SAMPLE AVERAGE APPROXIMATION AND MODEL PREDICTIVE CONTROL FOR INVENTORY OPTIMIZATION*

DOMINIC S. T. KEEHAN†, ANDY B. PHILPOTT†, AND EDWARD J. ANDERSON‡

Abstract. We study multistage stochastic optimization problems using sample average approximation (SAA) and model predictive control (MPC) as solution approaches. MPC is frequently employed when the size of the problem renders stochastic dynamic programming intractable, but it is unclear how this choice affects out-of-sample performance. To compare SAA and MPC out-of-sample, we formulate and solve an inventory control problem that is driven by random prices. Analytic and numerical examples are used to show that MPC can outperform SAA in expectation when the underlying price distribution is right-skewed. We also show that MPC is equivalent to a distributional robustification of the SAA problem with a first-moment based ambiguity set.

Key words. stochastic dynamic programming, sample average approximation, model predictive control, distributionally robust optimization

MSC codes. 90C15, 93E20, 90B05

1. Introduction.

Multistage stochastic optimization problems are in general very difficult to solve. Although one can create scenario-tree approximations of such problems based on samples of the random variables in each stage (called *sample average approximation* or SAA), the number of samples required to solve the true problem to ϵ-accuracy grows exponentially with the number of stages [10, 8] and the resulting optimization problems are computationally expensive to solve [3]. Beyond two-stage stochastic programming problems where the almost sure convergence of SAA has been thoroughly explored (see [9]), the performance of SAA on multistage problems has received little attention apart from the aforementioned negative results.

Multistage stochastic optimization problems become easier when the random variables are stage-wise independent or follow a Markov process and the problem can be formulated as a stochastic optimal control problem, where decisions are controls that affect state variables obeying some dynamics. In principle, such problems are amenable to solution by stochastic dynamic programming methods, or some approximate form of these, as long as the dimension of the state variable is not too large. Of course stochastic dynamic programming methods must compute expected values and so some discretization of the random variables is required to enable this. Here SAA provides a natural methodology and has the property that the sample expected values for a sample size N will converge almost surely by the strong law of large numbers to their true values as $N \to \infty$.

Stochastic optimal control problems do not have to be solved using a dynamic programming approach. In many practical settings (e.g., where state dimension is high and controls and states are subject to complicated constraints) *model predictive control* (MPC) can be used. There has been an enormous amount of work in control theory exploring the use of model predictive control in various contexts (see [5, 6]). In our situation we consider a relatively simple problem in which the state variables are fully observed, state constraints are simple, and we can find explicit solutions for the infinite horizon problems that we need. In this case the MPC approach fixes random

*Submitted to the editors DATE.

Funding: This work was funded by FUNDED.

†Electric Power Optimization Centre, University of Auckland (dkee331@aucklanduni.ac.nz, a.philpott@auckland.ac.nz).

‡Imperial College Business School, Imperial College London (e.anderson@imperial.ac.uk).

variables at their expectation and solves a deterministic optimal control problem. (One can either assume that the expectations are known exactly, or estimate them from a random sample. We focus on the second case in this work.) The optimal policy that solves this deterministic problem is applied in the first stage only and a new deterministic problem is formed from stage 2 onwards in a rolling horizon manner. There have been comparisons of SAA and MPC by simulation out-of-sample, and MPC does well in certain circumstances (see e.g. [4]). However, the reasons for this good performance have not been fully explored. Although the SAA and MPC solutions coincide when the certainty equivalence property holds [12, 13], this does not explain the success of MPC in more general conditions.

Our aim in this paper is to advance our understanding of SAA and MPC applied to stochastic control problems. To do this we restrict attention to a specific class of inventory problems with a one-dimensional state variable. This simple stochastic inventory control problem (SIC) seeks to maximize the expected reward from selling a fixed inventory of some item at a random and varying price over an infinite horizon. The price at each stage is assumed to be independent of other prices and identically distributed. At each stage the inventory held incurs an inventory cost that we assume is an increasing strictly convex function. This problem is simple enough to admit a closed-form optimal policy for any bounded price distribution, but complicated enough to provide a suitable laboratory to test the performance of SAA and MPC.

Given the SIC model and some ground-truth price distribution, for any price samples we can compute an SAA policy and compute its expected reward under the true price distribution. Similarly, we can compute an MPC policy based on the sample average of the random prices, and compute its expected reward under the true price distribution. The expectation of these two statistics over the sampling distribution gives a measure of out-of-sample performance of each approach. Our study is motivated by the question:

> Under what conditions does Model Predictive Control do better out of sample than the optimal dynamic programming solution based on Sample Average Approximation?

We observe that the performance of SAA is poor when price distributions have a long right tail. In this setting the price samples will occasionally contain a very high price, causing the SAA policy to anticipate high prices too frequently and pay too much in storage costs in the meantime. MPC policies attenuate this effect when it occurs and can perform better than SAA out-of-sample.

The paper is laid out as follows. We begin in section 2 by formulating our inventory problem and deriving a formula for its optimal solution as a function of the price probability distribution. This formula can be used to determine an SAA policy based on the empirical distribution of price samples, as well as an MPC policy based on the sample-average price. In section 3 we compare the out-of-sample performance of these two policies under some simple assumptions on the ground-truth price distribution, and provide conditions on the price samples which ensure that the MPC policy performs at least as well as the SAA policy. In section 4 we assume an exponential distribution for price and show that the expected out-of-sample improvement from using MPC instead of SAA becomes arbitrarily large as the discount factor approaches 1. In section 5 we report some numerical experiments that support the theoretical results of previous sections. We close the paper in section 6 by giving an interpretation of MPC as a distributional robustification of SAA that uses a moment-based ambiguity set, providing a different lens for viewing the performance differences of SAA and MPC.

2. A stochastic inventory control problem.

To study the performance of SAA and MPC, we will look at a particular stochastic inventory control problem that can be formulated as

$$\text{SIC:} \quad \max_{\{u_1,u_2,\ldots\}} \quad \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1}\left(P_t u_t - C(x_t)\right)\right]$$

where x_t and u_t satisfy

$$\begin{aligned} x_t &= x_{t-1} - u_t, \quad t = 1, 2, \ldots \\ u_t &\in [0, x_{t-1}], \quad t = 1, 2, \ldots, \end{aligned}$$

and u_t depends only on the price history $\{P_1, P_2, \ldots, P_t\}$ up to time t (i.e. the standard non-anticipativity constraints). The value of $x_0 \geq 0$, the initial inventory level, is given. Here $\beta \in (0,1)$ is a discount factor, P_t is a random price with finite expectation and C is an increasing strictly convex and differentiable function with derivative c. Because c is a strictly increasing continuous function, we may define an inverse function, c^{-1}, on the range of c. The problem SIC can be interpreted as the problem facing a merchant who maximizes expected discounted reward by selling at each time t an amount of stock u_t at a realisation of the random price P_t from their current inventory x_{t-1}, while incurring a storage cost $C(x_{t-1} - u_t)$ on their remaining inventory.

In what follows, we analyse the optimal solution of SIC and approximations of SIC that come from either an empirical distribution using a set of samples drawn from $\{P_t\}$ or assuming the price is fixed. To keep this analysis simple we make following assumptions:

ASSUMPTION 2.1. *The random prices P_t are independent and identically distributed on a bounded interval $[p_{\mathrm{L}}, p_{\mathrm{U}}]$, having probability distribution $\mathbb{P}$.*

ASSUMPTION 2.2. *The inventory cost is a continuously differentiable function $C : \mathbb{R}_+ \mapsto \mathbb{R}_+$ with $C(0) = 0$ and $\lim_{x\to\infty} c(x) = \infty$.*

Under Assumption 2.1, we drop dependence of the random price P_t on the index t and for $x \geq 0$ define the dynamic programming functional equation

$$\tilde{V}(x) = \mathbb{E}\left[\max_{0\leq u\leq x}\left\{Pu - C(x-u) + \beta\tilde{V}(x-u)\right\}\right]. \tag{2.1}$$

Observe that the mapping $(u,p) \mapsto pu - C(x-u)$ is bounded on the compact set $[0,x] \times [p_{\mathrm{L}}, p_{\mathrm{U}}]$ and $\beta < 1$. It follows that SIC has a finite optimal value, and by Theorem 9.2 of [11, p. 246] this is equal to $\tilde{V}(x_0)$. In addition, the mapping $x \mapsto pu - C(x-u)$ is continuous and strictly concave and the feasible region $[0,x]$ is a convex set. Strict concavity of $\tilde{V}(x)$ then follows by Theorem 9.8 of [11, p. 265]. With $\tilde{V}(x)$ strictly concave and bounded on bounded sets, it follows that $\tilde{V}(x)$ is also continuous and therefore must have a non-empty superdifferential which we denote by $\partial\tilde{V}(x)$.

For a given price p and current inventory x the optimum expected discounted reward from this point on is given by

$$V(x,p) = \max_{0 \le u \le x} \left\{ pu - C(x-u) + \beta \tilde{V}(x-u) \right\}, \tag{2.2}$$

where the optimal choice of action is given by the maximizing value u.

Denote the projection of $y \in \mathbb{R}$ onto the closed interval $[a,b]$ by $(y)_{[a,b]} = \max\{a, \min\{b,y\}\}$. We write $(y)_{[a,\infty)} = \max\{a,y\}$ and $(y)_+ = \max\{y,0\}$.

Proposition 2.3. *Under Assumptions* 2.1 *and* 2.2*, the right-hand side of* (2.2) *has optimal solution*

$$u(x,p) = x - c^{-1}\left((\beta \mathbb{E}[(P-p)_+] + \beta p \;\; - p)_{[c(0),c(x)]} \right).$$

Proof. Observe that the change of variables $w = x - u$ yields

$$V(x,p) = \max_{0 \le w \le x} \{p(x-w) - C(w) + \beta \tilde{V}(w)\}. \tag{2.3}$$

Let

$$\varphi_p(w) = p(x-w) - C(w) + \beta \tilde{V}(w).$$

For any values of x and p the mapping $w \mapsto \varphi_p(w)$ is strictly concave and has a nonempty superdifferential $\partial\varphi_p(w)$, so for $x \ge 0$ the optimization $\max_{0 \le w \le x} \varphi_p(w)$ has a unique solution $w^*(x,p) \in [0,x]$ satisfying

$$0 \in \partial\varphi_p(w^*(x,p)) + \mathcal{N}(w^*(x,p)),$$

where $\mathcal{N}(w^*(x,p))$ is the normal cone of $[0,x]$ at $w^*(x,p)$. Since the derivative $c(w)$ is strictly increasing and unbounded above, $\varphi_p(w)$ is decreasing for w large enough and there will be a unique solution $w(p)$ to $\max_{w \ge 0} \varphi_p(w)$ which is equal to $w^*(x,p)$ when projected onto $[0,x]$. Observe that the function $w(p)$ is decreasing, and it follows that for any x there exists some critical price $p_C(x)$ such that for $p \ge p_C(x)$ we have $w(p) \le x$ and for $p \le p_C(x)$ we have $w(p) \ge x$.

Denote by $\partial V_p(x)$ the superdifferential of the mapping $x \mapsto V(x,p)$. When $p \ge p_C(x)$, we have $w(p) \le x$, so $w^*(x,p) = (w(p))_+$ and

$$V(x,p) = p(x - (w(p))_+) - C((w(p))_+) + \beta \tilde{V}((w(p))_+).$$

In this case it follows that $p \in \partial V_p(x)$.

On the other hand, when $p \le p_C(x)$ we have $w(p) \ge x$, so $w^*(x,p) = x$ and

$$V(x,p) = -C(x) + \beta \tilde{V}(x). \tag{2.4}$$

For all $x > 0$, (2.4) implies that

$$-c(x) + \beta \partial \tilde{V}(x) \subseteq \partial V_p(x).$$

So any $\tilde{g} \in \partial \tilde{V}(x)$ defines a supergradient $-c(x) + \beta \tilde{g}$ in $\partial V_p(x)$. Let

$$h(\tilde{g},p) = \begin{cases} p, & p \ge p_C(x) \\ -c(x) + \beta \tilde{g}, & p < p_C(x) \end{cases}.$$

By Theorem 7.46 of [9, p. 371], $\tilde{V}(x) = \mathbb{E}[V(x, P)]$ has directional derivatives at every x, so

$$\mathbb{E}[h(\tilde{g}, P)] \in \partial \tilde{V}(x).$$

It is easy to see that the mapping $T : \partial\tilde{V}(x) \mapsto \partial\tilde{V}(x)$ defined by

$$T(\tilde{g}) = (\beta\tilde{g} - c(x))\mathbb{P}[P < p_{\mathrm{C}}(x)] + \mathbb{E}[P|P \geq p_{\mathrm{C}}(x)]\mathbb{P}[P \geq p_{\mathrm{C}}(x)]$$

is a contraction with Lipschitz constant strictly less than 1, since for any $\tilde{g}, \tilde{g}' \in \partial\tilde{V}(x)$

$$|T(\tilde{g}) - T(\tilde{g}')| = |\tilde{g} - \tilde{g}'|\,\beta\mathbb{P}[P < p_{\mathrm{C}}(x)] < |\tilde{g} - \tilde{g}'|\,.$$

As $\partial\tilde{V}(x)$ is a nonempty closed set, by the Banach fixed point theorem, there is a unique $\tilde{g}(x) \in \partial\tilde{V}(x)$ satisfying $T(\tilde{g}(x)) = \tilde{g}(x)$. But this implies

$$\tilde{g}(x) = (\beta\tilde{g}(x) - c(x))\mathbb{P}[P < p_{\mathrm{C}}(x)] + \mathbb{E}[P|P \geq p_{\mathrm{C}}(x)]\mathbb{P}[P \geq p_{\mathrm{C}}(x)]$$

so

$$\tilde{g}(x) = \frac{\mathbb{E}[P|P \geq p_{\mathrm{C}}(x)]\mathbb{P}[P \geq p_{\mathrm{C}}(x)] - c(x)\mathbb{P}[P < p_{\mathrm{C}}(x)]}{1 - \beta\mathbb{P}[P < p_{\mathrm{C}}(x)]} \in \partial\tilde{V}(x). \tag{2.5}$$

We now construct an optimal solution $w(p)$ to $\max_{w \geq 0} \varphi_p(w)$ as follows. First observe that $\beta(\mathbb{E}[(P - p)_+] + p) - p$ is a strictly decreasing continuous function of p. If

$$\beta(\mathbb{E}[(P - p)_+] + p) - p > c(0)$$

for all $p \in [p_{\mathrm{L}}, p_{\mathrm{U}}]$ then set $p_{\mathrm{Z}} = p_{\mathrm{U}}$. Otherwise let p_{Z} be the unique solution to $\beta(\mathbb{E}[(P - p)_+] + p) - p = c(0)$. We now define

$$w(p) = \begin{cases} c^{-1}(\beta(\mathbb{E}[(P - p)_+] + p) - p), & p < p_{\mathrm{Z}} \\ 0, & p \in [p_{\mathrm{Z}}, p_{\mathrm{U}}] \end{cases}$$

If $p < p_{\mathrm{Z}}$ then we have $w(p) > 0$ and

$$\begin{aligned} w(p) &= c^{-1}\left(\beta(\mathbb{E}[(P - p)_+] + p) - p\right) \\ &= c^{-1}\left(\beta\left(\mathbb{E}[P|P \geq p]\mathbb{P}[P \geq p] + p\mathbb{P}[P < p]\right) - p\right). \end{aligned}$$

We can rearrange this to give

$$(1 - \beta\mathbb{P}[P < p])p + c(w(p)) = \beta\mathbb{P}[P \geq p]\mathbb{E}[P \mid P \geq p]. \tag{2.6}$$

Thus

$$(1 - \beta\mathbb{P}[P < p])(p + c(w(p))) = -\beta c(w(p))\mathbb{P}[P < p] + \beta\mathbb{P}[P \geq p]\mathbb{E}[P \mid P \geq p],$$

and hence

$$-p - c(w(p)) + \beta\frac{-c(w(p))\mathbb{P}[P < p] + \mathbb{E}[P|P \geq p]\mathbb{P}[P \geq p]}{1 - \beta\mathbb{P}[P < p]} = 0. \tag{2.7}$$

The definition of $p_{\rm C}$ implies that $p = p_{\rm C}(w(p))$, and so (2.7) implies that if we define $\tilde{g}(w(p))$ by (2.5) then

$$-p - c(w(p)) + \beta\tilde{g}(w(p)) = 0,$$

and $0 \in \partial\varphi_p(w^*(x,p))$ showing that $w(p)$ solves $\max_{w\geq 0}\varphi_p(w)$.

If $p = p_{\rm Z}$ then a similar analysis shows that $\tilde{g}(0)$ satisfies

$$-p_{\rm Z} - c(0) + \beta\tilde{g}(0) = 0$$

so for $p \geq p_{\rm Z}$ the right-hand derivative of $p(x-w) - C(w) + \beta\mathbb{E}[V(w,P)]$ at $w = 0$ is less than or equal to 0 implying that $w(p) = 0$ solves $\max_{w\geq 0}\varphi_p(w)$.

Combining both cases and projecting $w(p)$ onto $[0,x]$ yields

$$w^*(x,p) = c^{-1}\left((\beta\mathbb{E}[(P-p)_+] + \beta p - p)_{[c(0),c(x)]}\right)$$

and

$$u(x,p) = x - c^{-1}\left((\beta\mathbb{E}[(P-p)_+] + \beta p - p)_{[c(0),c(x)]}\right).$$

□

Proposition 2.3 shows that SIC has an optimal target inventory level

$$w^*(x,p) = c^{-1}\left((\beta\mathbb{E}[(P-p)_+] + \beta p - p)_{[c(0),c(x)]}\right)$$

at which the marginal cost of storage is as close as possible to the discounted expected increase in price above p in the next stage. The optimal SIC policy is then to reduce the current inventory level to $w^*(x,p)$ if it is not already at $w^*(x,p)$ by selling surplus stock.

Proposition 2.3 makes no assumptions about the probability distribution $\mathbb{P}$, except that it has bounded support. Thus $\mathbb{P}$ could have a density f with bounded support giving the optimal policy

$$x - c^{-1}\left(\left(\beta\int_p^{p_{\rm U}}(q-p)f(q)dq + \beta p - p\right)_{[c(0),c(x)]}\right),$$

or could consist of an empirical distribution on N price samples $q_1, q_2, \ldots, q_N$ with $\mathbb{P}(q_i) = \frac{1}{N}$, giving the SAA policy

$$u_{\rm S}(x,p) := x - c^{-1}\left(\left(\beta\frac{1}{N}\sum_{i=1}^{N}(q_i - p)_+ + \beta p - p\right)_{[c(0),c(x)]}\right). \tag{2.8}$$

We can also obtain an MPC policy from the samples $q_1, q_2, \ldots, q_N$ by planning using the sample average $\bar{q} = \frac{1}{N}\sum_{i=1}^{N} q_i$. In this case Proposition 2.3 would use the probability distribution that assigns probability 1 to $\bar{q}$, giving $\mathbb{E}[(P-p)_+] = (\bar{q}-p)_+$ so

$$u_{\rm M}(x,p) := x - c^{-1}\left((\beta(\bar{q}-p)_+ + \beta p - p)_{[c(0),c(x)]}\right). \tag{2.9}$$

For an initial inventory level x, the sample-based policies each have a critical price (that we denote by $p_{\rm S}(x)$ and $p_{\rm M}(x)$ for the SAA and MPC policies, respectively) which is the minimum price required to be offered to the vendor for any stock to be sold. The critical price $p_{\mathrm{S}}(x)$ is the unique p that solves $\beta \frac{1}{N}\sum_{i=1}^{N}(q_i - p)_+ + \beta p - p = c(x)$ and a similar definition holds for $p_{\mathrm{M}}(x)$. Depending on the samples $q_1, q_2, \ldots, q_N$, each sample-based policy will either pay too much in storage costs by selling too little stock, or not be able to take full advantage of future high prices having sold too much stock. By Jensen's inequality, $(\mathbb{E}[P] - p)_+ \leq \mathbb{E}[(P-p)_+]$, whereby $p_{\mathrm{M}}(x) \leq p_{\mathrm{S}}(x)$ and $u_{\mathrm{M}}(x,p) \geq u_{\mathrm{S}}(x,p)$. In this way, the policy u_{M} requires a lower price to sell stock than the policy u_{S} and sells at least as much. We will explore the implications of this observation in the next section.

3. Out-of-sample performance.

The assumption that P lies within a bounded interval $[p_{\mathrm{L}}, p_{\mathrm{U}}]$ is restrictive. Assumption 3.1 allows us to study the out-of-sample performance of the sample-based policies (derived using Proposition 2.3 on sample-based distributions that are discrete and therefore bounded) even when the underlying distribution is unbounded.

Assumption 3.1. *The random prices P_t are independent and identically distributed, having a probability distribution $\mathbb{P}$ with support on $\mathbb{R}_+$, a finite mean, and no atoms.*

Suppose we observe N price samples $q_1, q_2, ..., q_N$ and use these to inform the sample-based policies as in (2.8) and (2.9). The value of the SIC problem if the (possibly sub-optimal) SAA policy is used is

$$\bar{V}_{\mathrm{S}}(x_0) := \sum_{t=1}^{\infty} \beta^{t-1}\mathbb{E}\left[P u_{\mathrm{S}}(x_{t-1}, P) - C(x_t)\right] \tag{3.1}$$

where the values x_t are random variables determined by successive prices and derived from an the initial value x_0 using the actions u_{S}. This is well-defined since the infinite series is easily shown to be convergent: the expectations at each stage are bounded and they are discounted by $\beta < 1$. To show boundedness, we note $x_t \leq x_0$, C is non-negative and an increasing function, and $u_{\mathrm{S}}(x_{t-1}, P) \leq x_{t-1} \leq x_0$, and thus

$$-C(x_0) \leq \mathbb{E}\left[P u_{\mathrm{S}}(x_{t-1}, P) - C(x_t)\right] \leq \mathbb{E}\left[P\right] x_0.$$

Having defined $\bar{V}_{\mathrm{S}}$ as a function of the initial inventory, we also have $\bar{V}_{\mathrm{S}}$ satisfying the associated functional equation

$$\bar{V}_{\mathrm{S}}(x) = \mathbb{E}\left[P u_{\mathrm{S}}(x,P) - C(x - u_{\mathrm{S}}(x,P)) + \beta \bar{V}_{\mathrm{S}}(x - u_{\mathrm{S}}(x,P))\right].$$

Similarly, the value of the SIC problem if the (possibly sub-optimal) MPC policy is used is well-defined and has an associated functional equation

$$\bar{V}_{\mathrm{M}}(x) = \mathbb{E}\left[P u_{\mathrm{M}}(x,P) - C(x - u_{\mathrm{M}}(x,P)) + \beta \bar{V}_{\mathrm{M}}(x - u_{\mathrm{M}}(x,P))\right].$$

It is convenient to define

$$B(x) := \frac{1}{1-\beta}\max\{C(x), \mathbb{E}\left[P\right] x\}.$$

Then the bounds on the individual terms in $\bar{V}_{\mathrm{S}}$ and $\bar{V}_{\mathrm{M}}$ show that $B(x)$ is an upper bound on both $|\bar{V}_{\mathrm{S}}(x)|$ and $|\bar{V}_{\mathrm{M}}(x)|$.

3.1. Derivative of the expected value function. Before making comparisons between $\bar{V}_{\mathrm{S}}$ and $\bar{V}_{\mathrm{M}}$ we will first calculate their derivatives with respect to the initial inventory. It will be helpful to use a result of [9, p. 369], who give the following result (Theorem 7.44). Suppose that $F : \mathbb{R}^n \times \Omega \to R$ is a random function with expected value $f(x) = \mathbb{E}[F(x, \omega)]$.

LEMMA 3.2. *If the following conditions hold:*

(A) The expectation $f(x_0)$ is well defined and finite valued at some point $x_0 \in \mathbb{R}^n$;

(B) There exists a positive valued random variable $L(\omega)$ such that $\mathbb{E}[L(\omega)] < \infty$, and for all x_1, x_2 in a neighbourhood of x_0 and almost every $\omega \in \Omega$, $|F(x_1, \omega) - F(x_2, \omega)| \le L(\omega)\|x_1 - x_2\|$;

(C) For almost every ω the function $F(x, \omega)$ is differentiable with respect to x at x_0;

then $f(x)$ is differentiable at x_0 and

$$\nabla f(x_0) = \mathbb{E}[\nabla_x F(x_0, \omega)].$$

Now we can establish the derivative values. Since $\bar{V}_{\mathrm{M}}$ is undefined for $x < 0$ the derivative $\frac{d}{dx}\bar{V}_{\mathrm{M}}(0)$ does not exist. However, at $x = 0$ the function $\bar{V}_{\mathrm{M}}(x)$ does have a right derivative, and for the rest of this paper the expression $\frac{d}{dx}\bar{V}_{\mathrm{M}}(x)$ implicitly refers to this right derivative when $x = 0$.

LEMMA 3.3. *Under Assumptions* 2.2 *and* 3.1*, each of the derivatives of the expected value functions exist and are given by*

$$\frac{d}{dx}\bar{V}_{\mathrm{S}}(x) = \frac{\mathbb{E}[P|P \ge p_{\mathrm{S}}(x)]\mathbb{P}[P \ge p_{\mathrm{S}}(x)] - c(x)\mathbb{P}[P < p_{\mathrm{S}}(x)]}{1 - \beta\mathbb{P}[P < p_{\mathrm{S}}(x)]}$$

and

$$\frac{d}{dx}\bar{V}_{\mathrm{M}}(x) = \frac{\mathbb{E}[P|P \ge p_{\mathrm{M}}(x)]\mathbb{P}[P \ge p_{\mathrm{M}}(x)] - c(x)\mathbb{P}[P < p_{\mathrm{M}}(x)]}{1 - \beta\mathbb{P}[P < p_{\mathrm{M}}(x)]}.$$

Proof. The proof proceeds by first showing that the derivatives exist and then determining their values by a recursion. We begin by considering $\bar{V}_{\mathrm{S}}(x_0)$. For a particular realisation $\omega = \{p_1, p_2, \ldots\}$ of the random variables $\{P_1, P_2, \ldots\}$ the value function is determined by

$$V_{\mathrm{S}}(x_0, \omega) = \sum_{t=1}^{\infty} \beta^{t-1}(p_t u_{\mathrm{S}}(x_{t-1}, p_t) - C(x_t)) \tag{3.2}$$

The expectation of this is $\bar{V}_{\mathrm{S}}(x_0)$ and is well-defined, satisfying condition (A) of Lemma 3.2. Consider a realization of (3.2) with prices $\{p_1, p_2, \ldots\}$. Assume that there is some minimal index, T such that $p_T \ge p_{\mathrm{S}}(x_0)$, the critical price. Since $\mathbb{P}$ has no atoms, we know that $p_T > p_{\mathrm{S}}(x_0) > \max\{p_1, p_2, \ldots, p_{T-1}\}$ almost surely. The SAA policy with this price realisation will sell no stock until period T and the inventory levels are fixed at $x_t = x_0$ up to this point. At time T the SAA policy sells stock $u_{\mathrm{S}}(x_{T-1}, p_T)$ for the price p_T. The resulting inventory level is $x_T = c^{-1}\big(\big(\beta\frac{1}{N}\sum_i (q_i - p_T)_+ + \beta p_T - p_T\big)_{[c(0),\infty)}\big)$ which is independent of x_0. Thus for all $t' > T$ the inventory levels $x_{t'}$ are also independent of x_0. Now, $p_{\mathrm{S}}(x)$ is a continuous function of x which means that $p_T > p_{\mathrm{S}}(x) > \max\{p_1, p_2, \dots, p_{T-1}\}$ also holds for x in a neighbourhood $\mathcal{N}$ about x_0. This allows us to track the change in $V_{\mathrm{S}}(x, \omega)$ for different initial inventories x in this neighbourhood. If $x_1 > x_2$ then

$$V_{\mathrm{S}}(x_1, \omega) - V_{\mathrm{S}}(x_2, \omega) = p_T(x_1 - x_2) - \sum_{t=1}^{T-1} \beta^{t-1}(C(x_1) - C(x_2)). \tag{3.3}$$

This has an absolute value upper bounded by $\theta(p_1, p_2, \dots)|x_1 - x_2|$ where

$$\theta(p_1, p_2, \dots) = p_T + \frac{1}{1-\beta} 2c(x_0)$$

and we choose $\mathcal{N}$ small enough so that for all $x \in \mathcal{N}$ we have the derivative $c(x) < 2c(x_0)$. In the case that $p_t < p_{\mathrm{S}}(x_0)$ for all t, so that p_T is not defined, we can find a neighbourhood of x_0 where (3.3) is replaced by

$$V_{\mathrm{S}}(x_1, \omega) - V_{\mathrm{S}}(x_2, \omega) = -\sum_{t=1}^{\infty} \beta^{t-1}(C(x_1) - C(x_2)) \tag{3.4}$$

and use a similar argument to show that $\theta(p_1, p_2, \dots)$ is also a Lipschitz constant for $V_{\mathrm{S}}(x_0, \omega)$ in a neighbourhood about x_0 in this case. Now

$$\mathbb{E}[\theta(P_1, P_2, \dots)] \le \mathbb{E}[P | P > p_{\mathrm{S}}(x_0)] + \frac{1}{1-\beta} 2c(x_0) < \infty.$$

So the existence of the function $\theta(p_1, p_2, \dots)$ verifies condition (B) of Lemma 3.2. Moreover, it is easy to see that (3.3) and (3.4) imply a well-defined derivative of $V_{\mathrm{S}}(x_0, \omega)$ for almost all ω, hence satisfying the final condition (C) of Lemma 3.2. Thus we can use this result to show that $\frac{d}{dx_0} \bar{V}_{\mathrm{S}}(x_0)$ exists and is finite. The proof is entirely similar for $\frac{d}{dx_0} \bar{V}_{\mathrm{M}}(x_0)$.

Let $\tilde{w}(p) = c^{-1}\big(\big(\beta \frac{1}{N} \sum_i (q_i - p)_+ + \beta p - p\big)_{[c(0), \infty)} \big)$. We can define

$$V_{\mathrm{S}}(x, p) = \begin{cases} -C(x) + \beta \bar{V}_{\mathrm{S}}(x) & p < p_{\mathrm{S}}(x) \\ p(x - \tilde{w}(p)) - C(\tilde{w}(p)) + \beta \bar{V}_{\mathrm{S}}(\tilde{w}(p)) & p \ge p_{\mathrm{S}}(x) \end{cases}.$$

Then $V_{\mathrm{S}}(x, p)$ is the expected value from following the SAA policy with initial inventory x and initial price p. So $\bar{V}_{\mathrm{S}}(x) = \mathbb{E}\left[V_{\mathrm{S}}(x, p)\right]$. We can use the same approach as above, making use of the fact that $\frac{d}{dx} \bar{V}_{\mathrm{S}}(x)$ is well-defined to show that $\frac{d}{dx} \bar{V}_{\mathrm{S}}(x) = \mathbb{E}\left[\frac{d}{dx} V_{\mathrm{S}}(x, p)\right]$. Thus

$$\frac{d}{dx} V_{\mathrm{S}}(x, p) = \begin{cases} -c(x) + \beta \frac{d}{dx} \bar{V}_{\mathrm{S}}(x) & p < p_{\mathrm{S}}(x) \\ p & p \ge p_{\mathrm{S}}(x) \end{cases}.$$

Taking expectations we derive

$$\frac{d}{dx} \bar{V}_{\mathrm{S}}(x) = \left(\beta \frac{d}{dx} \bar{V}_{\mathrm{S}}(x) - c(x)\right) \mathbb{P}[P < p_{\mathrm{S}}(x)] + \mathbb{E}[P | P \ge p_{\mathrm{S}}(x)] \mathbb{P}[P \ge p_{\mathrm{S}}(x)]$$

and rearranging gives the required expression:

$$\frac{d}{dx} \bar{V}_{\mathrm{S}}(x) = \frac{\mathbb{E}[P | P \ge p_{\mathrm{S}}(x)] \mathbb{P}[P \ge p_{\mathrm{S}}(x)] - c(x) \mathbb{P}[P < p_{\mathrm{S}}(x)]}{1 - \beta \mathbb{P}\left[P < p_{\mathrm{S}}(x)\right]}.$$

The expression for $\frac{d}{dx} \bar{V}_{\mathrm{M}}(x)$ can be derived via identical reasoning. □

3.2. Comparing MPC and SAA. Our approach to compare the two different policies is to consider starting with the MPC policy and then switching to the SAA policy after a certain number of stages.

DEFINITION 3.4. *Let*

$$\bar{D}_1(x) := \mathbb{E}\left[Pu_{\mathrm{S}}(x,P) - C(x - u_{\mathrm{S}}(x,P)) + \beta\bar{V}_{\mathrm{M}}(x - u_{\mathrm{S}}(x,P))\right],$$

and for $t > 1$*,*

$$\bar{D}_t(x) := \mathbb{E}\left[Pu_{\mathrm{S}}(x,P) - C(x - u_{\mathrm{S}}(x,P)) + \beta\bar{D}_{t-1}(x - u_{\mathrm{S}}(x,P))\right]. \tag{3.5}$$

The value $\bar{D}_t(x_0)$ is the value of the SIC problem if the policy u_{S} is used for t stages and the policy u_{M} is used forevermore. It is clear that $\bar{D}_t$ is bounded in the same way that $\bar{V}_{\mathrm{S}}$ and $\bar{V}_{\mathrm{M}}$ are bounded, so Theorem 9.2 of [11, p. 246] again holds.

PROPOSITION 3.5. $\lim_{t\to\infty}\left|\bar{D}_t(x) - \bar{V}_{\mathrm{S}}(x)\right| = 0$.

Proof. The values $\bar{D}_t(x_0)$ and $\bar{V}_{\mathrm{S}}(x_0)$ both implement the policy u_{S} for the first t periods when starting with initial inventory x_0. So

$$\left|\bar{D}_t(x_0) - \bar{V}_{\mathrm{S}}(x_0)\right| = \left|\mathbb{E}\left[\beta^t\left(\bar{V}_{\mathrm{M}}(x_t) - \bar{V}_{\mathrm{S}}(x_t)\right)\right]\right| \le \beta^t 2B(x_0)$$

where the expectation is taken with respect to the value x_t which is a random variable under the application of the policy u_{S}. As $t \to \infty$, the bound $\beta^t 2B(x_0) \to 0$. Thus, $\lim_{t\to\infty}\left|\bar{D}_t(x_0) - \bar{V}_{\mathrm{S}}(x_0)\right| = 0$. Replacing x_0 with x concludes the proof. ∎

LEMMA 3.6. *If* $\bar{V}_{\mathrm{M}}(x) \ge \bar{D}_1(x)$ *for all* $x \in [0, x_0]$*, then* $\bar{V}_{\mathrm{M}}(x_0) \ge \bar{V}_{\mathrm{S}}(x_0)$*.*

Proof. We will first show that $\bar{D}_t(x) \ge \bar{D}_{t+1}(x)$ for all t via induction. Since $x - u_{\mathrm{S}}(x,p) \in [0, x_0]$ for all $x \in [0, x_0]$, by the assumption in the statement of the lemma $\bar{V}_{\mathrm{M}}(x - u_{\mathrm{S}}(x,p)) \ge \bar{D}_1(x - u_{\mathrm{S}}(x,p))$. Thus

$$\begin{aligned}\bar{D}_1(x) &= \mathbb{E}\left[Pu_{\mathrm{S}}(x,P) - C(x - u_{\mathrm{S}}(x,P)) + \beta\bar{V}_{\mathrm{M}}(x - u_{\mathrm{S}}(x,P))\right]\\ &\ge \mathbb{E}\left[Pu_{\mathrm{S}}(x,P) - C(x - u_{\mathrm{S}}(x,P)) + \beta\bar{D}_1(x - u_{\mathrm{S}}(x,P))\right] = \bar{D}_2(x).\end{aligned} \tag{3.6}$$

We make the inductive hypothesis: $\bar{D}_{t-1}(x) \ge \bar{D}_t(x)$ for all $x \in [0, x_0]$. Of course $x - u_{\mathrm{S}}(x,p) \in [0, x_0]$ still holds, and by the inductive hypothesis $\bar{D}_{t-1}(x - u_{\mathrm{S}}(x,p)) \ge \bar{D}_t(x - u_{\mathrm{S}}(x,p))$ for all $x \in [0, x_0]$, so applying to (3.5) a similar line of reasoning as in (3.6) shows that $\bar{D}_t(x) \ge \bar{D}_{t+1}(x)$, as required. Setting $x = x_0$ then shows that $\bar{V}_{\mathrm{M}}(x_0) \ge \bar{D}_t(x_0)$ for all $t \ge 1$. Thus, $\bar{V}_{\mathrm{M}}(x_0) \ge \lim_{t\to\infty}\bar{D}_t(x_0) = \bar{V}_{\mathrm{S}}(x_0)$ where Proposition 3.5 yields the final equality. ∎

PROPOSITION 3.7. *Assume* $\mathbb{P}$ *has a density* f*. Under Assumptions* 2.2 *and* 3.1*, if* $c(x) \ge \beta\int_{p_{\mathrm{S}}(x)}^{\infty} pf(p)dp$ *for all* $x \in [0, x_0]$*, then* $\bar{V}_{\mathrm{M}}(x_0) \ge \bar{V}_{\mathrm{S}}(x_0)$*.*

Proof. In the context of the proposition we will first show that $\frac{d}{dx}\bar{V}_{\mathrm{M}}(x) \ge \frac{d}{dx}\bar{D}_1(x)$ for all $x \in [0, x_0]$. As in Lemma 3.3

$$\frac{d}{dx}\bar{V}_{\mathrm{M}}(x) = \left(\beta\frac{d}{dx}\bar{V}_{\mathrm{M}}(x) - c(x)\right)\int_{-\infty}^{p_{\mathrm{M}}(x)} f(p)dp + \int_{p_{\mathrm{M}}(x)}^{\infty} pf(p)dp.$$

Inspecting $\bar{D}_1(x)$ shows that the similar expression

$$\frac{d}{dx}\bar{D}_1(x) = \left(\beta\frac{d}{dx}\bar{V}_{\mathrm{M}}(x) - c(x)\right)\int_{-\infty}^{p_{\mathrm{S}}(x)} f(p)dp + \int_{p_{\mathrm{S}}(x)}^{\infty} pf(p)dp$$

also holds. Recalling $p_{\mathrm{S}}(x) \geq p_{\mathrm{M}}(x)$, it can be seen that

$$\frac{d}{dx}\bar{V}_{\mathrm{M}}(x) - \frac{d}{dx}\bar{D}_1(x) = -\left(\beta \frac{d}{dx}\bar{V}_{\mathrm{M}}(x) - c(x)\right)\int_{p_{\mathrm{M}}(x)}^{p_{\mathrm{S}}(x)} f(p)dp + \int_{p_{\mathrm{M}}(x)}^{p_{\mathrm{S}}(x)} pf(p)dp.$$

Using Lemma 3.3, we may write

$$\begin{aligned}\beta \frac{d}{dx}\bar{V}_{\mathrm{M}}(x) - c(x) &= \beta \frac{\int_{p_{\mathrm{M}}(x)}^{\infty} pf(p)dp - c(x)\int_{-\infty}^{p_{\mathrm{M}}(x)} f(p)dp}{1 - \beta\int_{-\infty}^{p_{\mathrm{M}}(x)} f(p)dp} - c(x) \\ &= \frac{\beta\int_{p_{\mathrm{M}}(x)}^{\infty} pf(p)dp - c(x)}{1 - \beta\int_{-\infty}^{p_{\mathrm{M}}(x)} f(p)dp}\end{aligned}$$

so applying the condition in the statement of the proposition yields

$$\beta \frac{d}{dx}\bar{V}_{\mathrm{M}}(x) - c(x) \leq \frac{\beta\int_{p_{\mathrm{M}}(x)}^{p_{\mathrm{S}}(x)} pf(p)dp}{1 - \beta\int_{-\infty}^{p_{\mathrm{M}}(x)} f(p)dp}. \tag{3.7}$$

Now

$$\frac{\beta\int_{p_{\mathrm{M}}(x)}^{p_{\mathrm{S}}(x)} pf(p)dp}{1 - \beta\int_{-\infty}^{p_{\mathrm{M}}(x)} f(p)dp}\int_{p_{\mathrm{M}}(x)}^{p_{\mathrm{S}}(x)} f(p)dp \leq \int_{p_{\mathrm{M}}(x)}^{p_{\mathrm{S}}(x)} pf(p)dp$$

since we can cancel $\int_{p_{\mathrm{M}}(x)}^{p_{\mathrm{S}}(x)} pf(p)dp$ and then rearrange to give $\beta\int_{-\infty}^{p_{\mathrm{S}}(x)} f(p)dp \leq 1$. Thus (3.7) yields

$$\left(\beta \frac{d}{dx}\bar{V}_{\mathrm{M}}(x) - c(x)\right)\int_{p_{\mathrm{M}}(x)}^{p_{\mathrm{S}}(x)} f(p)dp \leq \int_{p_{\mathrm{M}}(x)}^{p_{\mathrm{S}}(x)} pf(p)dp, \tag{3.8}$$

whereby

$$\frac{d}{dx}\bar{V}_{\mathrm{M}}(x) \geq \frac{d}{dx}\bar{D}_1(x),$$

as required. This implies that $\bar{V}_{\mathrm{M}}(x) \geq \bar{D}_1(x)$ for all $x \in [0, x_0]$. Lemma 3.6 then implies that $\bar{V}_{\mathrm{M}}(x_0) \geq \bar{V}_{\mathrm{S}}(x_0)$, as required. ∎

Recall the condition of Proposition 3.7: $c(x) \geq \beta\int_{p_{\mathrm{S}}(x)}^{\infty} pf(p)dp$ for all $x \in [0, x_0]$. This requires $c(0) > 0$. The critical price $p_{\mathrm{S}}(x)$ is strictly increasing in the maximum sampled price q_N in S. The term $\int_{p_{\mathrm{S}}(x)}^{\infty} pf(p)dp$ is then strictly decreasing in q_N and eventually vanishes. When f has infinite support we will occasionally encounter a q_N that is sufficiently large for the inequality $c(x) \geq \beta\int_{p_{\mathrm{S}}(x)}^{\infty} pf(p)dp$ to hold for all $x \in [0, x_0]$, as long as $\int_{p_{\mathrm{S}}(x)}^{\infty} pf(p)dp$ is not too large. In other words we can expect to encounter samples where $\bar{V}_{\mathrm{M}}(x_0) > \bar{V}_{\mathrm{S}}(x_0)$ when f has a small amount of probability at high prices.

As an example application of Proposition 3.7, suppose that $\beta = 0.95$, $C(x) = \frac{1}{2}x^2 + \frac{1}{2}x$, $x_0 = 1$, and $P \sim \mathrm{LogNormal}\left(\mu = -\frac{1}{2}, \sigma^2 = 1\right)$ with probability density f. Let $N = 2$ with $q_1 = \frac{1}{2}$ and $q_2 = 3$. Numerically evaluating $c(x) - \beta\int_{p_{\mathrm{S}}(x)}^{\infty} pf(p)dp$ for $x \in [0, 1]$, Figure 1 shows that this difference is always positive which means that the condition of Proposition 3.7 is satisfied.

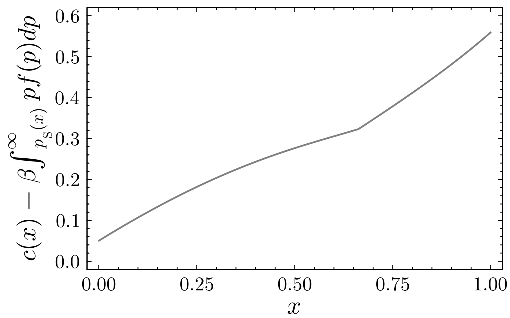

FIG. 1. *The difference* $c(x) - \beta \int_{p_S(x)}^{\infty} pf(p)dp$ *over* $x \in [0, 1]$.

It follows that the MPC policy performs at least as well as the SAA policy does for the sampled prices $q_1 = \frac{1}{2}$ and $q_2 = 3$ for the initial inventory level $x_0 = 1$. The SAA and MPC policies in question are included in Figure 2, and they differ for certain values of the sales price p.

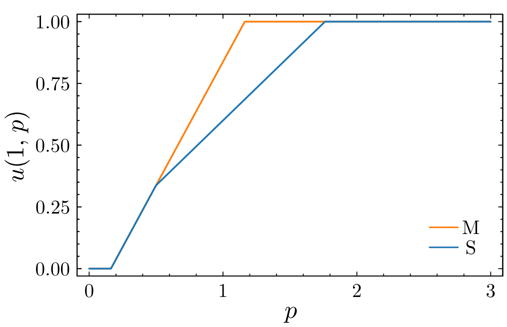

FIG. 2. *Stock sold by the SAA and MPC policies from the initial inventory level* $x_0 = 1$ *over* $p \in [0, 3]$. *Note that the stock sold is constrained to be less than* 1 *which causes the policies to coincide at* $p \approx 1.8$ *rather than* $p = q_2 = 3$.

If $c(0) = 0$, then the premise of Proposition 3.7 is not true. Despite this we present examples below which show that $\bar{V}_M(x_0) > \bar{V}_S(x_0)$ can still occur when $c(0) = 0$. These examples all involve densities having a small amount of probability at high prices.

4. Exponentially distributed prices. In this section we compare the expected out-of-sample rewards of the sample-based policies when $C(x) = \frac{1}{2}x^2$ (so $c(x) = x$) and P has an exponential density with mean 1. Here $c(0) = 0$, so Proposition 3.7 does not apply.

For $N \geq 2$, let S be a sample of size N drawn from the exponential distribution. First we consider the SAA solution to SIC using sample S when $x_0 = 1$. The result below shows that the SAA solution performs very poorly. In fact the expected out of sample value approaches $-\infty$ as β approaches 1. We will then compare this with the result if the MPC policy is used, instead of SAA.

PROPOSITION 4.1. *When* $S = \{q_1, q_2, \ldots, q_N\}$ *is a sample of size* $N \geq 2$ *drawn* *from the exponential distribution, then* $\mathbb{E}[\bar{V}_{\mathrm{S}}(1)] \to -\infty$ *as* $\beta \to 1$*, where the expectation is taken with respect to the sample* S*.*

Proof. We begin by considering fixed $q_1, q_2, \ldots, q_{N-1}$. Without loss of generality, it may be assumed that $q_1 \leq q_2 \ldots \leq q_{N-1}$. Consider first those samples where $q_N > \frac{N}{\beta} + q_{N-1}(\frac{N}{\beta} - (N-1))$. This gives a policy that, on observing price p, aims for inventory target $w_{\mathrm{S}}(p)$. If $p > q_{N-1}$, then from (2.8), $w_{\mathrm{S}}(p) = \beta\frac{(q_N - p)_+}{N} - (1-\beta)p$. Now the critical value $p_{\mathrm{S}}(x)$ occurs when $w_{\mathrm{S}}(p) = x$ and so $p_{\mathrm{S}}(x) = \frac{q_N\beta - Nx}{N+\beta-N\beta}$. We are considering values of q_N large enough so that $p_{\mathrm{S}}(x) \in (q_{N-1}, q_N]$ since $x \in [0,1]$.

From Lemma 3.3,

$$\begin{aligned}
\frac{d}{dx}\bar{V}_{\mathrm{S}}(x) &= \frac{\mathbb{E}[P|P \geq p_{\mathrm{S}}(x)]\mathbb{P}[P \geq p_{\mathrm{S}}(x)] - c(x)\mathbb{P}[P < p_{\mathrm{S}}(x)]}{1 - \beta\mathbb{P}[P < p_{\mathrm{S}}(x)]} \\
&= \frac{\int_{p_{\mathrm{S}}(x)}^{\infty} pe^{-p}dp - x(1 - e^{-p_{\mathrm{S}}(x)})}{1 - \beta(1 - e^{-p_{\mathrm{S}}(x)})} \\
&= \frac{e^{-p_{\mathrm{S}}(x)}(p_{\mathrm{S}}(x) + 1) - x(1 - e^{-p_{\mathrm{S}}(x)})}{1 - \beta(1 - e^{-p_{\mathrm{S}}(x)})} \\
&= \frac{p_{\mathrm{S}}(x) + x + 1 - xe^{p_{\mathrm{S}}(x)}}{(1-\beta)e^{p_{\mathrm{S}}(x)} + \beta}.
\end{aligned}$$

Since $\bar{V}_{\mathrm{S}}(1) = 0$ we deduce

$$\bar{V}_{\mathrm{S}}(1) = \bar{V}_{\mathrm{S}}(1)^+ - \bar{V}_{\mathrm{S}}(1)^-,$$

where

$$\bar{V}_{\mathrm{S}}(1)^+ = \int_0^1 \frac{p_{\mathrm{S}}(x) + x + 1}{(1-\beta)e^{p_{\mathrm{S}}(x)} + \beta}dx > 0$$

and

$$\bar{V}_{\mathrm{S}}(1)^- = \int_0^1 \frac{xe^{p_{\mathrm{S}}(x)}}{(1-\beta)e^{p_{\mathrm{S}}(x)} + \beta}dx > 0.$$

We will show that

$$\lim_{\beta\to 1}\int_{\frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1))}^{\infty} \bar{V}_{\mathrm{S}}(1)e^{-q_N}dq_N = -\infty. \tag{4.1}$$

First we show that $\bar{V}_{\mathrm{S}}(1)^+$ is bounded for all $\beta \in (0.5, 1)$. We have $p_{\mathrm{S}}(x) \in [q_{N-1}, q_N]$, so $p_{\mathrm{S}}(x) + x + 1$ is bounded. If $\beta > 0.5$, then $(1-\beta)e^{p_{\mathrm{S}}(x)} + \beta$ is bounded away from 0, which shows that $\bar{V}_{\mathrm{S}}(1)^+$ is bounded for all $\beta \in (0.5, 1)$. Thus the component of the integral in (4.1) from $\bar{V}_{\mathrm{S}}(1)^+$ is bounded.

Now assume that there is some M such that for all $\beta \in (0.5, 1)$ we have

$$\int_{\frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1))}^{\infty} \bar{V}_{\mathrm{S}}(1)^- e^{-q_N}dq_N < M, \tag{4.2}$$

and seek a contradiction. We write

$$\begin{aligned}&\int_{\frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1))}^{\infty} \bar{V}_{\mathrm{S}}(1)^{-} e^{-q_N} dq_N\\ &= \int_{\frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1))}^{\infty} \int_0^1 \frac{xe^{p_{\mathrm{S}}(x)}e^{-q_N}}{(1-\beta)e^{p_{\mathrm{S}}(x)}+\beta} dx dq_N\end{aligned}$$

Observe first that both the numerator and denominator of the integrand are positive, and if $\beta \in (0.5, 1)$ then

$$(1-\beta)e^{p_{\mathrm{S}}(x)} + \beta \leq (1-\beta)e^{q_N} + \beta.$$

Since

$$p_{\mathrm{S}}(x) = \frac{q_N\beta - Nx}{N+\beta-N\beta}$$

the numerator is

$$\begin{aligned} xe^{p_{\mathrm{S}}(x)}e^{-q_N} &= xe^{\frac{q_N\beta-Nx}{N+\beta-N\beta}}e^{-q_N}\\ &= xe^{-N\frac{q_N(1-\beta)+x}{N+\beta-N\beta}}\\ &= e^{-N\frac{q_N(1-\beta)}{N+\beta-N\beta}}xe^{-\frac{Nx}{N+\beta-N\beta}}. \end{aligned}$$

Now for any $\beta \in (0,1)$, since $N+\beta-N\beta > 1$, we have

$$\begin{aligned}\int_0^1 xe^{-\frac{Nx}{N+\beta-N\beta}}dx &\geq \int_0^1 xe^{-Nx}dx\\ &= \frac{1}{N^2}\left(1-(N+1)e^{-N}\right),\end{aligned}$$

so

$$\begin{aligned}&\int_{\frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1))}^{\infty} \int_0^1 \frac{e^{-q_N}xe^{p_{\mathrm{S}}(x)}}{(1-\beta)e^{p_{\mathrm{S}}(x)}+\beta} dx dq_N\\ &\geq \frac{1}{N^2}\left(1-(N+1)e^{-N}\right)\int_{\frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1))}^{\infty} \frac{e^{-N\frac{q_N(1-\beta)}{N+\beta-N\beta}}}{(1-\beta)e^{q_N}+\beta}dq_N\end{aligned}$$

For all $q > 0$ we have

$$\begin{aligned}&\frac{\partial}{\partial q}\left(\frac{e^{-N\frac{q(1-\beta)}{N+\beta-N\beta}}}{(1-\beta)e^q+\beta}\right)\\ &= -e^{-Nq\frac{(1-\beta)}{N+\beta-N\beta}}\frac{1-\beta}{(N+\beta-N\beta)\left(\beta+e^q-\beta e^q\right)^2}\left(N\beta+\beta e^q+2Ne^q(1-\beta)\right)\end{aligned}$$

which is negative, so the integrand is decreasing. Moreover for any $q > \frac{N}{\beta}+q_{N-1}(\frac{N}{\beta} - (N-1))$, $\lim_{\beta\to 1}\frac{e^{-N\frac{q(1-\beta)}{N+\beta-N\beta}}}{(1-\beta)e^q+\beta} = 1$, so there is some $\beta < 1$ with

$$\frac{e^{-N\frac{q(1-\beta)}{N+\beta-N\beta}}}{(1-\beta)e^q+\beta} > \frac{1}{2}.$$

It follows that for any such q we can find $\beta < 1$ so that

$$\int_{\frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1))}^{q} \frac{e^{-N\frac{q(1-\beta)}{N+\beta-N\beta}}}{(1-\beta)e^{q}+\beta} dq_N > \frac{1}{2}(q-(\frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1)))).$$

By choosing q large enough we can make

$$\frac{1}{N^2}\left(1-(N+1)e^{-N}\right)\int_{\frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1))}^{\infty} \frac{e^{-N\frac{q(1-\beta)}{N+\beta-N\beta}}}{(1-\beta)e^{q}+\beta} dq_N > M$$

contradicting (4.2).

Now for all q_N in the range $[0, \frac{N}{\beta}+q_{N-1}(\frac{N}{\beta}-(N-1))]$ it is easy to show that $\bar{V}_{\mathrm{S}}(1)$ is bounded for all $\beta \in (0,1)$. It follows for every fixed $q_1, q_2, \ldots, q_{N-1}$ that

$$\int_0^{\infty} \bar{V}_{\mathrm{S}}(1)e^{-q_N} dq_N$$

is unbounded below as $\beta \to 1$.

This statement is true independent of the values of $q_1, q_2, \ldots, q_{N-1}$. So if we take an expectation with respect to the (joint exponential) sampling distribution on $q_1, q_2, \ldots, q_{N-1}$ then this will also be unbounded below as $\beta \to 1$. Thus the out-of-sample losses incurred by the sample average approximation solution u_{S} are unbounded as $\beta \to 1$, regardless of the choice of N. ∎

In contrast to the SAA result, the expected value of the out-of-sample cost for the MPC policy is bounded as $\beta \to 1$. For simplicity we demonstrate this in the case $N = 2$, although it can be shown to hold in general. The expected value of the out-of-sample cost for the MPC policy is

$$\int_0^{\infty}\left(\int_0^{\infty} \bar{V}_{\mathrm{M}}(1)e^{-q_2} dq_2\right) e^{-q_1} dq_1. \tag{4.3}$$

where Lemma 3.3 gives

$$\bar{V}_{\mathrm{M}}(1) = \int_0^1 \frac{p_{\mathrm{M}}(x)+x+1-xe^{p_{\mathrm{M}}(x)}}{(1-\beta)e^{p_{\mathrm{M}}(x)}+\beta} dx.$$

The negative part of $\bar{V}_{\mathrm{M}}(1)$ is

$$\bar{V}_{\mathrm{M}}(1)^- = \int_0^1 \frac{xe^{p_{\mathrm{M}}(x)}}{(1-\beta)e^{p_{\mathrm{M}}(x)}+\beta} dx.$$

Let $\overline{q} = \frac{1}{2}(q_1+q_2)$. Recall that $p_{\mathrm{M}}(x) = (\beta\overline{q}-x)_+$, so

$$\begin{aligned}
\bar{V}_{\mathrm{M}}(1)^- &= \int_0^{\min\{\beta\overline{q},1\}} \frac{xe^{\beta\overline{q}-x}}{(1-\beta)e^{\beta\overline{q}-x}+\beta} dx + \int_{\min\{\beta\overline{q},1\}}^1 x dx \\
&\le \int_0^{\min\{\beta\overline{q},1\}} \frac{xe^{\beta\overline{q}-x}}{\beta} dx + \int_0^1 x dx \\
&\le \frac{e^{\overline{q}}}{e\beta} + \frac{1}{2}.
\end{aligned}$$

Therefore

$$\int_0^\infty \left(\int_0^\infty \bar{V}_{\mathrm{M}}(1)^- e^{-q_2} dq_2\right) e^{-q_1} dq_1 \leq \frac{1}{e\beta} \int_0^\infty \left(\int_0^\infty e^{\frac{1}{2}(q_1+q_2)} e^{-q_2} dq_2\right) e^{-q_1} dq_1 + \frac{1}{2}$$
$$= \frac{4}{e\beta} + \frac{1}{2}.$$

Thus, as long as $\beta \in (0,1)$ is bounded away from 0, we have

$$\int_0^\infty \left(\int_0^\infty \bar{V}_{\mathrm{M}}(1)^- e^{-q_2} dq_2\right) e^{-q_1} dq_1 < \infty$$

so

$$\int_0^\infty \left(\int_0^\infty \bar{V}_{\mathrm{M}}(1) e^{-q_2} dq_2\right) e^{-q_1} dq_1 > -\infty.$$

Moreover, identical reasoning as in the SAA case shows that (4.3) has a finite-valued positive part. Thus, when $N = 2$, the expected out-of-sample loss incurred under the MPC policy is bounded as $\beta \to 1$.

5. Numerical studies. In this section we use numerical simulation to study the performance of the two sample-based policies (SAA and MPC) on different price distributions. In section 4 we showed that MPC is far better then SAA with an exponential distribution. But this is an exception—we do not usually find this extreme behaviour with the two expected out-of-sample values differing by an amount that is unbounded as $\beta \to 1$. However this case does suggest that the amount of skew in the underlying distribution is important, and we will explore this in this section.

To compute the expected out-of-sample performance of the sample-based policies under the sampling distribution of $q_1, q_2, \ldots, q_N$, we use a simulation coded in the `Julia` programming language [2]. Although the true problem has an infinite number of stages, simulation with a finite number of stages (say T) will give a realistic estimate as long as it is sufficiently large. We set $T = 1000$ and efficiently simulate the repeated sales process by terminating any instances as soon as the inventory level reaches 0. Setting $\beta = 0.95$, $x_0 = 1$ and $C(x) = \frac{1}{2}x^2$, for each policy we:

1. Sample N random prices from $\mathbb{P}$ to construct $q_1, q_2, \ldots, q_N$ which then determines the sample-based policy u (either SAA or MPC).
2. Sample a random price p_t from $\mathbb{P}$, accrue reward $\beta^{t-1}(p_t u(x_{t-1}, p_t) - C(x_{t-1} - u(x_{t-1}, p_t)))$ and set $x_t = x_{t-1} - u(x_{t-1}, p_t)$.
3. Repeat Step 2 from stage $t = 1$ to $T-1$ and sell any remaining stock at stage T to generate $\sum_{t=1}^{T} \beta^{t-1}(p_t u(x_{t-1}, p_t) - C(x_{t-1} - u(x_{t-1}, p_t)))$.

We repeat Steps 1 through 3 to generate realisations for use as an estimate of the expected value of the SIC problem when a policy u is used out-of-sample. In our experiments we used 50000 realisations to generate the estimate of the expected value and found that this was sufficient to achieve accurate values. In Figures 3-5 and 7 the standard error ranges are smaller than the markers and so are not shown. Also note that for $N = 1$ the two sample-based policies coincide.

5.1. Triangularly distributed prices. Suppose $P \sim \mathrm{Triangular}(a, m, b)$, with lower limit a, mode m, and upper limit b. This is not a particularly realistic distribution but serves to illustrate the effect of skew on the performance of SAA and MPC. In what follows we select a, m, and b such that $\mathbb{E}[P] = 1$ and $\mathrm{Var}[P] = \frac{1}{8}$; the

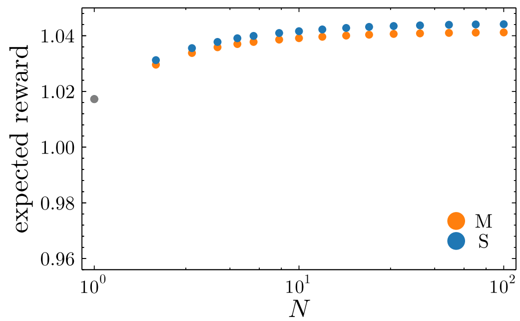

FIG. 3. *Expected out-of-sample reward of SAA and MPC for $P \sim$ Triangular$(0, \frac{3}{2}, \frac{3}{2})$, a left-skewed distribution.*

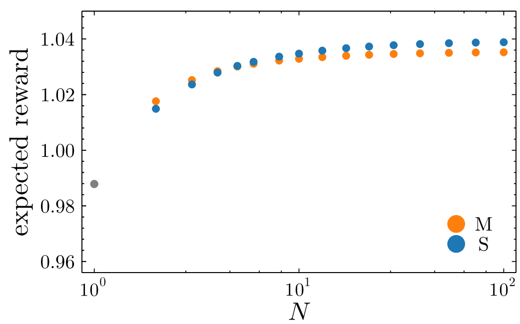

FIG. 4. *Expected out-of-sample reward of SAA and MPC for $P \sim$ Triangular$(1 - \frac{1}{2}\sqrt{3}, 1, 1 + \frac{1}{2}\sqrt{3})$, a symmetric distribution.*

intention being to confine differences between SAA and MPC to the sampling effects of skew only and compare them on different distributions as fairly as possible.

Figure 3 shows SAA outperforming MPC for all N on a price distribution that is triangular and left-skewed. This is in contrast to Figure 4, which shows MPC outperforming SAA for $N \leq 5$ on a price distribution that is triangular and symmetric. Replacing the left-skewed price distribution that yields Figure 3 with a symmetric distribution increases the value of b. Samples with high prices then cause the SAA policy to under-sell and pay too much in storage costs. The MPC policy attenuates this effect since $u_{\mathrm{M}} \geq u_{\mathrm{S}}$.

Further increasing c to 2 increases the range where MPC outperforms SAA, as can be seen in Figure 5, which shows MPC outperforming SAA for $N \leq 6$ on a price distribution that is triangular and right-skewed.

5.2. Log-normally distributed prices. Suppose that $P \sim \text{LogNormal}\left(\mu, \sigma^2\right)$, with mean μ and variance σ^2. Log-Normal distributions are often used to model prices in financial applications and have a significant right-skew (see e.g. Figure 6).

Figure 7 shows MPC outperforming SAA for all N less than about 50, a significantly larger range than that in Figure 5. The significant right-skew of the Log-Normal distribution increases the propensity for a single very large price sample to be included in $q_1, q_2, \ldots, q_N$ which degrades the quality of the approximate price dis-

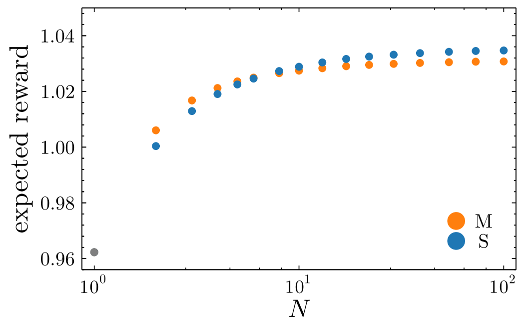

FIG. 5. *Expected out-of-sample reward of SAA and MPC for* $P \sim$ *Triangular*$(\frac{1}{2}, \frac{1}{2}, 2)$*, a right-skewed distribution.*

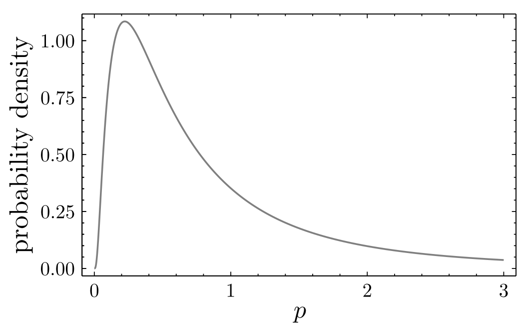

FIG. 6. *Probability density of* $p \in [0, 3]$ *for* $P \sim$ *LogNormal*$\left(-\frac{1}{2}, 1\right)$.

tribution informing the SAA policy. Figure 8 demonstrates this explicitly in the case where $N = 2$; typical price samples result in the SAA policy outperforming the MPC policy, but for a small proportion of more extreme samples, where one of the samples is very large, the reverse occurs and the MPC policy significantly outperforms the SAA policy.

6. A distributionally robust interpretation of MPC. Proposition 3.7 and the examples in sections 4 and 5 show that the lower target inventory of the MPC policy can be beneficial as it reduces sensitivity to large price samples. In the following section we show that this effect can be seen as an example of *distributional robustness*.

Distributionally robust optimisation (DRO) is an approach to stochastic optimization that intends to protect decision-makers from ambiguity in the specification of the underlying probability distributions. DRO problems optimise against the worst case element of an *ambiguity set*, in which the true distribution is believed to lie. By considering the worst cases, distributionally robust estimates are usually less sensitive to outliers and in some cases give better out-of-sample expected performance [1].

The seminal work [7] specified an ambiguity set by requiring its elements have certain first and second moments. We will show that the MPC optimization problem is equivalent to a multistage DRO problem with an ambiguity set specified by the first moment of the empirical price distribution.

Let $\mathcal{P}(\mathbb{R})$ denote the set of possible probability distributions on the real line

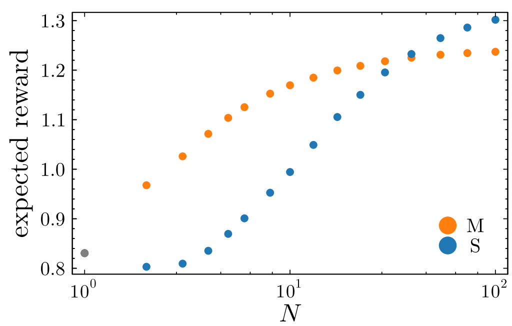

FIG. 7. *Expected out-of-sample reward of SAA and MPC for $P \sim LogNormal\left(-\frac{1}{2}, 1\right)$. Note* $\mathbb{E}[P] = 1$.

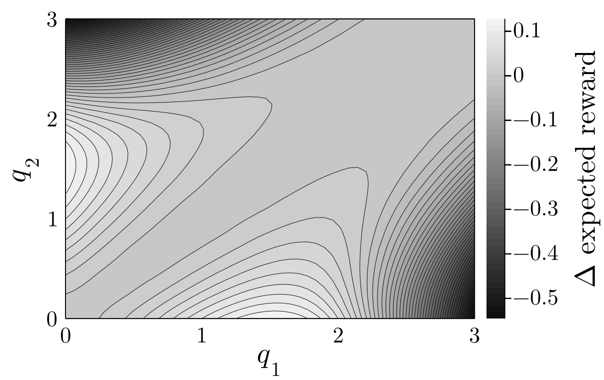

FIG. 8. *Expected out-of-sample reward of SAA minus that of MPC as a function of q_1 and q_2 over $[0,3] \times [0,3]$ for $P \sim LogNormal\left(-\frac{1}{2}, 1\right)$. Darker contours indicate regions where the MPC policy outperforms the SAA policy and vice versa. The contour that the right diagonal lies in is at elevation 0 since the SAA and MPC policies are identical when $q_1 = q_2$.*

for a random variable P. For some probability distribution μ, define $\mathcal{M}_1(\mu) := \{\nu \in \mathcal{P}(\mathbb{R}) : \mathbb{E}_\nu[P] = \mathbb{E}_\mu[P]\}$, this being the set of probability distributions having the same first moment as μ. Now define the distributionally robust functional equation

$$V_{\mathrm{R}}(x,p) := \sup_{0 \le u \le x} \left\{ pu - C(x-u) + \beta \inf_{\nu \in \mathcal{M}_1(\mu)} \mathbb{E}_\nu[V_{\mathrm{R}}(x-u, P)] \right\}. \tag{6.1}$$

(We defer showing that a function satisfying (6.1) actually exists until the proof of Proposition 6.1.) The distributionally robust functional equation (6.1) selects the worst-case distribution in $\mathcal{M}_1(\mu)$ for each candidate policy u. This process propagates through the definition of the functional equation, such that the resulting optimal policy is protected against the worst case distribution in the current stage and the worst case distributions in all future stages, simultaneously. Although this is inconsistent with the modeling assumption that the price distribution at each stage is independent and identically distributed, in this case the worst case distribution is unique, and we have the following result.

PROPOSITION 6.1. *The solution $V_{\mathrm{M}}(x,p)$ to the MPC recursion*

$$V_{\mathrm{M}}(x,p) = \max_{0\le u\le x} \{pu - C\,(x-u) + \beta V_{\mathrm{M}}(x-u, \mathbb{E}_{\mu}[P])\} \tag{6.2}$$

is the unique solution to (6.1)*.*

Proof. For any V_{R} satisfying (6.1) and any $\nu \in \mathcal{M}_1(\mu)$ it follows that

$$\begin{aligned}
&\mathbb{E}_{\nu}[V_{\mathrm{R}}(x,P)]\\
&= \mathbb{E}_{\nu}\left[\sup_{0\le u\le x}\left\{Pu - C(x-u) + \beta \inf_{\nu'\in\mathcal{M}_1(\mu)} \mathbb{E}_{\nu'}[V_{\mathrm{R}}(x-u,P')]\right\}\right]\\
&\ge \sup_{0\le u\le x}\left\{\mathbb{E}_{\nu}\left[Pu - C(x-u) + \beta \inf_{\nu'\in\mathcal{M}_1(\mu)} \mathbb{E}_{\nu'}[V_{\mathrm{R}}(x-u,P')]\right]\right\}\\
&= \sup_{0\le u\le x}\left\{\mathbb{E}_{\mu}[P]u - C(x-u) + \beta \inf_{\nu'\in\mathcal{M}_1(\mu)} \mathbb{E}_{\nu'}[V_{\mathrm{R}}(x-u,P')]\right\}\\
&= V_{\mathrm{R}}(x,\mathbb{E}_{\mu}[P]).
\end{aligned}$$

where the second equality follows since $\mathbb{E}_{\nu}[P] = \mathbb{E}_{\mu}[P]$.

But the probability distribution with all of its mass at $\mathbb{E}_{\mu}[P]$ is in $\mathcal{M}_1(\mu)$, which means that $\inf_{\nu\in\mathcal{M}_1(\mu)} \mathbb{E}_{\nu}[V_{\mathrm{R}}(x,P)] = V_{\mathrm{R}}(x,\mathbb{E}_{\mu}[P])$, and so

$$\beta \inf_{\nu\in\mathcal{M}_1(\mu)} \mathbb{E}_{\nu}[V_{\mathrm{R}}(x-u,P)] = \beta V_{\mathrm{R}}(x-u,\mathbb{E}_{\mu}[P]).$$

This shows that (6.1) is equivalent to the recursion

$$V_{\mathrm{R}}(x,p) = \sup_{0\le u\le x} \{pu - C(x-u) + \beta V_{\mathrm{R}}(x-u,\mathbb{E}_{\mu}[P])\}$$

which has solution $V_{\mathrm{M}}(x,p)$. Lastly, we know that V_{M} exists by Theorem 9.2 of [11, p. 246], concluding the proof. □

When μ is the empirical distribution on the samples $q_i, q_2, \ldots, q_N$, Proposition 6.1 shows that the MPC policy u_{M} is distributionally robust. This can be helpful as a lens for understanding MPC: when viewed as distributionally robust we expect to see a a shrinkage effect, which occurs here because $u_{\mathrm{M}} \ge u_{\mathrm{S}}$. This can yield an improvement in out-of-sample expected reward when variance reduction outweighs any increase in bias.

7. Conclusions. We studied the performance of SAA and MPC on a multistage stochastic inventory control problem, finding that MPC can outperform SAA when the underlying price distribution is right-skewed and N is not too large. In the case where the underlying price distribution is exponential, MPC can outperform SAA regardless of the size of N. The good performance of MPC can be explained by viewing it through the lens of a distributional robustification, challenging the assumption that stochastic dynamic programming is always the right solution approach.

Acknowledgments. We would like to acknowledge the helpful suggestions of Andrew J. Mason.

REFERENCES

[1] E. Anderson and A. Philpott, *Improving sample average approximation using distributional robustness*, INFORMS Journal on Optimization, 4(1) (2022), pp. 90–124.

[2] J. Bezanson, A. Edelman, S. Karpinski, and V. Shah, *Julia: A fresh approach to numerical computing*, SIAM Review, 59 (2017), pp. 65–98.

[3] M. Dyer and L. Stougie, *Computational complexity of stochastic programming problems*, Mathematical Programming, 106 (2006), pp. 423–.

[4] T. Martin, *Stochastic optimization for the procurement of crude oil in refineries*, PhD thesis, École des Ponts ParisTech, 2021.

[5] D. Mayne, *Model predictive control: Recent developments and future promise*, Automatica, 50 (2014), pp. 2967–2986.

[6] A. Mesbah, *Stochastic model predictive control: An overview and perspectives for future research*, IEEE control systems, 36 (2016), pp. 30–44.

[7] H. Scarf, *A min-max solution of an inventory problem*, Studies in the Mathematical Theory of Inventory and Production, (1958), pp. 201—-209.

[8] A. Shapiro, *On complexity of multistage stochastic programs*, Operations Research Letters, 34 (2006), pp. 1–8.

[9] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 3rd ed., 2021.

[10] A. Shapiro and A. Nemirovski, *On complexity of stochastic programming problems*, in Continuous Optimization: Current Trends and Modern Applications, V. Jeyakumar and A. Rubinov, eds., Springer US, Boston, MA, 2005, pp. 111–146.

[11] N. Stokey, R. Lucas, and E. Prescott, *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, MA, 1989.

[12] H. Theil, *A note on certainty equivalence in dynamic planning*, Econometrica, 25 (1957), pp. 346–349.

[13] W. Ziemba, *Transforming stochastic dynamic programming problems into nonlinear programs*, Management Science, 17 (1971), pp. 450–462.