Model predictive control and distributionally robust stochastic dynamic programming

D. S. T. Keehan*
University of Auckland

A. B. Philpott
University of Auckland

E. J. Anderson
University of Sydney and Imperial College London

Abstract

Stochastic dynamic programming problems are difficult to solve to optimality. Model predictive control (where a simpler deterministic problem is solved as an alternative) has enjoyed a wide range of application for this reason. We study an instance of model predictive control where each random value is held at its expectation. Though surprising, this approach yields policies that perform competitively in some applications. For example, Fonterra Co-operative Group Limited utilise this control policy to govern dairy product sales driven by periodic price fluctuations. Even more surprisingly, when solving stochastic dynamic programs using data-driven approximations of the true random dynamics, model predictive control can perform better out-of-sample. We study this phenomenon in an idealised version of Fonterra's application. We observe that model predictive control can be interpreted as distributionally robust stochastic dynamic programming. After demonstrating out-of-sample value in its own right, we further hypothesise that this distributional robustness explains the strong performance of model predictive control and support this with numerical studies.

1 Introduction

Recursive decision problems of the form

$$V(x,p) = \min_{u \in U} \{F(u,x,p) + \mathbb{E}[V(T(u,x),q)]\} \tag{1}$$

are of fundamental importance to planners because of their ability to capture how immediate decisions affect future utility in the face of uncertainty. This problem can be solved by Stochastic Dynamic Programming (SDP) (Ross, 2014). In general, the recursion (1) is difficult to solve, even approximately. A simpler alternative is to solve a deterministic optimisation problem that models

*dkee331@aucklanduni.ac.nz

the randomness involved, without accounting for the full random dynamics. This is termed Model Predictive Control (MPC) (Kouvaritakis and Cannon, 2016)

One model predictive alternative is to hold each random value at its expectation. The recursion would become

$$V(x, p) = \min_{u \in U} \{F(u, x, p) + V(T(u, x), \mathbb{E}[q])\}, \tag{2}$$

so that the optimisation is deterministic in the sense that V is evaluated only at a single value. In some cases recursions have simplified dynamics and are *certainty equivalent*, so that the optimal u from (2) solves (1) (Theil, 1957) and is therefore an optimal policy for the immediate decision to be made.

1.1 Our idealised model

Yielding sensible policies from MPC is not relegated to problems that strictly exhibit certainty equivalence. Many complicated problems faced in industry utilise this simplified control law to a surprisingly pleasant conclusion (see (Camacho and Alba, 2013) for some examples). Of particular interest is Fonterra Co-operative Group Limited's application, where MPC is used to plan how much raw milk to produce and refine into various products, before deciding when to export these products globally. In an internal technical report Guan (2019) demonstrated that MPC may be on par with SDP out-of-sample for this task. Fonterra's setting has the general features:

- Goods that can be immediately bought and sold for a known price.
- Good prices set by an international market that fluctuate in a random way.
- Known storage costs that are required to maintain goods over time.

We can represent this by the following recursive optimisation problem:

$$V(x, p) = \min_{u \in U} \{-pu + C(x - u) + \mathbb{E}[V(x - u, q)]\} \tag{3}$$

where x is the current stock of goods, p is the immediate unit stock price, u is the amount of stock sold (or purchased if $u < 0$), $C(x)$ is the cost of storage of stock amount level x, and the expectation is taken with respect to the underlying price distribution. We assume C is strictly convex and has derivative c, where $C(0) = c(0) = 0$. Since C is strictly convex, c is strictly increasing and has an inverse, which we denote by c^{-1}.

Our idealised problem is an inventory management problem. There is a wealth of inventory management literature surrounding planning problems where demand is random. Typically these problems feature shortage and surplus costs (see for example (Bellman, Glicksberg, and Gross, 1955)). The situation with random prices has not been addressed so thoroughly, and is distinguished from the afforementioned case by being less concerned with satisfying consumers and more with identifying abnormally low prices. Therefore, as this paper progresses we will identify analytic expressions for the minimiser of (3).

The contributions of this paper are as follows: Firstly, we provide a brief outline of the sample average approximation method and distributional robustness. Secondly, we solve recursion (3) when stock purchasing is allowed and show that it is certainty equivalent, to then show that distributional robustness can decrease costs out-of-sample. Lastly, we solve recursion (3) when stock purchasing is disallowed to then show that MPC produces the same policy as a distributional robustification, and examine how this explains the performance of MPC in various simulations.

2 Supporting theory

It is rare that the complete definition of the distribution of a random variable required to evaluate a general expectation $\mathbb{E}[F(u,\xi)]$ is available. Instead, one may have independent samples $S = \{\xi_1, \xi_2, \dots, \xi_N\}$ of the random variable ξ. These comprise an empirical measure $\mathbb{P}_0$ of ξ, which we can use to form a Sample Average Approximation (SAA) of the expectation, so that

$$\mathbb{E}[F(u,\xi)] \approx \sum_{i=1}^{N} \frac{1}{N} F(u,\xi_i).$$

When using this approximation in place of true expectation to solve

$$\min_{u\in U} \mathbb{E}[F(u,\xi)] \approx \min_{u\in U} \sum_{i=1}^{N} \frac{1}{N} F(u,\xi_i),$$

we are left with an approximate u_0 in place of the true u^*. Under quite general conditions it is known that as $N \to \infty$, $u_0 \to u^*$ (Shapiro, Dentcheva, and Ruszczynski, 2021)—of course, an infinite amount of observations is just as unknowable as a full definition of the distribution. One must contend with randomness in any N samples to yield an out-of-sample cost $\mathbb{E}_S[\mathbb{E}[F(u_0,\xi)]]$.

While it is obviously the case that $\mathbb{E}[F(u^*,\xi)] \leq \mathbb{E}_S[\mathbb{E}[F(u_0,\xi)]$, there are approaches to reducing this gap. One approach is to perform a Distributionally Robust Optimisation (DRO) (Shapiro, 2017), by considering problems of the form

$$\min_{u\in U} \max_{\mu\in\mathbb{P}_\theta} \mathbb{E}_\mu[F(u,\xi)].$$

If one forms a measure uncertainty set $\mathbb{P}_\theta$ in a ball of radius θ about $\mathbb{P}_0$, one can protect against realisations of S which are unrepresentative of the underlying distribution by being conservative against other measures "close" to $\mathbb{P}_0$. Indeed, Anderson and Philpott (2022) showed that if the underlying distribution was skewed, distributional robustness could protect against realised observations with large outliers, resulting in a better cost out-of-sample.

For a radius θ we will consider measure uncertainty sets of the form

$$\mathbb{P}_\theta(S) := \left\{ \mu \middle| \mu(\zeta_i) = \frac{1}{N}, |\zeta_i - \xi_i| \leq \theta, i \in \{1, 2, \dots, N\} \right\}, \tag{4}$$

so that sample ξ_i may vary between $\xi_i - \theta$ and $\xi_i + \theta$. The form of $\mathbb{P}_\theta$ can be quite important, however the optimal choice of $\mathbb{P}_\theta$ is not the focus of this paper, and (4) is simple enough to demonstrate the general idea.

3 Purchasing allowed

Suppose that the feasible region in (3) is $U := \{u : u \leq x\}$, so that we cannot sell more stock than we initially have, but we can buy as much as we like. The value function is then

$$V(x,p) = \min_{u\leq x} \left\{ -pu + C(x-u) + \mathbb{E}[V(x-u,q)] \right\}. \tag{5}$$

Theorem 1. *The optimal u defined by (5) is*

$$u(x,p) = x - c^{-1}((\mathbb{E}[q] - p)_+).$$

Proof. Expanding (5) using the recursive definition yields $V(x,p)$ as

$$\min_{u_1 \le x} \left\{ -pu_1 + C(x-u_1) + \mathbb{E}\left[\min_{u_2 \le x-u_1} \left\{ -qu_2 + C(x-u_1-u_2) + \mathbb{E}[V(x-u_1-u_2,q)] \right\} \right] \right\}.$$

Introduce the substitutions $w_1 = x - u_1$ and $w_2 = x - u_1 - u_2$. We then have

$$\begin{aligned} V(x,p) &= \min_{0 \le w_1} \left\{ -px + pw_1 + C(w_1) + \mathbb{E}\left[\min_{0 \le w_2} \left\{ -qw_1 + qw_2 + C(w_2) + \mathbb{E}[V(w_2,q)] \right\} \right] \right\} \\ &= -px + \min_{0 \le w_1} \left\{ pw_1 + C(w_1) - \mathbb{E}[q]w_1 \right\} + \mathbb{E}\left[\min_{0 \le w_2} \left\{ qw_2 + C(w_2) + \mathbb{E}[V(w_2,q)] \right\} \right]. \end{aligned}$$

We are only concerned with the terms involving w_1 since this completely defines our immediate decision. The optimal solution to $\min_{0 \le w_1} \{pw_1 + C(w_1) - \mathbb{E}[q]w_1\}$ is $w_1 = c^{-1}((\mathbb{E}[q]-p)_+)$, so that $u(x,p) = x - c^{-1}((\mathbb{E}[q]-p)_+)$. □

Corollary 1. *Since the expectation in $u(x,p) = x - c^{-1}((\mathbb{E}[q]-p)_+)$ is of a linear function, the recursion (5) is certainty equivalent.*

The substitution utilised in the proof of Theorem 1 is illuminating. The value $w = x - u$ represents the optimal target inventory, at which point the marginal increase in storage cost is equal to the expected increase in price when waiting for the next period. Although surprising, the potential for recursions to be decoupled in this way has been known since at least (Ziemba, 1971). What this shows is that the nature of the problem is in actuality single stage, and since the randomness is linear, certainty equivalence is assured. Having identified that recursion (5) is certainty equivalent, one may take this as a hint that certainty equivalent policies in Fonterra's application will be competitive.

3.1 Bias and variance reduction via robustification

When $C(x-u) = \frac{1}{2}(x-u)^2$, we have the following observation

$$w^*(x,p) = (\mathbb{E}[q]-p)_+ = \left(\mathbb{E}_S\left[\sum_{i=1}^{N} \frac{1}{N} q_i \right] - p \right)_+ \le \mathbb{E}_S\left[\left(\sum_{i=1}^{N} \frac{1}{N} q_i - p \right)_+ \right] = \mathbb{E}_S[w_0(p)] \qquad (6)$$

where the inequality follows by Jensen's inequality. Since the optimal SDP solution $w_0(x,p)$ to (5) is biased when SAA is used, there is potential to beat it out-of-sample. If we can show that DRO reduces this bias, it will be very easy to show improvement out-of-sample.

Consider the distributionally robust version of (5),

$$V(x,p) = \min_{u \le x} \max_{\mu \in \mathbb{P}_\theta} \left\{ -pu + C(x-u) + \mathbb{E}_\mu[V(x-u,q)] \right\}.$$

Here we are not able to decouple the recursion since the choice of μ affects the immediate target inventory and the future target inventory. That being said, all we care about out-of-sample is the objective $\min_{0 \le w} \{pw + C(w) - \mathbb{E}[q]w\}$ since it is all that will ever depend on our current decision. One could instead formulate the distributionally robust problem:

$$\min_{0 \le w} \max_{\mu \in \mathbb{P}_\theta} \left\{ pw + C(w) - \mathbb{E}_\mu[q]w \right\}. \qquad (7)$$

Although at first this may seem dissatisfying, the removal of terms that do not depend on the decision variable means that we only consider robustness from the perspective of actions we can control. This situation was discussed in Anderson and Philpott, 2022, and is not the main focus of the current paper.

Proposition 1. *Suppose $\mathbb{P}_\theta$ is specified according to (4). Then the solution to (7) is*

$$u_\theta(x,p) = x - c^{-1}((\mathbb{E}_{\mathbb{P}_0}[q-\theta] - p)_+).$$

Proof. Consider the inner optimisation $\max_{\mu \in \mathbb{P}_\theta}\{pw + C(w) - \mathbb{E}_\mu[q]w\}$. Since $w \geq 0$, the term $-\mathbb{E}_\mu[q]w$ is the opposite sign to $\mathbb{E}_\mu[q]$ and we wish to make each q_i as small as possible. Hence, all q_i should be reduced to $q_i - \theta$ and μ^* is such that $\mathbb{E}_{\mu^*}[q] = \mathbb{E}_{\mathbb{P}_0}[q-\theta]$. Applying Theorem 1 then proves the theorem. □

Proposition 1 means that for small enough θ, the SAA solution bias will be reduced since $u_\theta(x,p) \geq u_0(x,p)$. We will now briefly prove that estimators of this form also reduce variance:

Lemma 1. *Suppose X is a random variable on the interval (α,β) and θ is non-negative. Then $Var(Y)$ where $Y = \max(X-\theta,\alpha)$ is less than or equal to $Var(X)$.*

Proof. Noting that variance is translation invariant, without loss of generality assume X is a random variable on $(0, \beta-\alpha)$. Let $Y = X - Z$ where $Z = X - \max(X-\theta, 0)$. We wish to show

$$\mathbb{E}[(X-Z)^2] - \mathbb{E}[X-Z]^2 \leq \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

This is equivalent to

$$\begin{aligned}
\mathbb{E}[X^2 - 2XZ + Z^2] - \mathbb{E}[X]^2 + 2\mathbb{E}[X]\mathbb{E}[Z] - \mathbb{E}[Z]^2 &\leq \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
\mathbb{E}[-2XZ + Z^2] + 2\mathbb{E}[X]\mathbb{E}[Z] - \mathbb{E}[Z]^2 &\leq 0 \\
\mathbb{E}[(2X-Z)Z] - (2\mathbb{E}[X] - \mathbb{E}[Z])\mathbb{E}[Z] &\geq 0 \\
Cov((2X-Z), Z) &\geq 0 \\
Cov(X + \max(X-\theta,0), X - \max(X-\theta,0)) &\geq 0. \qquad (8)
\end{aligned}$$

Then, (8) is true since $X + \max(X-\theta,0)$ and $X - \max(X-\theta,0)$ are monotone increasing functions in X. □

A reduction in bias and variance is sufficient to lower the objective out-of-sample in the quadratic costs case, as the following theorem shows:

Theorem 2. *Suppose $C(x-u) = \frac{1}{2}(x-u)^2$. Then there exists a ball of θ values such that the robust policy $u_\theta(x,p)$ has a lower out-of-sample cost than $u_0(x,p)$ for (5).*

Proof. For initial stock x and price p, the out-of-sample cost defined by (5) when applying the policy $u_\theta(x,p)$ derived from samples S in a rolling-horizon manner is

$$V(x,p) = \mathbb{E}_S\left[-pu_\theta(x,p) + C(x - u_\theta(x,p)) + \mathbb{E}_q\left[V(x - u_\theta(x,p), q)\right]\right].$$

Decoupling this yields

$$V(x,p) = -px$$
$$+\mathbb{E}_S\left[pw_\theta(p) + C(w_\theta(p)) - \mathbb{E}_q\left[\mathbb{E}_S\left[qw_\theta(p)\right]\right] + \mathbb{E}_q\left[\mathbb{E}_S\left[qw_\theta(q) + C(w_\theta(q)) - \mathbb{E}_r\left[V(w_\theta(q), r)\right]\right]\right]\right].$$

It suffices to show that $\mathbb{E}_S\Big[pw_\theta(p) + C(w_\theta(p)) - \mathbb{E}_q\left[\mathbb{E}_S\left[qw_\theta(p)\right]\right]\Big]$ is reduced for all p. After noting that $\mathbb{E}_q\left[\mathbb{E}_S\left[qw_\theta(p)\right]\right] = \mathbb{E}_q[q]\mathbb{E}_S\left[w_\theta(p)\right]$, we may rewrite this expression as

$$\mathbb{E}_S\left[p\bar{w}_\theta(p) + \frac{1}{2}\left(\bar{w}_\theta(p)\right)^2 - \mathbb{E}_q[q]\bar{w}_\theta(p) + \frac{1}{2}\left((w_\theta(p))^2 - \bar{w}_\theta(p)^2\right)\right] \tag{9}$$

where $\mathbb{E}_S[w_\theta(p)] = \bar{w}_\theta(p)$. This is a sum of the bias and variance of the optimal solution. As $w_\theta(p) = (\mathbb{E}_{\mathbb{P}_0}[q-\theta] - p)_+ = \max(w_0(p) - \theta, 0)$, we know from Lemma 1 that the variance term in (9) must not increase. Furthermore, because $\bar{w}_0(p)$ is biased high (as shown in (6)), for small enough θ the robust term $\bar{w}_\theta(p) = \mathbb{E}_S[(\mathbb{E}_{\mathbb{P}_0}[q-\theta] - p)_+]$ will be less biased, and the bias term in (9) must decrease. This proves the theorem. □

The advantage of Theorem 2 is that it is true regardless of the underlying price distribution. However, the amount of improvement possible depends on the size of the ball of radii that strictly decrease bias, since a large θ may increase bias and negate cost savings from variance reduction. We demonstrate this bias-variance trade-off with the following simulated example: suppose that we have 100 stages of purchasing and selling with prices uniformly distributed between 10 and 50, and an initial stock level of 50 units. Figure 1 shows that the increase in revenue as a result of robustification can be significant for small N. As $N \to \infty$, SAAs become accurate so it is unsurprising that all robust policies are eventually beaten.

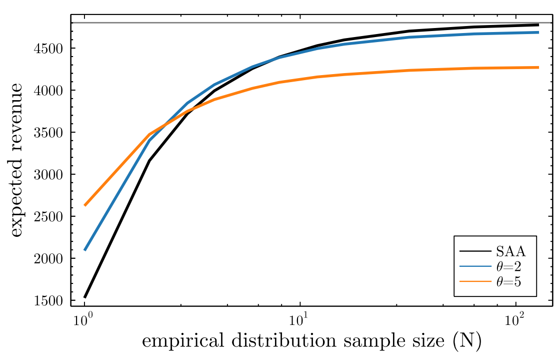


Figure 1: Expected revenue accrued over 10000 simulations as a function of sample size, for the SAA policy and robust policies from Proposition 1 with $\theta = 2, 5$. The horizontal line is the optimal expected revenue.

4 Purchasing disallowed

Suppose now that the feasible region in (3) is $U := \{u : 0 \le u \le x\}$, so that we cannot sell more stock than we initially have and we cannot purchase any stock. The value function becomes

$$V(x,p) = \min_{0\le u\le x} \{-pu + C(x-u) + \mathbb{E}[V(x-u,q)]\} . \tag{10}$$

Theorem 3. *The optimal u defined by (10) is*

$$u(x,p) = (x - c^{-1}(\mathbb{E}[(q-p)_+]))_+.$$

Proof. We will work with the sub-derivative operator to verify $u(x,p) = (x - c^{-1}(\mathbb{E}[(q-p)_+]))_+$. In the usual way we use the notation $\partial f + \partial g := \{x + y : x \in \partial f, y \in \partial g\}$. Note that $\partial(f+g) \supset \partial f + \partial g$, and in the context of SAA we have $\partial\mathbb{E}[f] \supset \mathbb{E}[\partial f]$. One can see that the recursion (10) is convex in u. Our first order condition is then

$$-p - c(x-u) + \partial\mathbb{E}[V(x-u,q)] \supset -p - c(x-u) + \mathbb{E}[\partial V(x-u,q)] \ni 0.$$

If $x \ge c^{-1}(\mathbb{E}\left[(q-p)_+\right])$ we must show $-p-\mathbb{E}\left[(q-p)_+\right]+\mathbb{E}\left[\partial V\left(c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right),q\right)\right] \ni 0$. In the region $x \ge c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right)$ the term $\partial V(x,p)$ becomes

$$\begin{aligned}&= \partial\left(-p\left(x - c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right)\right) + C\left(c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right)\right) + \mathbb{E}\left[V\left(c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right),q\right)\right]\right)\\ &\ni -p.\end{aligned} \tag{11}$$

In the region $x \le c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right)$

$$\begin{aligned}\partial V(x,p) &= \partial\left(C(x) + \mathbb{E}\left[V(x,q)\right]\right)\\ &\supset c(x) + \mathbb{E}\left[\partial V(x,q)\right].\end{aligned} \tag{12}$$

(11) means that $\partial V\left(c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right),r\right) \ni -r$ if $c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right) \ge c^{-1}\left(\mathbb{E}\left[(q-r)_+\right]\right)$, which is an equivalent condition to $r \ge p$. On the other hand, (12) means that $\partial V\left(c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right),r\right) \supset \mathbb{E}\left[(q-p)_+\right]+\mathbb{E}\left[\partial V\left(c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right),q\right)\right]$ if $r < p$. Utilising the previous two observations and letting $\int_{-\infty}^{p} 1 d\mu(q) = \rho$, we have that $\mathbb{E}\left[\partial V\left(c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right),q\right)\right]$ is

$$\begin{aligned}&\supset \int_p^\infty -q d\mu(q) + \int_{-\infty}^{p}\left(\mathbb{E}\left[(q-p)_+\right] + \mathbb{E}\left[\partial V\left(c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right),q\right)\right]\right)d\mu(q)\\ &= -\mathbb{E}\left[(q-p)_+\right] - (1-\rho)p + \rho\mathbb{E}\left[(q-p)_+\right] + \rho\mathbb{E}\left[\partial V\left(c^{-1}\left(\mathbb{E}\left[(q-p)_+\right]\right),q\right)\right]\\ &= \frac{-\mathbb{E}\left[(q-p)_+\right] - (1-\rho)p + \rho\mathbb{E}\left[(q-p)_+\right]}{1-\rho}\\ &= \mathbb{E}\left[(q-p)_+\right] + p.\end{aligned} \tag{13}$$

Subbing (13) into the first order condition yields $-p - \mathbb{E}\left[(q-p)_+\right] + p + \mathbb{E}\left[(q-p)_+\right] = 0$, as required.

If $x < c^{-1}(\mathbb{E}\left[(q-p)_+\right])$, we have $c(x) < \mathbb{E}\left[(q-p)_+\right]$ and the marginal cost of storage is less than the expected gain in price of waiting for the next period, so $u(x,p) = 0 = (x - c^{-1}(\mathbb{E}[(q-p)_+]))_+$.

□

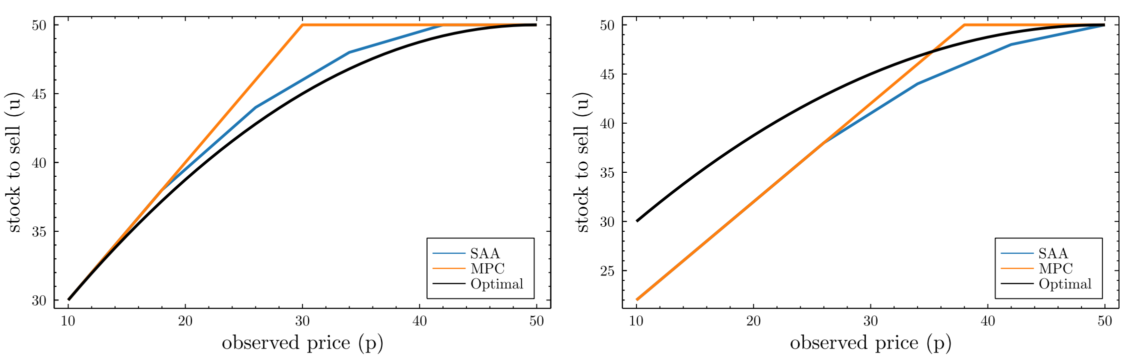
Corollary 2. *The MPC solution to (10) is*

$$u(x,p) = (x - c^{-1}((\mathbb{E}[q] - p)_+))_+.$$

Since $\mathbb{E}[(q-p)_+] \geq (\mathbb{E}[q] - p)_+$ is true by Jensen's inequality, we have the following sequence of inequalities:

$$\begin{aligned} c^{-1}(\mathbb{E}[(q-p)_+]) &\geq c^{-1}((\mathbb{E}[q]-p)_+) \\ -c^{-1}(\mathbb{E}[(q-p)_+]) &\leq -c^{-1}((\mathbb{E}[q]-p)_+) \\ (x - c^{-1}((\mathbb{E}[q]-p)_+))_+ &\geq (x - c^{-1}(\mathbb{E}[(q-p)_+]))_+ \end{aligned} \tag{14}$$

where the first inequality follows because the inverse of a strictly increasing function is strictly increasing. This means that the MPC policy holds less stock in inventory. Intuitively this can be interpreted as a robust policy, since it accepts more of the immediate price known with certainty and gambles less on high future prices. Figure 2 shows SAA and MPC policies as functions of p for different empirical distributions. Figure 2a shows that for a "typical" set of price samples, the SAA policy is a better approximation of the optimal policy than the MPC policy. In contrast, Figure 2b shows that when the price samples have a mean larger than the true expected price, the MPC policy can be a better approximation of the optimal policy than the SAA policy.



(a) $S = \{18, 26, 34, 42\}$.

(b) $S = \{26, 34, 42, 50\}$.

Figure 2: Optimal policies induced by (10) as a function of p for prices uniformly distributed between 10 to 50, an initial stock of 50 units, and different empirical distributions.

4.1 Model predictive control as robustification

We can also formulate a distributionally robust version of (10):

$$R(x,p) = \min_{0 \leq u \leq x} \left\{ -pu + C(x-u) + \max_{\mu \in \mathbb{P}_\theta} \mathbb{E}_\mu[V(x-u,q)] \right\}. \tag{15}$$

where V is defined according to (10) but with each expectation taken with respect to μ. Since purchasing is not allowed and the recursion (10) does not decouple like (5), we are not able to drop irrelevant terms and must compute $\max_{\mu \in \mathbb{P}_\theta} \mathbb{E}_\mu[V(x-u,q)]$.

Theorem 4. *The optimal solution to (15) is*

$$u_\theta(x,p) = (x - c^{-1}(\mathbb{E}_{\mathbb{P}_0}[((q-\theta)-p)_+]))_+.$$

Proof. Consider the term $\max_{\mu\in\mathbb{P}_\theta} \mathbb{E}_\mu[V(x,q)]$ from (15). For general $\mu_1 \in \mathbb{P}_\theta$ we have

$$\mathbb{E}_{\mu_1}\left[V(x,q) = \min_{0\le u\le x}\{-pu + C(x-u) + \mathbb{E}_{\mu_1}[V(x-u,q)]\}\right]. \tag{16}$$

Theorem 3 shows that the optimal policy for (16) is $u_1(x,p) = (x - c^{-1}(\mathbb{E}_{\mu_1}[(q-p)_+]))_+$. For some $\mu_2 \in \mathbb{P}_\theta$ which has each sample ζ_i the same as μ_1 except for some index j where the sample is fixed at $\xi_j - \theta$, Theorem 3 would yield the policy $u_2(x,p) = (x - c^{-1}(\mathbb{E}_{\mu_2}[(q-p)_+]))_+$. If one were to set $u_2'(x,\zeta_j) = (x - c^{-1}(\mathbb{E}_{\mu_2}[(q-(\xi_j-\theta))_+]))_+$ and $u_2'(x,\zeta_i) = u_2(x,\zeta_i)$ for $i \neq j$, applying the amended policy $u_2'(x,p)$ to compute

$$\mathbb{E}_{\mu_1}\left[V(x,p) = -pu_2'(x,p) + C(x - u_2'(x,p)) + \mathbb{E}_{\mu_1}[V(x - u_2'(x,p), q)]\ \right]$$

would achieve a lower value than

$$\mathbb{E}_{\mu_2}\left[V(x,p) = -pu_2(x,p) + C(x - u_2(x,p)) + \mathbb{E}_{\mu_2}[V(x - u_2(x,p), q)]\ \right]$$

since any sales at the price ζ_j would lower the objective by an additional $-(\zeta_j - (\xi_j - \theta))$ while storing an unchanged amount of stock. Since u_2' is a feasible policy for (16), the optimal u_1 must be at least as good. Therefore, $\mathbb{E}_{\mu_1}[V(x,q)] \le \mathbb{E}_{\mu_2}[V(x,q)]$. Because the previous argument works for general μ, the optimal (worst case) μ^* must have each price at $\xi_i - \theta$. Thus, $\max_{\mu\in\mathbb{P}_\theta} \mathbb{E}_\mu[V(x-u,q)] = \mathbb{E}_{\mathbb{P}_0}[V(x-u,q-\theta)]$. Subbing this into (16) and applying Theorem 3 yields

$$u_\theta(x,p) = (x - c^{-1}(\mathbb{E}_{\mathbb{P}_0}[((q-\theta)-p)_+]))_+.$$

□

Corollary 3. *For each x and p there exists a radius θ at which the solution to (15) coincides with the MPC solution to (10), such that*

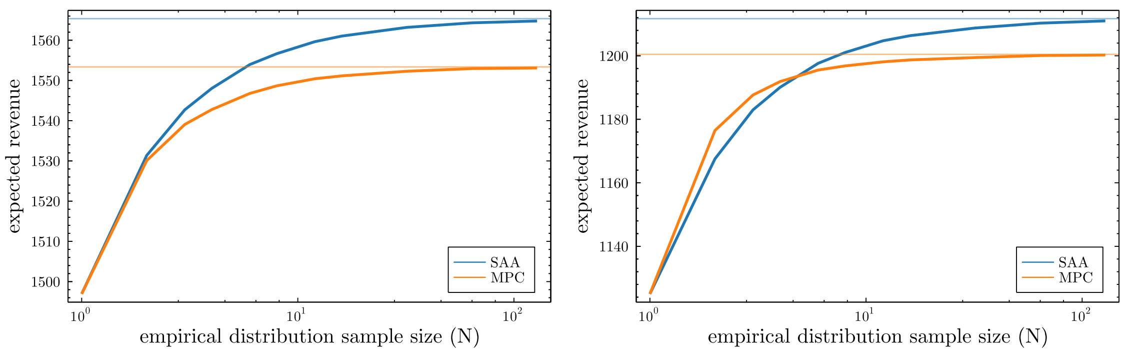
$$(x - c^{-1}(\mathbb{E}_{\mathbb{P}_0}[((q-\theta)-p)_+]))_+ = (x - c^{-1}((\mathbb{E}_{\mathbb{P}_0}[q] - p)_+))_+.$$

Proof. At $\theta = 0$ we have $(x - c^{-1}(\mathbb{E}_{\mathbb{P}_0}[(q-p)_+]))_+ \le (x - c^{-1}((\mathbb{E}_{\mathbb{P}_0}[q]-p)_+))_+$ from (14). For $\theta = \max \xi_i$ we have $(x - c^{-1}(0))_+ = x \ge (x - c^{-1}((\mathbb{E}_{\mathbb{P}_0}[q]-p)_+))_+$. Lastly, since $(x - c^{-1}(\mathbb{E}_{\mathbb{P}_0}[((q-\theta)-p)_+]))_+$ is a composition of continuous functions it is continuous in θ, and by the intermediate value theorem there exists some $0 \le \theta \le \max \xi_i$ such that

$$(x - c^{-1}(\mathbb{E}_{\mathbb{P}_0}[((q-\theta)-p)_+]))_+ = (x - c^{-1}((\mathbb{E}_{\mathbb{P}_0}[q] - p)_+))_+.$$

□

We contend that the interpretation provided by Corollary 3 explains the surprising performance of MPC in (Guan, 2019). Its robust nature means that it will do passably well in most applications. When the price distribution is right-skewed so that there are occasionally very high prices that the SAA policy is willing to risk waiting for, the robustness of the MPC policy protects against waiting too long under empirical distributions which are poor representations of the true price distribution. This is supported by Figure 2; if the price distribution is right-skewed the sample mean will often be much greater than the expected price, and so the MPC policy will often be a better approximation of the optimal policy that that derived from SAA. We demonstrate this by running two simulations, one without right-skewed prices and one with right-skewed prices. Suppose that we have 100 stages of selling and an initial stock level of 50 units. Figure 3a shows that without a right-skew, the SAA policy always outperforms the MPC policy. Figure 3b shows that for modest N, right-skewed prices mean that the MPC policy outperforms the SAA policy. Indeed, the study of Guan (2019) was limited in N due to severe computational expense, so this explanation is plausible.



(a) Prices uniformly distributed between 10 and 50.

(b) Prices triangularly distributed between 10 and 50 with a mode at 10.[1]

Figure 3: Expected revenue accrued over 10000 simulations as a function of sample size, for the SAA policy and MPC policy. The horizontal lines are the expected revenue of these policies under the true price distributions.

The right-skew used to generate Figure 3b is not that extreme. More general settings may involve price distributions with fat tails and increase the range of N where MPC is the better choice. Furthermore, taking inspiration from Figure 1, robust policies with θ less than the value in Corollary 3 will also increase this range and may prove to be useful in applications with minimal right-skew.

5 Conclusion

We have shown that MPC can be interpreted as distributionally robust SDP in the context of Fonterra Co-operative Group Limited. This is a feasible explanation of the competitive nature of MPC, which we supported with numerical studies showing that right-skew in the underlying price distribution created opportunity for robustification to be beneficial. We have also shown that distributional robustness offers out-of-sample value in its own right, which suggests that

[1] The increase in revenue is statistically significant. At $N = 3$ the difference in revenue is 1.60. Using a common random number approach gives a standard error of 0.45.

additional competitive policies could be achieved by exploring policies that are between the SAA policy and the MPC policy.

6 Acknowledgements

We would like to thank A. J. Mason for lengthy discussions, help, and encouragement.

References

Anderson, Edward and Andy Philpott (2022). "Improving Sample Average Approximation Using Distributional Robustness". In: *INFORMS Journal on Optimization* 4(1), pp. 90–124. URL: http://www.jstor.org/stable/2284155.

Bellman, R., I. Glicksberg, and O. Gross (1955). "On the Optimal Inventory Equation". In: *Management Science* 2.1, pp. 83–104. ISSN: 00251909, 15265501. URL: http://www.jstor.org/stable/2627240.

Camacho, Eduardo F and Carlos Bordons Alba (2013). *Model predictive control*. Springer science & business media.

Guan, Ziming (2019). *SDDP production planning model and SDDP hedge contract model*. Tech. rep. Fonterra Co-operative Group Limited, Ingredients Portfolio Management.

Kouvaritakis, Basil and Mark Cannon (2016). *Model Predictive Control Classical, Robust and Stochastic*. eng. 1st ed. 2016. Advanced Textbooks in Control and Signal Processing. Cham: Springer International Publishing. ISBN: 3-319-24853-7.

Ross, Sheldon M (2014). *Introduction to stochastic dynamic programming*. Academic press.

Shapiro, Alexander (2017). "Distributionally Robust Stochastic Programming". In: *SIAM Journal on Optimization* 27.4, pp. 2258–2275. DOI: 10.1137/16M1058297. URL: https://doi.org/10.1137/16M1058297.

Shapiro, Alexander, Darinka Dentcheva, and Andrzej Ruszczynski (2021). *Lectures on Stochastic Programming: Modeling and Theory, Third Edition*. Philadelphia, PA: Society for Industrial and Applied Mathematics. DOI: 10.1137/1.9781611976595. URL: https://epubs.siam.org/doi/abs/10.1137/1.9781611976595.

Theil, H. (1957). "A Note on Certainty Equivalence in Dynamic Planning". In: *Econometrica* 25.2, pp. 346–349. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1910260.

Ziemba, William T. (1971). "Transforming Stochastic Dynamic Programming Problems into Nonlinear Programs". In: *Management Science* 17.7, pp. 450–462. ISSN: 00251909, 15265501. URL: http://www.jstor.org/stable/2629541.