{X}.$$

An average-cost optimal policy is defined by a feedback law $f_*: \mathbf{X} \rightarrow \mathbf{U}$, that is characterized by the discrete-time dynamic programming equations. In analogy with the CBM model, we define the pair of regions,

$$(2.24) \quad \mathcal{R}^p = \{x \in \mathbf{X} : f_*(x) = u^{p+}\}, \quad \mathcal{R}^a = \{x \in \mathbf{X} : f_*(x) = u^{a+}\},$$

where u^{p+} and u^{a+} are defined in (3.27).

Consider a simple instance of the CRW model in which an optimal policy can be computed using value iteration. The marginal distribution of the increment process \mathcal{E} is symmetric, with support contained in the finite set $\{0, \pm 3, \pm 6\}$. The cost parameters are taken to be $c^{bo} = 100$, $c^a = 10$, $c^p = 1$, and we take $\zeta^{p+} = 1$, $\zeta^{a+} = 2$. The state process \mathbf{X} is restricted to an integer lattice to facilitate computation.

Three cases are considered to show how an optimal policy is influenced by variability: In each case the marginal distribution of \mathcal{E} is uniform on its respective support. The support and respective variance values are given by,

$$\begin{aligned} \text{(a)} \quad & \{-3, 0, 3\}, \sigma_a^2 = 6 & \text{(b)} \quad & \{-6, -3, 0, 3, 6\}, \sigma_b^2 = 18 \\ \text{(c)} \quad & \{-6, 0, 6\}, \sigma_c^2 = 24. \end{aligned}$$

The marginal distribution has zero mean since the support is symmetric in each case.

The average-cost optimal policy was computed for these three different models using value iteration. Results from these experiments are illustrated in Figure 3: The constant \bar{r}^p is defined as the maximum of $r \geq 0$ such that $U^p(t) = 1$ when $X(t) = (r, 0)^T$. The grey region represents the set \mathcal{R}^a defined in (2.24), and the constant \bar{r}^a is an approximation of the right-hand boundary of \mathcal{R}^a .

Also shown in Figure 3 is a representation of the optimal policy for the CBM model with first and second order statistics consistent with the CRW model. That is, the demand process \mathbf{D} was taken to be a drift-less Brownian motion with variance σ_D^2 equal to 6, 18, or 24 as shown in the figure. The constants \bar{r}^{p*} , \bar{r}^{a*} indicated in the figure are the optimal parameters for the CBM model given in (2.19). The optimal policy for the CBM model closely matches the optimal policy for the discrete-time model in each case.

We consider now a simulation experiment based on a family of affine policies for the CRW model.

The model parameters used in this simulation are as follows: The ‘value of consumption’ defined in the introduction is $v = 0$, $c^p = 1$, $c^a = 20$, and $c^{bo} = 400$. The ramp-up rates were taken as $\zeta^{p+} = 1/10$ and $\zeta^{a+} = 2/5$. The marginal distribution of the increment distribution was taken symmetric on $\{\pm 1\}$.

Shown at right in Figure 2 is the average cost obtained from Theorem 2.2 (ii) for the CBM model with first and second order statistics identical to those of the CRW model. For these numerical values, (2.19) gives $(\bar{r}^{p*}, \bar{r}^{a*}) = (17.974, 2.996)$.

Affine policies for the CRW model were constructed based on threshold values $\{\bar{r}^p, \bar{r}^a\}$. The average cost was approximated under several values of (\bar{r}^p, \bar{r}^a) based on the (unbiased) *smoothed estimator* of [35] (for details see [12].) In the simulation shown at left in Figure 2 the time-horizon was $n = 8 \times 10^5$. Among the affine parameters considered, the best policy for the discrete time model is given by $(\bar{r}^{p*}, \bar{r}^{a*}) = (19, 3)$, which almost coincides with the optimal values for the CBM model obtained using (2.19).

The thesis [12] contains similar simulations in which \mathcal{E} is Markov rather than i.i.d.. Similar solidarity is seen in these experiments when σ_D^2 is taken to be the asymptotic variance appearing in the Central Limit Theorem for \mathcal{E} .

In conclusion, in spite of the drastically different demand statistics, the best affine policy for the discrete-time model is remarkably similar to the average-cost optimal policy for the continuous-time model with Gaussian demand. Moreover, the optimal average cost for the two models are in close agreement.

We now consider in further detail the CBM model.

3. DIFFUSION MODEL

Here we present proofs of the main results as well as necessary background that may be of independent interest. Without loss of generality we take $D(0) = 0$ throughout.

3.1. Poisson's equation. It is convenient to introduce two “generators” for \mathbf{X} under a given Markov policy. The *extended generator*, denoted \mathcal{A} , is defined as follows: We write $\mathcal{A}f = g$ and say that f is in the domain of \mathcal{A} if the stochastic process \mathbf{M}_f defined below is a local martingale for each initial condition,

$$(3.25) \quad M_f(t) := f(X(t)) - f(X(0)) + \int_0^t g(X(s)) ds, \quad t \geq 0.$$

That is, there exists a sequence of stopping times $\{\tau^n\}$ satisfying $\tau^n \uparrow \infty$, and for each n the stochastic process $\{M_f^n(t) = M_f(t \wedge \tau^n) : t \geq 0\}$ satisfies the martingale property,

$$\mathbb{E}[M_f^n(t+s) | \mathcal{F}_t] = M_f^n(t), \quad t, s \geq 0,$$

where $\mathcal{F}_t = \sigma(X(s), D(s) : s \leq t)$. See [24, 22] for background.

The *differential generator* is defined on C^2 functions $f : \mathbf{X} \rightarrow \mathbb{R}$ via,

$$(3.26) \quad \mathcal{D}f := \langle \nabla f, Bu^{a+} \rangle + \frac{1}{2} \sigma_D^2 \frac{\partial^2}{\partial r^2} f,$$

with

$$(3.27) \quad u^{p+} = (\zeta^{p+}, 0)^T, \quad u^{a+} = (\zeta^{p+}, \zeta^{a+})^T.$$

Suppose that \mathbf{X} is controlled using an affine policy, and that the C^2 function f satisfies the boundary conditions,

$$(3.28) \quad \langle \nabla f(x), B1^1 \rangle = 0, \quad r = \bar{r}^p, \quad g^a = 0, \quad \langle \nabla f(x), B1^2 \rangle = 0, \quad r = \bar{r}^a, \quad g^a \geq 0.$$

where $\{1^i\}$ are the standard basis vectors. It then follows from Itô's formula that f is in the domain of \mathcal{A} with $\mathcal{A}f = \mathcal{D}f$.

In (3.28) and throughout the paper the vector 1^i denotes the i th basis vector in Euclidean space.

Suppose that \mathbf{X} is defined by a Markov policy with steady-state cost $\eta := \pi(c) < \infty$, where c is defined in (2.11). *Poisson's equation* is then defined to be the identity,

$$(3.29) \quad \mathcal{A}h = -c + \eta$$

The function $h : \mathbf{X} \rightarrow \mathbb{R}$ is known as the *relative value function*. If (3.29) holds then the stochastic process defined below is a local martingale for each initial condition,

$$(3.30) \quad M_h(t) = h(X(t)) - h(X(0)) + \int_0^t (c(X(s)) - \eta) ds, \quad t \geq 0.$$

The following result is an extension of results of [13], following [52, 29] and [53, Chapter 17].

Define for any policy the two stopping times,

$$(3.31) \quad \tau_p := \inf\{t \geq 0 : I^p(t) > 0\}, \quad \tau_a := \inf\{t \geq 0 : I^a(t) > 0\}.$$

For an affine policy with thresholds $\bar{r}^a < \bar{r}^p$ these stopping times have the equivalent representation,

$$\tau_p = \inf\{t \geq 0 : R(t) \geq \bar{r}^p\}, \quad \tau_a = \inf\{t \geq 0 : R(t) \geq \bar{r}^a\}.$$

In this case it is shown that the function $h : \mathbf{X} \rightarrow \mathbb{R}$ defined as

$$(3.32) \quad h(x) = \mathbb{E}_x \left[\int_0^{\tau_p} (c(X(t)) - \eta) dt \right], \quad x \in \mathbf{X},$$

solves Poisson's equation for \mathbf{X} , where $\eta(\bar{r})$ is defined in Theorem 2.2.

Proposition 3.1. *Suppose that \mathbf{X} is controlled using an affine policy. Then,*

- (i) *The following bound holds for each $m \geq 2$, some constant $b_m < \infty$, and any stopping time τ satisfying $\tau \leq \tau_p$,*

$$\mathbb{E}_x \left[\|X(\tau)\|^m + \int_0^\tau \|X(t)\|^{m-1} dt \right] \leq b_m (\|x\|^m + 1) \quad x \in \mathbf{X}.$$

- (ii) *One solution to Poisson's equation is given by (3.32). Moreover, for this solution the stochastic process \mathbf{M}_h is a martingale.*

- (iii) *The function h given in (3.32) satisfies for some $b_0 < \infty$,*

$$-b_0 \leq h(x) \leq b_0 (\|r - g^a\|^2 + 1), \quad x \in \mathbf{X}.$$

Proof. Part (i) is a minor extension of the proof of Proposition A.2 in [13]. Parts (ii) and (iii) are given in [13, Proposition A.2].

Since the proof is short we include a proof of (i): Consider for $m \geq 2$ the C^2 function $V_m(x) := m^{-1}|r - g^a - \bar{r}^p|^m$, $x = (r, g^a)^T \in \mathbf{X}$. Applying the differential generator (3.26) we obtain,

$$\mathcal{D}V_m(x) = -\zeta^{p+}|r - g^a - \bar{r}^p|^{m-1} + \sigma_D^2(m-1)|r - g^a - \bar{r}^p|^{m-2}, \quad x \in \mathcal{R}(\bar{r}),$$

where $\mathcal{R}(\bar{r})$ is defined above (2.7). This function also satisfies the boundary conditions given in (3.28) since $m \geq 2$, so that V_m is in the domain of \mathcal{A} and $\mathcal{A}V_m = \mathcal{D}V_m$ (this follows from Itô's formula — See [24, Theorem 2.9, p. 287], [23, 59] and their references). Consequently, one can find a compact set $S_m \subset \mathbf{X}$, $c_m < \infty$, and $\varepsilon_m > 0$ such that,

$$\mathcal{A}V_m \leq -\varepsilon_m V_{m-1} + c_m \mathbb{I}_{S_m}, \quad \text{on } \mathcal{R}(\bar{r}).$$

The bound in (i) then follows from standard arguments (see [13, Proposition A.2] and also [52]). \square

3.2. Height process. Proposition 3.1 asserts that the function h defined in (3.32) is a solution to Poisson's equation. To prove Theorem 2.4 we construct representations for the gradient of h in terms of the stopping times defined in (3.31), and the pair of sensitivity functions,

$$(3.33) \quad \begin{aligned} \lambda_p(x) &= \langle \nabla c(x), B1^1 \rangle = c^p - \mathbb{I}\{r \leq 0\}c^{b_0}, \\ \lambda_a(x) &= \langle \nabla c(x), B1^2 \rangle = c^a - \mathbb{I}\{r \leq 0\}c^{b_0}. \end{aligned}$$

The analysis is based upon a reduction to a pair of one-dimensional reflected processes: the *height processes* with respect to each of the thresholds \bar{r}^p, \bar{r}^a ,

$$(3.34) \quad H^p(t) := \bar{r}^p - R(t) + G^a(t),$$

$$(3.35) \quad H^a(t) := \bar{r}^a - R(t), \quad t \geq 0.$$

We have $H^p(t) \geq 0$ for all $t > 0$ under the affine policy, and moreover,

$$dH^p(t) = -\zeta^{p+}dt + dI^p(t) + dD(t), \quad t \geq 0.$$

Hence \mathbf{H}^p is a one-dimensional reflected Brownian motion (RBM), and \mathbf{H}^a evolves as an RBM up until the first time that $G^a(t) = 0$.

To analyze these height processes we first present some results for a standard RBM satisfying the the Itô equation,

$$(3.36) \quad dH = -\delta_H dt - dI(t) + dD(t), \quad t \geq 0,$$

where \mathbf{D} is a driftless Brownian motion, and the reflection process \mathbf{I} is non-decreasing and satisfies,

$$\int_0^\infty \mathbb{I}\{H(t) > 0\} dI(t) = 0.$$

When $\delta_H > 0$ the Markov process \mathbf{H} is positive recurrent, and its unique invariant probability measure is exponential with parameter $\theta_H := 2\delta_H/\sigma_H^2$ [11].

For a given constant $r_0 \geq 0$ define

$$\tau_{r_0} = \min\{t \geq 0 : H(t) = r_0\}$$

and consider the convex, C^1 function defined by,

$$(3.37) \quad \Psi(r) = \begin{cases} e^{\theta_H r} - \theta_H r & r < r_0 \\ e^{\theta_H r_0} - \theta_H r_0 + m_H(r - r_0) & r \geq r_0, \end{cases}$$

where $m_H = \Psi'(r_0) = \theta_H(e^{\theta_H r_0} - 1)$.

Proposition 3.2. *Suppose that the reflected Brownian motion has negative drift so that $\delta_H > 0$. Then for each initial condition $H(0) = r \geq 0$,*

(i) $\mathbb{E}[\tau_0] = \delta_H^{-1} r$.

(ii) For any $r_0 \geq r$,

$$\mathbb{P}\{\tau_{r_0} < \tau_0\} = (e^{\theta_H r_0} - 1)^{-1}(e^{\theta_H r} - 1).$$

(iii) For any constant $r_0 > 0$,

$$(3.38) \quad \mathbb{E}_x \left[\int_0^{\tau_0} \mathbb{I}\{H(t) \geq r_0\} dt \right] = \left(\Psi(r) - 1 + \theta_H r \right) \left(\delta_H \theta_H e^{\theta_H r_0} \right)^{-1}.$$

Proof. These formulae can be found in or derived from results in [11]. In particular (ii) is given as formula 3.0.4 (b), p. 309, and (iii) follows from formula 3.46 (a), p. 313 of [11]. We provide a brief proof based on invariance equations for the differential generator,

$$\mathcal{D}_H = -\delta_H \nabla + \frac{1}{2} \sigma_H^2 \nabla^2.$$

To prove (i) let $g(r) = \delta_H^{-1} r$ so that $\mathcal{D}_H g = -1$. It follows that $M_i(t) = t \wedge \tau_0 + g(H(t \wedge \tau_0))$ is a martingale,

$$\mathbb{E}[t \wedge \tau_0 + \delta_H^{-1} H(t \wedge \tau_0)] = \mathbb{E}[M_i(t)] = \mathbb{E}[M_i(0)] = \delta_H^{-1} H(0).$$

Letting $r \rightarrow \infty$ and applying the Dominated Convergence Theorem gives (i).

The function $g: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ defined by $g(r) := e^{\theta_H r}$, $r \in \mathbb{R}$, is in the null space of the differential generator for \mathbf{H} , which implies that $M_{ii}(t) = e^{\theta_H H(t \wedge \tau_0)}$ is a local martingale. Uniform integrability can be established to conclude that with $\tau = \min(\tau_{r_0}, \tau_0)$,

$$\mathbb{E}[M_{ii}(\tau)] = M_{ii}(0), \quad 0 \leq r \leq r_0.$$

Rearranging terms gives (ii).

To see (iii) we apply the differential generator to Ψ to obtain,

$$\mathcal{D}_H \Psi = -\delta_H (-\theta_H \mathbb{I}_{r < r_0} + m_H \mathbb{I}_{r \geq r_0}).$$

Here we have again used the fact that $\mathcal{D}_H g = 0$. Substituting the definition of m_H and writing $\mathbb{I}_{r < r_0} + \mathbb{I}_{r \geq r_0} = 1$ then gives,

$$\mathcal{D}_H \Psi = \delta_H (\theta_H - (m_H + \theta_H) \mathbb{I}_{r \geq r_0}).$$

The function Ψ satisfies $\Psi'(0) = 0$, so we can replace \mathcal{D}_H by \mathcal{A}_H in the expression above, which means that the process defined below is a local martingale,

$$M_{\text{iii}}(t) := \Psi(H(t)) - \delta_H \int_0^t (\theta_H - (m_H + \theta_H) \mathbb{I}_{H(s) \geq r_0}) ds, \quad t \geq 0.$$

It is in fact a martingale since it is uniformly integrable on any bounded interval. It is also uniformly integrable on $[0, \tau_0]$ which implies that,

$$\Psi(r) = M_{\text{iii}}(0) = \mathbb{E}[M_{\text{iii}}(\tau_0)] = \Psi(0) - \delta_H \mathbb{E} \left[\int_0^{\tau_0} (\theta_H - (m_H + \theta_H) \mathbb{I}\{H(s) \geq r_0\}) ds \right]$$

Rearranging terms gives for any $r_0 > 0$,

$$\mathbb{E} \left[\int_0^{\tau_0} \mathbb{I}\{H(s) \geq r_0\} ds \right] = \left(\delta_H (m_H + \theta_H) \right)^{-1} \left(\Psi(r) - \Psi(0) + \delta_H \theta_H \mathbb{E}[\tau_0] \right).$$

The expression (3.38) then follows. \square

Based on the foregoing we now show that the function h defined in (3.32) is smooth. We consider the normalized function $h_{\bullet}(x) = h(x) - h(x^p)$, $x \in \mathbf{X}$, with $x^p := (\bar{r}^p, 0)^T$. We obtain a useful representation for h_{\bullet} through a particular construction of the state processes starting from various initial conditions. Based on a single Brownian motion \mathbf{D} , we define on the same probability space the entire family of solutions to (2.3), denoted $\{X(t; x) : t \geq 0, x \in \mathbf{X}\}$.

The processes $X(t; x)$ and $X(t; x^p)$ have corresponding height processes $H^p(t; x)$, $H^p(t; x^p)$ satisfying $H^p(0; x^p) = 0$ and $H^p(0; x) \geq 0$. Consequently, the two processes couple at time $\tau_p(x) = \min\{t : H^p(t; x) = 0\} = \min\{t : H^p(t; x) = H^p(t; x^p)\}$. This combined with (3.32) implies the representation,

$$\begin{aligned} (3.39) \quad h_{\bullet}(x) &= \mathbb{E}_x \left[\int_0^{\tau_p(x)} (c(X(t; x)) - c(X(t; x^p))) dt \right] \\ &= \mathbb{E}_x \left[\int_0^{\infty} (c(X(t; x)) - c(X(t; x^p))) dt \right], \quad x \in \mathbf{X}. \end{aligned}$$

In Proposition 3.3, the function (3.39) is compared to the function obtained when \mathbf{X} is replaced by the fluid model in which $\sigma_D^2 = 0$. This deterministic process is denoted $\mathbf{x} = (r, g^a)^T$, and satisfies for $t > 0$,

$$\frac{d}{dt} \mathbf{x}(t; x) = \begin{cases} Bu^{a+} & r(t; x) < \bar{r}^a; \\ Bu^{a-} & r(t; x) = \bar{r}^a, g^a(t; x) > 0; \\ Bu^{p+} & \bar{r}^a \leq r(t; x) < \bar{r}^p, g^a(t; x) = 0; \\ 0 & r(t; x) = \bar{r}^p, \end{cases}$$

where u^{p+} and u^{a+} are defined in (3.27), and $u^{a-} = (\zeta^{p+}, -\zeta^{p+})^T$. The potential jump at time $t = 0$ is identical to its stochastic counterpart (see (2.6).)

Proposition 3.3. *Suppose that \mathbf{X} is controlled using an affine policy and that h_\bullet is defined by (3.39). Denote by h_0 the corresponding function when $\mathbf{D} = 0$,*

$$(3.40) \quad h_0(x) = \int_0^\infty (c(x(t; x)) - c(x^p)) dt \quad x \in \mathbf{X}.$$

Then,

(i) *The function h_0 is C^1 , and satisfies,*

$$\mathcal{D}h_0 = -c + c(x^p) + \frac{1}{2}\sigma_D^2 \frac{\partial^2}{\partial r^2} h_0.$$

(ii) *The function h_\bullet has the explicit form,*

$$h_\bullet = h_0 + \ell + m$$

where ℓ is a continuous piecewise-linear function of $(r, g^a)^T$, and m is a piecewise exponential function of r . The following identities hold whenever the derivatives are defined.

$$\mathcal{D}\ell = -\left(c(x^p) + \frac{1}{2}\sigma_D^2 \frac{\partial^2}{\partial r^2} h_0\right) + \eta, \quad \mathcal{D}m = 0.$$

(iii) *h_\bullet is C^1 and satisfies the boundary conditions,*

$$(3.41) \quad \begin{aligned} \frac{\partial}{\partial r} h_\bullet(x) + \frac{\partial}{\partial g^a} h_\bullet(x) &= 0, \quad \text{when } r \geq \bar{r}^a; \\ \frac{\partial}{\partial r} h_\bullet(x) &= 0, \quad \text{when } r + g^a \geq \bar{r}^p. \end{aligned}$$

Proof. The proof of Part (i) is similar to Proposition 4.2 of [50]. We first establish that h_0 is smooth. A representation of the derivative is obtained by differentiating the expression for h_0 :

$$\nabla h_0(x) = \int_0^\infty \nabla c(x(t; x)) dt \quad x \in \mathbf{X},$$

where the derivative is with respect to the initial condition x . Following arguments in [50], it can be shown that this expression is continuous and piecewise-linear on \mathbf{X} . The expression $\langle \nabla h_0(x), Bu^{a+} \rangle = -c(x) + c(x^p)$ follows from the Fundamental Theorem of Calculus and the representation,

$$h_0(x(s; x)) = \int_s^\infty (c(x(t; x)) - c(x^p)) dt \quad x \in \mathbf{X},$$

and this completes the proof of (i).

Parts (ii) and (iii) are proved in Appendix A of Chen's thesis [12]. The explicit formulae require complex computations, but the qualitative results are straightforward: We have from (i),

$$\mathcal{D}h_0 = -c + b_0,$$

where b_0 is piecewise constant. Following the proof of (i) we can give an explicit expression for ℓ ,

$$\ell(x) = \int_0^{\tau_p^0} (b_0(x(t; x)) - \eta) dt \quad x \in \mathbf{X},$$

where $\tau_p^0 = H^p(0; x)/\zeta^{p+} = (\bar{r}^p - r + g^a)/\zeta^{p+}$. This function is piecewise-linear and continuous, and satisfies,

$$\mathcal{D}\ell = -b_0 + \eta,$$

whenever ℓ is differentiable. Hence $h_0 + \ell$ satisfies Poisson's equation for the differential generator. The function m is in the null space of the differential generator, and is constructed so that $h_0 + \ell + m$ is C^1 . It is of the specific form,

$$m(r) = \begin{cases} A_p + B_p e^{-\theta_p r} & \bar{r}^a \leq r \leq \bar{r}^p; \\ A_a + B_a e^{-\theta_a r} & 0 \leq r < \bar{r}^a; \\ A_a + B_a & r < 0, \end{cases}$$

where the parameters $\{\theta_a, \theta_p\}$ are defined in (2.15), and (A_a, A_p, B_a, B_p) are constants. \square

We can now obtain representations of the derivatives:

Proposition 3.4. *Under an affine policy using any thresholds \bar{r}^p, \bar{r}^a , the directional derivatives of the function h defined in (3.32) can be expressed as follows for $x \in \mathcal{R}(\bar{r})$,*

$$(3.42) \quad \begin{aligned} \langle \nabla h(x), I^2 \rangle &= -\mathbf{E}_x \left[\int_0^{\tau_p} \lambda^p(X(t)) dt \right] \\ &= -c^p \mathbf{E}[\tau_p] + c^a \mathbf{E}_x \left[\int_0^{\tau_a} \mathbb{I}\{R(t) \leq \bar{r}^a\} dt \right] \end{aligned}$$

$$(3.43) \quad \begin{aligned} \langle \nabla h(x), B I^2 \rangle &= \mathbf{E}_x \left[\int_0^{\tau_a} \lambda^a(X(t)) dt \right] \\ &= c^a \mathbf{E}[\tau_a] - c^{b_0} \mathbf{E}_x \left[\int_0^{\tau_a} \mathbb{I}\{R(t) \leq 0\} dt \right]. \end{aligned}$$

Proof. We omit the proof of (3.42) since it is similar (and simpler) than the proof of (3.43).

Equation (3.43) is based on the following representation of h ,

$$(3.44) \quad h(x) = \mathbf{E}_x \left[\int_0^{\tau_a} (c(X(t)) - \eta) dt + h(X(\tau_a)) \right], \quad x \in \mathbf{X}.$$

This follows from the martingale property for \mathbf{M}_h and the bound Proposition 3.1 (i) with $m = 3$ (which implies that the martingale is uniformly integrable on $[0, \tau_a]$).

Suppose that $X(0) = x$ lies in the interior of $\mathcal{R}^a(\bar{r})$ and consider the perturbation $X^\varepsilon(0) = x^\varepsilon := x - \varepsilon B 1^2$ for small $\varepsilon > 0$. The state processes $\{\mathbf{X}^\varepsilon : \varepsilon \geq 0\}$ are defined on a common probability space, with common demand process \mathbf{D} . We then have,

$$X^\varepsilon(t) = X(t) - \varepsilon B 1^2, \quad 0 \leq t \leq \tau_a,$$

and from (3.44) it follows that for all $x \in \mathbf{X}$,

$$(3.45) \quad h(x) - h(x^\varepsilon) = \mathbf{E}_x \left[\int_0^{\tau_a} (c(X(t)) - c(X(t) - \varepsilon B 1^2)) dt + h(X(\tau_a)) - h(X(\tau_a) - \varepsilon B 1^2) \right].$$

Proposition 3.3 and the Mean Value Theorem give,

$$|h(X(\tau_a)) - h(X(\tau_a) - \varepsilon B 1^2)| = \varepsilon |\langle \nabla h(\tilde{X}), B 1^2 \rangle|,$$

where $\tilde{X} = X(\tau_a) - \tilde{\varepsilon} B 1^2$ for some $\tilde{\varepsilon} \in (0, \varepsilon)$. We have $R(\tau_a) = \bar{r}^a$, which implies that $|\langle \nabla h(\tilde{X}), B 1^2 \rangle| \rightarrow 0$ with probability one as $\varepsilon \rightarrow 0$.

An application of Proposition 3.3 (iii) gives the following bound,

$$\sup_{0 \leq \varepsilon \leq 1} \|\nabla h(y - \varepsilon B1^2)\| \leq b_0 \|y\|, \quad y = (\bar{r}^a, g^a), \quad g^a \geq 0,$$

where the constant b_0 is independent of g^a . Proposition 3.1 (i) with $m = 3$ implies uniform integrability, so that

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E}[\langle \nabla h(\tilde{X}), B1^2 \rangle] = 0.$$

Multiplying the identity (3.45) by ε^{-1} and applying Proposition 3.1 (i) once more gives,

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (h(x) - h(x^\varepsilon)) = \mathbf{E}_x \left[\int_0^{\tau_a} \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (c(X(t)) - c(X(t) - \varepsilon B1^2)) dt \right], \quad x \in \mathcal{R}^a(\bar{r}),$$

which gives (3.43). \square

3.3. Optimization. In this section we apply Proposition 3.1 to show that the affine policy described in Theorem 2.4 is average-cost optimal. The treatment of the discounted case is identical - we omit the details.

The dynamic programming equations for the CBM model are written as follows,

Average cost:

$$(3.46) \quad (\mathcal{D}h_* + c - \eta_*) \wedge \left(\inf \{ \langle \nabla h_*(x), -Bu \rangle : u \in \mathbb{R}_+^2 \} \right) = 0$$

Discounted cost:

$$(3.47) \quad (\mathcal{D}K_* + c - \gamma K_*) \wedge \left(\inf \{ \langle \nabla K_*(x), -Bu \rangle : u \in \mathbb{R}_+^2 \} \right) = 0$$

where the differential generator \mathcal{D} is defined in (3.26) (see [3, 4, 10]). The function $h_*: \mathbf{X} \rightarrow \mathbb{R}_+$ in (3.46) is known as the relative value function, and η_* is the optimal average cost. For the models considered here, however, we do not know if these value functions are C^2 on all of \mathbf{X} . Hence the dynamic programming equation (3.46) or (3.47) is interpreted in the viscosity sense [26, 17] (alternatively, one can replace \mathcal{D} by \mathcal{A} in these definitions.)

The relative value function defines a constraint region for \mathbf{X} as follows: Define in analogy with (2.7),

$$\mathcal{R}^p = \{x \in \mathbf{X} : \langle \nabla h_*(x), B1^1 \rangle < 0\}, \quad \mathcal{R}^a = \{x \in \mathbf{X} : \langle \nabla h_*(x), B1^2 \rangle < 0\}.$$

Then, with $\mathcal{R}^* := \text{closure} \{ \mathcal{R}^a \cup \mathcal{R}^p \}$, the optimal policy maintains for each initial condition,

$$(3.48) \quad \text{(i) } X(t) \in \mathcal{R}^* \text{ for all } t > 0, \quad \text{(ii) With probability one (2.8) holds.}$$

A representation for h_* can be obtained through a generalization of the ‘‘stochastic shortest path’’ formulation of [25] to the continuous time case (see also [66, 67, 9, 45], and in particular [47, Theorem 1.7] where the extensions to continuous time are spelled-out.) Consider for any $x \in \mathbf{X}$,

$$h_\circ(x) := \inf \mathbf{E}_x \left[\int_0^{\tau_p} (c(X(t)) - \eta_*) dt \right],$$

where the stopping time τ_p is defined for a general policy in (3.31), and the infimum is over all admissible \mathbf{I} . Under the optimal policy, the value function h_\circ solves the same

martingale problem as h_* , that is $\mathcal{A}h_* = -c + \eta_*$. It is the unique solution to (3.46) (up to an additive constant) over all functions with quadratic growth.¹

Conversely, if a solution to (3.46) can be found with quadratic growth then this defines an optimal policy:

Proposition 3.5. *Suppose that (3.46) holds for a function h_* satisfying for some $b_0 < \infty$,*

$$-b_0 \leq h_*(x) \leq b_0(1 + \|x\|^2), \quad x \in \mathbf{X}.$$

Then for any Markov policy that gives rise to a positive recurrent process \mathbf{X} with invariant probability measure π we have $\int c(x)\pi(dx) \geq \eta_$. Moreover, this lower bound is attained for the process defined in \mathcal{R}^* satisfying (3.48).*

Proof. Proposition 3.5 is a minor extension of [45, Theorem 5.2]. We sketch the proof here.

The essence of the dynamic programming equation (3.46) is that the process defined by

$$M_{h_*}(t) = h_*(X(t)) - h_*(X(0)) + \int_0^t (c(X(s)) - \eta_*) ds, \quad t \geq 0,$$

is a local submartingale for any solution \mathbf{X} obtained using an admissible idleness process \mathbf{I} : There exists a sequence of stopping times $\{\tau^n\}$ satisfying $\tau^n \uparrow \infty$, and the stochastic process defined by $M_{h_*}^n(t) = M_{h_*}(t \wedge \tau^n)$ satisfies the sub-martingale property,

$$\mathbb{E}[M_{h_*}^n(t+s) \mid \mathcal{F}_t] \geq M_{h_*}^n(t), \quad t, s \geq 0.$$

We can in fact take $\tau_n = \min\{t \geq 0 : h_*(X(t)) \geq n\}$. From the Monotone Convergence Theorem we obtain the bound,

$$(3.49) \quad \mathbb{E}_x \left[h_*(X(t)) + \int_0^t c(X(s)) ds \right] \geq t\eta_* + h_*(x), \quad t \geq 0, x \in \mathbf{X}.$$

That is, the modifier ‘local’ can be removed: M_{h_*} is a sub-martingale.

Arguments used in [45, Theorem 5.2] imply that the following limit holds for a.e. $x \in \mathbf{X}$ whenever $\pi(c) < \infty$,

$$\lim_{t \rightarrow \infty} t^{-1} \mathbb{E}_x[h_*(X(t))] = \lim_{t \rightarrow \infty} t^{-1} \mathbb{E}_x[\|X(t)\|^2] = 0.$$

Consequently, letting $t \rightarrow \infty$ in (3.49) gives, for a.e. $X(0) = x \in \mathbf{X}$,

$$\eta = \lim_{t \rightarrow \infty} t^{-1} \mathbb{E}_x \left[\int_0^t c(X(s)) ds \right] \geq \eta_*.$$

Moreover, if \mathbf{X} is defined under the optimal policy then M_{h_*} is a local martingale since Poisson’s equation holds,

$$(3.50) \quad \mathcal{A}h_* = -c + \eta_*.$$

If h_* has quadratic growth then M_{h_*} is a martingale (see prior footnote), and hence (3.49) can be strengthened to an equality,

$$\mathbb{E}_x \left[h_*(X(t)) + \int_0^t c(X(s)) ds \right] = t\eta_* + h_*(x), \quad t \geq 0, x \in \mathbf{X}.$$

¹Uniqueness is established in [45, Theorem A3]. Although stated in discrete time, Section 6 of [45] describes how to translate to continuous time. Related results are obtained for constrained diffusions in [3].

This shows that $\int c(x)\pi(dx) = \eta_*$ under the policy defined in (3.48). \square

We can now state the main result of this section. Recall that the thresholds \bar{r}^{p*} and \bar{r}^{a*} are defined in (2.19).

Proposition 3.6. *The following hold for the CBM model under an affine policy:*

- (i) *Suppose that primary service is specified using the threshold $\bar{r}^p > \bar{r}^{a*}$. If $\bar{r}^a = \bar{r}^{a*}$ then the solution to Poisson's equation (3.32) satisfies,*

$$\langle \nabla h(x), B1^2 \rangle < 0, \quad x \in \mathcal{R}^a.$$

- (ii) *If $\bar{r}^p = \bar{r}^{p*}$ and $\bar{r}^a = \bar{r}^{a*}$ then h satisfies in addition,*

$$\langle \nabla h(x), B1^1 \rangle < 0, \quad x \in \mathcal{R}^p.$$

Consequently, h solves the dynamic programming equation (3.46).

Corollary 3.7. *For any given $\bar{r}^p > \bar{r}^{a*}$, the optimal policy over all G^a is the affine policy obtained using the same \bar{r}^{a*} .*

Proof of Proposition 3.6. Recall that under an affine policy, \mathbf{H}^p is a one-dimensional RBM. For an initial condition $x \in \mathcal{R}(\bar{r})$ satisfying $g^a = G^a(0) > 0$, the height process \mathbf{H}^p evolves as a one-dimensional RBM up to the first time $t > 0$ that $G^a(t) = 0$.

Part (i). We begin by considering the right hand side of (3.43). We show that this is strictly negative on $\mathcal{R}^a(\bar{r})$ when $\bar{r}^a = \bar{r}^{a*}$ through an analysis of the height process \mathbf{H}^a .

The gradient formula (3.43) can be expressed in terms of the height process (3.35) via,

$$(3.51) \quad \langle \nabla h(x), B1^2 \rangle = c^a \mathbf{E}[\tau_a] - c^{bo} \mathbf{E}_x \left[\int_0^{\tau_a} \mathbb{I}\{H^a(t) \geq \bar{r}^a\} dt \right].$$

The stopping time τ_a can be interpreted as the first hitting time to the origin for \mathbf{H}^a . The following identities are obtained in Proposition 3.2:

$$\mathbf{E}[\tau_a] = \delta_H^{-1} r, \quad \mathbf{E}_x \left[\int_0^{\tau_a} \mathbb{I}\{H^a(t) \geq \bar{r}^a\} dt \right] = \left(\Psi(r) - 1 + \theta_H r \right) \left(\delta_H \theta_H e^{\theta_H \bar{r}^a} \right)^{-1},$$

where Ψ is defined in (3.37) by

$$\Psi(r) = \begin{cases} e^{\theta_H r} - \theta_H r & r < r_0 \\ e^{\theta_H r_0} - \theta_H r_0 + m_H(r - r_0) & r \geq r_0, \end{cases}$$

with $r_0 = \bar{r}^a$, $r = H^a(0) = \bar{r}^a - r \geq 0$, and $\delta_H = \zeta^{p+} + \zeta^{a+}$. Consequently, (3.51) can be expressed,

$$\langle \nabla h(x), B1^2 \rangle = \Phi(r) := c^a \delta_H^{-1} r - c^{bo} \left(\Psi(r) - 1 + \theta_H r \right) \left(\delta_H \theta_H e^{\theta_H \bar{r}^a} \right)^{-1}, \quad r \geq 0.$$

The function Ψ is convex, with $\Psi(0) = 1$ and $\Psi'(0) = 0$. Consequently, the function Φ defined above is *concave*, strictly concave on $[0, \bar{r}^a]$, with $\Phi(0) = 0$. To show that Φ is negative on $(0, \infty)$ it suffices to show that $\Phi'(0) \leq 0$.

The derivative at zero is expressed,

$$\Phi'(0) = c^a \delta_H^{-1} - c^{bo} \left(\Psi'(0) + \theta_H \right) \left(\delta_H \theta_H e^{\theta_H \bar{r}^a} \right)^{-1} = \delta_H^{-1} \left(c^a - c^{bo} e^{-\theta_H \bar{r}^a} \right).$$

When $\bar{r}^a = \bar{r}^{a*}$ we have $e^{-\theta_H \bar{r}^a} = c^a / c^{bo}$, so that $\Phi'(0) = 0$.

Part (ii). We next consider $\langle \nabla h, B1^1 \rangle$ for $x \in \mathcal{R}^a$ when \bar{r}^{p*} and \bar{r}^{a*} are given by (2.19).

Consider the height process relative to \bar{r}^p defined in (3.34). By the foregoing analysis we have on \mathcal{R}^a ,

$$\langle \nabla h, 1^1 + 1^2 \rangle = \langle \nabla h, B1^2 \rangle < 0.$$

Consequently, to show that $\langle \nabla h, B1^1 \rangle < 0$ it is sufficient to show that $\langle \nabla h, 1^2 \rangle \geq 0$. This derivative is given in (3.42), which can be expressed in terms of the height process,

$$\langle \nabla h, 1^2 \rangle = c^a \mathbf{E} \left[\int_0^{\tau_p} \mathbb{I}\{H^p(t) \geq r_0\} dt \right] - c^p \mathbf{E}[\tau_p],$$

where $r_0 = \bar{r}^p - \bar{r}^a$. Proposition 3.2 then gives, with $r = H^p(0)$,

$$\langle \nabla h, 1^2 \rangle = c^a \left(\Psi(r) - 1 + \theta_H r \right) \left(\delta_H \theta_H e^{\theta_H r_0} \right)^{-1} - c^p \delta_H^{-1} r.$$

The drift parameter for \mathbf{H}^p is $\delta_H = \zeta^{p+}$. The parameter \bar{r}^{p*} is chosen so that $e^{\theta_H r_0} = c^a / c^p$, which on substitution gives,

$$(3.52) \quad \langle \nabla h, 1^2 \rangle = c^p \delta_H^{-1} \left(\theta_H^{-1} (\Psi(r) - 1 + \theta_H r) - r \right) = c^p \delta_H^{-1} \theta_H^{-1} (\Psi(r) - 1).$$

For $r \geq r_0 > 0$ (equivalently, $r < \bar{r}^a$), we obtain from the definition of Ψ ,

$$\langle \nabla h, 1^2 \rangle \geq c^p \delta_H^{-1} \theta_H^{-1} (e^{\theta_H r_0} - \theta_H r_0 - 1) > 0.$$

We conclude that $\langle \nabla h, 1^1 \rangle < 0$ on \mathcal{R}^a .

Finally we demonstrate that $\langle \nabla h, B1^1 \rangle < 0$ on $\mathcal{R}^p \setminus \mathcal{R}^a$. For an initial condition $x \in \mathcal{R}^p$ satisfying $\bar{r}^a \leq r < \bar{r}^p$ we can write,

$$h(x) = \mathbf{E}_x \left[\int_0^\tau (c(X(t)) - \eta) dt + h(X(\tau)) \right]$$

where $\tau = \min(\tau_a, \tau_p)$. Recall that $\langle \nabla h(x), 1^1 \rangle = 0$ if $\tau = \tau_p$. Consequently, using familiar arguments,

$$\begin{aligned} \langle \nabla h(x), 1^1 \rangle &= \mathbf{E} [c^p \tau + \langle \nabla h(X(\tau)), 1^1 \rangle] \\ &= \mathbf{E} [(c^p \tau_a + \langle \nabla h(x^a), 1^1 \rangle) \mathbb{I}\{\tau_a \leq \tau_p\} + c^p \tau_p \mathbb{I}\{\tau_a > \tau_p\}], \end{aligned}$$

where $x^a = (\bar{r}^a, 0)^T$. On rearranging terms and applying the strong Markov property we obtain,

$$\langle \nabla h(x), 1^1 \rangle = c^p \mathbf{E}_x[\tau_p] + (\langle \nabla h(x^a), 1^1 \rangle - c^p \mathbf{E}_{x^a}[\tau_p]) \mathbf{P}\{\tau_a > \tau_p\}.$$

Consideration of the height process \mathbf{H}^p then gives,

$$(3.53) \quad \langle \nabla h(x), 1^1 \rangle = c^p \delta_H^{-1} r + (\langle \nabla h(x^a), 1^1 \rangle - c^p \delta_H^{-1} r_0) \mathbf{P}^H\{\tau_{r_0} > \tau_0\},$$

where $r_0 = \bar{r}^p - \bar{r}^a$ and the probability on the right hand side is with respect to the height process: Proposition 3.2 gives $\mathbf{P}^H\{\tau_{r_0} < \tau_0\} = (e^{\theta_H r_0} - 1)^{-1} (e^{\theta_H r} - 1)$.

Equation (3.52) provides an expression for the derivative of h at x^a ,

$$\langle \nabla h(x^a), 1^1 \rangle = -\langle \nabla h(x^a), 1^2 \rangle = -c^p \delta_H^{-1} \theta_H^{-1} (e^{\theta_H r_0} - \theta_H r_0 - 1).$$

Combining this identity with (3.53) we obtain,

$$\begin{aligned} \langle \nabla h, 1^1 \rangle &= c^p \delta_H^{-1} \left(r - \theta_H^{-1} (e^{\theta_H r_0} - 1) \right) \mathbf{P}^H\{\tau_{r_0} > \tau_0\} \\ &= c^p \delta_H^{-1} \left(r - \theta_H^{-1} (e^{\theta_H r} - 1) \right), \quad 0 \leq r \leq r_0. \end{aligned}$$

The right hand side is strictly negative for $r > 0$. \square

The proof of Theorem 2.6 is identical to the simpler setting of Section 3.3. We demonstrate that the solution to Poisson's equation h under the affine policy solves the dynamic programming equation,

$$\left(\mathcal{D}h_*(x) + c - \eta_* \right) \wedge \left(\inf \{ \langle \nabla h_*(x), -Bu \rangle : u \in \mathbb{R}_+^{K+1} \} \right) = 0, \quad x \in \mathbb{X},$$

where here we let B denote the $(K+1) \times m$ matrix defined by $B(1, i) = 1 = B(i+1, i) = 1$ for each i , and $B(i, j) = 0$ for all other indices (i, j) .

Proof of Theorem 2.6. The proof that $\langle \nabla h(x), B1^i \rangle$ is non-positive for $i = 1$ and $i = K+1$ is identical to the proof of Theorem 2.4. To obtain the analogous result for $i \in \{2, \dots, m\}$ we apply similar reasoning. Exactly as in (3.43), it can be shown that for $x \in \mathcal{R}^{a_i}$,

$$(3.54) \quad \langle \nabla h(x), B1^{i+1} \rangle = c^{a_i} \mathbb{E}[\tau_{a_i}] - c^{a_{i+1}} \mathbb{E}_x \left[\int_0^{\tau_{a_i}} \mathbb{I}\{R(t) \leq \bar{r}^{a_{i+1}}\} dt \right],$$

where $\tau_{a_i} := \inf\{t \geq 0 : R(t) = \bar{r}^{a_i}\}$.

To compute the right hand side of (3.54) we again construct a one-dimensional Brownian motion to represent these expectations. Define for $R(0) = q < \bar{r}^{a_i}$,

$$H^{a_i}(t) := \bar{r}^{a_i} - R(t) + \sum_{j>i} G^{a_j}(t), \quad t \geq 0.$$

This is described by,

$$dH^{a_i}(t) = -\zeta_i^+ dt + dI^{a_i}(t) + dD(t), \quad t \geq 0 \quad \text{while } G^{a_i}(t) > 0.$$

Consequently, we can write using (3.54),

$$(3.55) \quad \langle \nabla h(x), B1^{i+1} \rangle = c^{a_i} \mathbb{E}[\tau_{a_i}] - c^{a_{i+1}} \mathbb{E}_x \left[\int_0^{\tau_{a_i}} \mathbb{I}\{H^{a_i}(t) \geq (\bar{r}^{a_i} - \bar{r}^{a_{i+1}})\} dt \right],$$

and τ_{a_i} coincides with the first hitting time to the origin for H^{a_i} . Consequently, applying Proposition 3.2,

$$\langle \nabla h(x), B1^{i+1} \rangle = c^{a_i} \delta_H^{-1} r - c^{a_{i+1}} \left(\Psi(r) - 1 + \theta_H r \right) \left(\delta_H \theta_H e^{\theta_H (\bar{r}^{a_i} - \bar{r}^{a_{i+1}})} \right)^{-1}, \quad r \geq 0,$$

where $\theta_H = 2\zeta_i^+ / \sigma_D^2$, and the function Ψ is defined in (3.37). The remainder of the proof is identical to the proof of Proposition 3.6 using the formula $e^{-\theta_H \bar{r}^{a_i}} = c^{a_i} / c^{a_{i+1}}$. \square

4. CONCLUSIONS

Theorem 2.4 establishes an explicit formula for reserves in the dynamic newsboy model. Optimal reserves are high whenever there is high variability in demand, or significant ramping constraints on production. These conclusions are (qualitatively) consistent with the substantial reserves maintained in any major power market.

Many conclusions are reasonably robust to modeling assumptions. An investigation of the impact of uncertainty of supply is carried out in [44] based on ideas in this paper. Some preliminary numerical studies to investigate the impact of correlation are contained in [12]. In most cases a Brownian model predicts with reasonable accuracy approximate values for optimal hedging points, where the variance chosen in the diffusion model is the asymptotic variance for a discrete model (i.e. the Central Limit Theorem variance.)

Our focus in current research is the decentralized market problem, for which Theorem 2.4 and Theorem 2.5 have clear implications: If the prices $\{c^p, c^a\}$ are the prices charged to the consumer by the supplier, the supplier will assume that the consumer will optimize based on what it is charged. The formulae given in these two theorems quantify the observation: *When θ_p is small, then the fair price for ancillary service may be extremely high.* The parameter θ_p is small if there is significant variability in demand, or if the maximum ramp-up rate for primary service is small. This is an important observation critical for interpreting a market outcome sustaining the optimal allocation [15, 44].

It is a task of fundamental importance to build robust market rules that can withstand considerable volatility and possible strategic manipulation by the players. We need to move beyond a static analysis in order to address issues surrounding reliability and dynamics in a market setting.

REFERENCES

- [1] E. Altman, B. Gaujal, and A. Hordijk. Multimodularity, convexity, and optimization properties. *Math. Oper. Res.*, 25(2):324–347, 2000.
- [2] K. J. Arrow, Theodore Harris, and Jacob Marschak. Optimal inventory policy. *Econometrica*, 19:250–272, 1951.
- [3] R. Atar and A. Budhiraja. Singular control with state constraints on unbounded domain. *Ann. Probab.*, 34(5):1864–1909, 2006.
- [4] Rami Atar, Amarjit Budhiraja, and Ruth J. Williams. HJB equations for certain singularly controlled diffusions. *Adv. Appl. Probab.*, 17(5-6):1745–1776, 2007.
- [5] C. Bell. Characterization and computation of optimal policies for operating an M/G/1 queuing system with removable server. *Operations Res.*, 19:208–218, 1971.
- [6] S. L. Bell and R. J. Williams. Dynamic scheduling of a system with two parallel servers in heavy traffic with complete resource pooling: Asymptotic optimality of a continuous review threshold policy. *Ann. Appl. Probab.*, 11:608–649, 2001.
- [7] R. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.
- [8] R. Bellman, I. Glicksberg, and O. Gross. On the optimal inventory equation. *Management Sci.*, 2:83–104, 1955.
- [9] D. P. Bertsekas. A new value iteration method for the average cost dynamic programming problem. *SIAM J. Control Optim.*, 36(2):742–759 (electronic), 1998.
- [10] V. Borkar and A. Budhiraja. Ergodic control for constrained diffusions: characterization using HJB equations. *SIAM J. Control Optim.*, 43(4):1467–1492 (electronic), 2004/05.
- [11] A. N. Borodin and P. Salminen. *Handbook of Brownian motion—facts and formulae*. Probability and its Applications. Birkhäuser Verlag, Basel, first (second ed. published 2002) edition, 1996.
- [12] M. Chen. *Modelling and control of complex stochastic networks, with applications to manufacturing systems and electric power transmission networks*. PhD thesis, University of Illinois at Urbana Champaign, University of Illinois, Urbana, IL, USA, 2005.
- [13] M. Chen, I.-K. Cho, and S.P. Meyn. Reliability by design in a distributed power transmission network. *Automatica*, 42:1267–1281, August 2006. (invited).
- [14] M. Chen, C. Pandit, and S. P. Meyn. In search of sensitivity in network optimization. *Queueing Syst. Theory Appl.*, 44(4):313–363, 2003.
- [15] I.-K. Cho and S. P. Meyn. Efficiency and marginal cost pricing in dynamic competitive markets. To appear in *J. Theo. Economics*, 2006.
- [16] I.-K. Cho and S. P. Meyn. Optimization and the price of anarchy in a dynamic newsboy model. Stochastic Networks, invited session at the INFORMS Annual Meeting, November 13-16, 2005.
- [17] M. G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.
- [18] P. Dasgupta, L. Moser, and P. Melliard-Smith. Dynamic pricing for time-limited goods in a supplier-driven electronic marketplace. *Electronic Commerce Research*, 5:267–292, 2005.

- [19] C. L. DeMarco. Electric power network tutorial: Basic steady state and dynamic models for control, pricing, and optimization. http://www.ima.umn.edu/talks/workshops/3-7.2004/demarco/IMA_Power_Tutorial_3_2004.pdf, 2004.
- [20] E.A. DeMeo, G.A. Jordan, C. Kalich, J. King, M.R. Milligan, C. Murley, B. Oakleaf, and M.J. Schuerger. Accommodating wind's natural behavior. *Power and Energy Magazine, IEEE*, 5(6):59–67, Nov.-Dec. 2007.
- [21] B. T. Doshi. Optimal control of the service rate in an M/G/1 queueing system. *Adv. Appl. Probab.*, 10:682–701, 1978.
- [22] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23(4):1671–1691, 1995.
- [23] P. Dupuis and R. J. Williams. Lyapunov functions for semimartingale reflecting Brownian motions. *Ann. Appl. Probab.*, 22(2):680–702, 1994.
- [24] S. N. Ethier and T. G. Kurtz. *Markov Processes : Characterization and Convergence*. John Wiley & Sons, New York, 1986.
- [25] A. Federgruen, A. Hordijk, and H. C. Tijms. Denumerable state semi-Markov decision processes with unbounded costs, average cost criterion unbounded costs, average cost criterion. *Stoch. Proc. Applns.*, 9(2):223–235, 1979.
- [26] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1993.
- [27] S. B. Gershwin. *Manufacturing Systems Engineering*. Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [28] S.B. Gershwin. Production and subcontracting strategies for manufacturers with limited capacity and volatile demand. *IIE Transactions on Design and Manufacturing*, 32(2):891–906, 2000. Special Issue on Decentralized Control of Manufacturing Systems.
- [29] P. W. Glynn and S. P. Meyn. A Liapounov bound for solutions of the Poisson equation. *Ann. Probab.*, 24(2):916–931, 1996.
- [30] S. C. Graves. Safety stocks in manufacturing systems. *J. Manuf. Oper. Management*, 1(1):67–101, 1988.
- [31] B. Hajek. Optimal control of two interacting service stations. *IEEE Trans. Automat. Control*, AC-29:491–499, 1984.
- [32] B. Hajek. Extremal splittings of point processes. *Math. Oper. Res.*, 10(4):543–556, 1985.
- [33] J. M. Harrison and J. A. Van Mieghem. Dynamic control of Brownian networks: state space collapse and equivalent workload formulations. *Ann. Appl. Probab.*, 7(3):747–771, 1997.
- [34] J.M. Harrison. *The BIGSTEP approach to flow management in stochastic processing networks*, pages 57–90. *Stochastic Networks: Theory and Applications*. Oxford University Press, Oxford, UK, 1996.
- [35] S. G. Henderson, S. P. Meyn, and V. B. Tadić. Performance evaluation and policy selection in multiclass networks. *Discrete Event Dynamic Systems: Theory and Applications*, 13(1-2):149–189, 2003. Special issue on learning, optimization and decision making (invited).
- [36] D.P. Heyman. Optimal operating policies for M/G/1 queueing systems. *Operations Res.*, 16:362–382, 1968.
- [37] P. Hines, J. Apt, and S. Talukdar. Trends in the history of large blackouts in the united states. In *Proc. of the 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–8, jul. 2008.
- [38] IBM. Call center analytics and optimization. <http://domino.research.ibm.com/odis/odis.nsf/pages/micro.01.02.html>, 2005.
- [39] M. Ilic and J. Zaborszky. *Dynamics and Control of Large Electric Power Systems*. Wiley-Interscience, New York, 2000.
- [40] P. L. Joskow and J. Tirole. *Reliability and Competitive Electricity Market*. IDEI, University of Toulouse, 2004.
- [41] D. L. Kaufman, H.-S. Ahn, and M. E. Lewis. On the introduction of an agile, temporary workforce into a tandem queueing system. *Queueing Syst. Theory Appl.*, 51:135–171, 2005.
- [42] F.P. Kelly and C.N. Laws. Dynamic routing in open queueing networks: Brownian models, cut constraints and resource pooling. *Queueing Syst. Theory Appl.*, 13:47–86, 1993.
- [43] L. F. Martins, S. E. Shreve, and H. M. Soner. Heavy traffic convergence of a controlled, multiclass queueing system. *SIAM J. Control Optim.*, 34(6):2133–2171, November 1996.

- [44] S. Meyn, M. Negrete-Pincetic, G. Wang, A. Kowli, and E. Shafieepoorfard. The value of volatile resources in electricity markets. In *Proc. of the 49th Conf. on Dec. and Control*, pages 4921–4926, Atlanta, GA, 2010.
- [45] S. P. Meyn. The policy iteration algorithm for average reward Markov decision processes with general state space. *IEEE Trans. Automat. Control*, 42(12):1663–1680, 1997.
- [46] S. P. Meyn. Stability and optimization of queueing networks and their fluid models. In *Mathematics of stochastic manufacturing systems (Williamsburg, VA, 1996)*, pages 175–199. Amer. Math. Soc., Providence, RI, 1997.
- [47] S. P. Meyn. Stability, performance evaluation, and optimization. In E. Feinberg and A. Shwartz, editors, *Markov Decision Processes: Models, Methods, Directions, and Open Problems*, pages 43–82. Kluwer, Holland, 2001.
- [48] S. P. Meyn. Sequencing and routing in multiclass queueing networks. Part II: Workload relaxations. *SIAM J. Control Optim.*, 42(1):178–217, 2003.
- [49] S. P. Meyn. Dynamic safety-stocks for asymptotic optimality in stochastic networks. *Queueing Syst. Theory Appl.*, 50:255–297, 2005.
- [50] S. P. Meyn. Workload models for stochastic networks: Value functions and performance evaluation. *IEEE Trans. Automat. Control*, 50(8):1106–1122, August 2005.
- [51] S. P. Meyn. *Control Techniques for Complex Networks*. Cambridge University Press, Cambridge, 2007.
- [52] S. P. Meyn and R. L. Tweedie. Generalized resolvents and Harris recurrence of Markov processes. *Contemporary Mathematics*, 149:227–250, 1993.
- [53] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. Published in the Cambridge Mathematical Library. 1993 edition online: <http://black.csl.uiuc.edu/~meyn/pages/book.html>.
- [54] B. Mitchell. Optimal service-rate selection in an $M/G/\hat{1}$ queue. *SIAM J. Appl. Math.*, 24(1):19–35, 1973.
- [55] Sara Robinson. Math model explains volatile prices in power markets. *SIAM News*, Oct. 2005.
- [56] S. M. Ross. Arbitrary state Markovian decision processes. *Ann. Math. Statist.*, 39(1):2118–2122, 1968.
- [57] H. E. Scarf. The optimality of (S, s) policies in the dynamic inventory problem. In *Mathematical methods in the social sciences, 1959*, pages 196–202. Stanford Univ. Press, Stanford, Calif., 1960.
- [58] Herbert E. Scarf. Some remarks on Bayes solutions to the inventory problem. *Naval Res. Logist. Quart.*, 7:591–596, 1960.
- [59] E. Scherer. A linear programming approach to the steady-state analysis of reflected Brownian motion. *Stoch. Models*, 17(3):341–368, 2001.
- [60] S. P. Sethi and G. L. Thompson. *Optimal Control Theory: Applications to Management Science and Economics*. Kluwer Academic Publishers, Boston, 2000.
- [61] S.P. Sethi, H. Yan, J.H. Yan, and H. Zhang. An analysis of staged purchases in deregulated time-sequential electricity markets. *Journal of Industrial and Management Optimization*, 1:443–463, November 2005.
- [62] M. J. Sobel. Optimal average-cost policy for a queue with start-up and shut-down costs. *Operations Res.*, 17:145–162, 1969.
- [63] S. Stidham, Jr. and R. Weber. A survey of Markov decision models for control of networks of queues. *Queueing Syst. Theory Appl.*, 13(1-3):291–314, 1993.
- [64] A. L. Stolyar. Maxweight scheduling in a generalized switch: state space collapse and workload minimization in heavy traffic. *Adv. Appl. Probab.*, 14(1):1–53, 2004.
- [65] L. M. Wein. Dynamic scheduling of a multiclass make-to-stock queue. *Operations Res.*, 40(4):724–735, 1992.
- [66] P. Whittle. *Optimization over time. Vol. I*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1982. Dynamic programming and stochastic control.
- [67] P. Whittle. *Optimization over time. Vol. II*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1983. Dynamic programming and stochastic control.
- [68] M. Yadin and P. Naor. Queueing systems with a removable service station. *Operations Research Quarterly*, 14(4):393–405, 1963.

DEPARTMENT OF ECONOMICS, UNIVERSITY OF ILLINOIS, 1206 S. 6TH STREET, CHAMPAIGN, IL 61820
USA

E-mail address: inkoocho@uiuc.edu

URL: <http://www.business.uiuc.edu/inkoocho>

COORDINATED SCIENCE LABORATORY AND DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING,
UNIVERSITY OF ILLINOIS, 1308 W. MAIN STREET, URBANA, IL 61801 USA

E-mail address: meyn@illinois.edu

URL: <https://netfiles.uiuc.edu/meyn/www/spm.html>