Online Supplement

William L. Cooper and Bharath Rangarajan. "Performance guarantees for empirical Markov decision processes with applications to multiperiod inventory models". *Operations Research*, 60(5), 1267-1281, 2012.

Online Supplement

S-1 Proofs and Supplementary Material for Section 4

The proof of Theorem 1 relies on a series of preliminary results. The main idea is to compare MDPs that differ only in their disturbance distributions, and to show that when the disturbance distributions are suitably "close", then so are solutions of the MDPs. We then complete the proof by bounding the probability with which the true disturbance distributions and the empirical disturbance distributions are close.

We begin by presenting a result due to Massart (1990, Corollary 1 and Comment 2(iii)).

Lemma S-1 (Massart 1990) *Suppose $n \geq 1$ and $Z^1, \ldots, Z^n$ are i.i.d. random variables with common distribution function F. Let F^n denote the empirical distribution function determined by $Z^1, \ldots, Z^n$; i.e., $F^n(z) = n^{-1} \sum_{i=1}^n 1\{Z^i \leq z\}$ for $z \in \mathbb{R}$. Then $P(\|F - F^n\| > \gamma) \leq 2\exp(-2\gamma^2 n)$ for any $\gamma > 0$.*

We will also need the following technical lemma.

Lemma S-2 *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz with constant ρ, and that $a, b \in \mathbb{R}$ satisfy $a \leq b$. Suppose also that F, G are distributions such that $F(a-) = G(a-)$, $F(b) = G(b)$, and $\|F - G\| \leq \gamma$. Then $\left| \int_{[a,b]} f(z)dF(z) - \int_{[a,b]} f(z)dG(z) \right| \leq \gamma(b-a)\rho$.*

Proof. The function f is Lipschitz with constant ρ, so it is continuous. Hence, $\int_{[a,b]} f(z)dF(z) = f(b)F(b) - f(a-)F(a-) - \int_{[a,b]} F(z)df(z)$ by an integration-by-parts formula for Lebesgue-Stieltjes integrals; see Folland (1999, p. 108, Ex. 34b). Moreover, f is absolutely continuous and its derivative f' exists almost everywhere on $[a, b]$ and satisfies $|f'(z)| \leq \rho$ almost everywhere; see Folland (1999, p. 106, Thm. 3.35 and p. 108, Ex. 37). Hence, $\int_{[a,b]} F(z)df(z) = \int_{[a,b]} F(z)f'(z)dm(z)$ where $m(\cdot)$ is Lebesgue measure. It follows that $\int_{[a,b]} f(z)dF(z) = f(b)F(b) - f(a-)F(a-) - \int_{[a,b]} F(z)f'(z)dm(z)$. Similarly, $\int_{[a,b]} f(z)dG(z) = f(b)G(b) - f(a-)G(a-) - \int_{[a,b]} G(z)f'(z)dm(z)$. Therefore, we have $\left| \int_{[a,b]} f(z)dF(z) - \int_{[a,b]} f(z)dG(z) \right| = \left| \int_{[a,b]} [G(z) - F(z)]f'(z)dm(z) \right| \leq \int_{[a,b]} \|F - G\| \rho dm(z) \leq \gamma(b-a)\rho$, which completes the proof □

Below we consider two separate finite-horizon MDPs that are identical except for the disturbance distributions. In one model, the disturbance distributions are $\{F_t\}$ and in the other the disturbance distributions are $\{G_t\}$. We append superscripts F and G below to indicate this dependence; for $\Phi = F$ or G, we have $C_t^\Phi(x,q) = \int c_t(x,q,z)d\Phi_t(z)$ and $V_t^\Phi(x) = \inf_{q\in\mathcal{A}(x)} W_t^\Phi(x,q)$, where $W_\tau^\Phi(x,q) = C_\tau^\Phi(x,q)$ and $W_t^\Phi(x,q) = C_t^\Phi(x,q) + \int V_{t+1}^\Phi(\psi_t(x,q,z))d\Phi_t(z)$ for $t < \tau$. For $\varepsilon \geq 0$ let $\{q_t^{\varepsilon,\Phi}(x)\}$ be an ε-optimal policy for the problem with demand distributions $\{\Phi_t\}$ so that $W_t^\Phi(x, q_t^{\varepsilon,\Phi}(x)) \leq V_t^\Phi(x) + \varepsilon/\tau$. The following result states that similar disturbance distributions yield similar value functions.

Proposition S-1 *Suppose $F_k(\alpha_k-) = G_k(\alpha_k-) = 0$ and $F_k(\beta_k) = G_k(\beta_k) = 1$ and $\|F_k - G_k\| \leq \gamma_k$ for $k = t,\ldots,\tau$ where $\{\gamma_k \geq 0\}$ are non-negative numbers. Suppose also that for $k = t,\ldots,\tau-1$, we have that $V_{k+1}^F(\psi_k(x,q,\cdot))$ is Lipschitz with constant H_{k+1} and for $t = k,\ldots,\tau$, we have that $c_k(x,q,\cdot)$ is Lipschitz with constant ρ_k for all x,q. Define $\theta_k = \gamma_\tau(\beta_\tau - \alpha_\tau)\rho_\tau + \sum_{j=k}^{\tau-1}\gamma_j(\beta_j - \alpha_j)(\rho_j + H_{j+1})$ for $k = t,\ldots,\tau$. Then $\|V_k^F - V_k^G\| \leq \|W_k^F - W_k^G\| \leq \theta_k$ for all $k = t,\ldots,\tau$.*

Proof. The proof is by induction. For the base case, $W_\tau^\Phi(x,q) = C_\tau^\Phi(x,q) = \int_{[\alpha_\tau,\beta_\tau]} c_\tau(x,q,z)d\Phi_\tau(z)$ for $\Phi = F, G$. We have assumed that $c_\tau(x,q,\cdot)$ is Lipschitz with constant ρ_τ; Lemma S-2 now implies that $|C_\tau^F(x,q) - C_\tau^G(x,q)| \leq \gamma_\tau(\beta_\tau - \alpha_\tau)\rho_\tau$, and hence we have $\|W_\tau^F - W_\tau^G\| \leq \gamma_\tau(\beta_\tau - \alpha_\tau)\rho_\tau$. To see that $\|V_\tau^F - V_\tau^G\| \leq \|W_\tau^F - W_\tau^G\|$, observe that for any $\varepsilon > 0$ we have

$$V_\tau^F(x) \geq W_\tau^F(x, q_\tau^{\varepsilon,F}(x)) - \varepsilon/\tau \geq W_\tau^G(x, q_\tau^{\varepsilon,F}(x)) - \|W_\tau^F - W_\tau^G\| - \varepsilon/\tau \geq V_\tau^G(x) - \|W_\tau^F - W_\tau^G\| - \varepsilon/\tau\,.$$

Letting $\varepsilon \downarrow 0$, we see that $V_\tau^F(x) \geq V_\tau^G(x) - \|W_\tau^F - W_\tau^G\|$. Reversing the roles of F and G shows that $V_\tau^G(x) \geq V_\tau^F(x) - \|W_\tau^F - W_\tau^G\|$, and hence $\|V_\tau^F - V_\tau^G\| \leq \|W_\tau^F - W_\tau^G\|$. This completes the base case. For general t we have $\|V_t^F - V_t^G\| \leq \|W_t^F - W_t^G\|$ by an identical argument. Moreover,

$$\begin{aligned}
&\left|W_t^F(x,q) - W_t^G(x,q)\right| \\
&\qquad = \left|C_t^F(x,q) + \int_{[\alpha_t,\beta_t]} V_{t+1}^F(\psi_t(x,q,z))dF_t(z) - C_t^G(x,q) - \int_{[\alpha_t,\beta_t]} V_{t+1}^G(\psi_t(x,q,z))dG_t(z)\right| \\
&\qquad \leq \left|\int_{[\alpha_t,\beta_t]} c_t(x,q,z)dF_t(z) - \int_{[\alpha_t,\beta_t]} c_t(x,q,z)dG_t(z)\right| \\
&\qquad\quad + \left|\int_{[\alpha_t,\beta_t]} V_{t+1}^F(\psi_t(x,q,z))dF_t(z) - \int_{[\alpha_t,\beta_t]} V_{t+1}^F(\psi_t(x,q,z))dG_t(z)\right| \\
&\qquad\quad + \left|\int_{[\alpha_t,\beta_t]} V_{t+1}^F(\psi_t(x,q,z))dG_t(z) - \int_{[\alpha_t,\beta_t]} V_{t+1}^G(\psi_t(x,q,z))dG_t(z)\right| \\
&\qquad \leq \gamma_t(\beta_t - \alpha_t)\rho_t + \gamma_t(\beta_t - \alpha_t)H_{t+1} + \theta_{t+1} = \theta_t
\end{aligned}$$

where the final inequality follows from Lemma S-2 (applied twice) and the induction hypothesis. In view of the above, $\|W_t^F - W_t^G\| \leq \theta_t$, which completes the proof. □

Next we examine how well an ε-optimal policy from the model with distributions $\{G_t\}$ performs in a situation that actually has disturbance distributions $\{F_t\}$. To this end, let $V_\tau^{F,\varepsilon,G}(x) = C_\tau^F(x, q_\tau^{\varepsilon,G}(x))$ and $V_t^{F,\varepsilon,G}(x) = C_t^F(x, q_t^{\varepsilon,G}(x)) + \int V_{t+1}^{F,\varepsilon,G}(\psi_t(x, q_t^{\varepsilon,G}(x), z))dF_t(z)$ for $t < \tau$. The quantity $V_t^{F,\varepsilon,G}(x)$ represents the expected total cost from time t onward in the model with disturbance distributions $\{F_t\}$, given that the state at the start of period t is x, and assuming that actions are selected according to $\{q_t^{\varepsilon,G}(x)\}$.

Proposition S-2 *Suppose $\varepsilon \geq 0$ and $\|W_k^F - W_k^G\| \leq \eta_k$ for $k = t, \ldots, \tau$, where $\{\eta_k \geq 0\}$ is a collection of non-negative numbers. Then $V_k^{F,\varepsilon,G}(x) \leq V_k^F(x) + 2\sum_{j=k}^{\tau} \eta_j + (\tau - k + 1)\varepsilon/\tau$ for all $x \in \mathcal{X}$ and $k = t, \ldots, \tau$*

Proof. Observe first that $\|W_k^F - W_k^G\| \leq \eta_k$ implies that for any $\varepsilon' > 0$ we have $V_k^F(x) \geq W_k^F(x, q_k^{\varepsilon',F}(x)) - \varepsilon'/\tau \geq W_k^G(x, q_k^{\varepsilon',F}(x)) - \eta_k - \varepsilon'/\tau \geq V_k^G(x) - \eta_k - \varepsilon'/\tau \geq W_k^G(x, q_k^{\varepsilon,G}(x)) - \varepsilon/\tau - \eta_k - \varepsilon'/\tau \geq W_k^F(x, q_k^{\varepsilon,G}(x)) - \varepsilon/\tau - 2\eta_k - \varepsilon'/\tau$. Letting $\varepsilon' \downarrow 0$ we obtain

$$W_k^F(x, q_k^{\varepsilon,G}(x)) \leq V_k^F(x) + 2\eta_k + \varepsilon/\tau \ . \tag{S-1}$$

The proof of the proposition is by backwards induction on t. By (S-1), the result is true for

$t = \tau$ because $W_\tau^F(x, q_\tau^{\varepsilon,G}(x)) = V_\tau^{F,\varepsilon,G}(x)$. Proceeding by induction, for arbitrary $t < \tau$ we have

$$
\begin{aligned}
V_t^{F,\varepsilon,G}(x) &= C_t^F(x, q_t^{\varepsilon,G}(x)) + \int V_{t+1}^{F,\varepsilon,G}(\psi_t(x, q_t^{\varepsilon,G}(x), z))dF_t(z) \\
&\le C_t^F(x, q_t^{\varepsilon,G}(x)) + \int \Big[V_{t+1}^F(\psi_t(x, q_t^{\varepsilon,G}(x), z)) + 2\sum_{j=t+1}^{\tau} \eta_j + \frac{(\tau-t)\varepsilon}{\tau}\Big] dF_t(z) \\
&= W_t^F(x, q_t^{\varepsilon,G}(x)) + 2\sum_{j=t+1}^{\tau} \eta_j + \frac{(\tau-t)\varepsilon}{\tau} \\
&\le V_t^F(x) + 2\eta_t + \frac{\varepsilon}{\tau} + 2\sum_{j=t+1}^{\tau} \eta_j + \frac{(\tau-t)\varepsilon}{\tau},
\end{aligned}
$$

where the first inequality follows from the inductive hypothesis and the second follows from (S-1). This completes the proof. □

We are now ready for the proof of the main theorem of Section 4.

Proof of Theorem 1. Let $\Omega' = \{\omega \in \Omega : Z_k^i(\omega) \in [\alpha_k, \beta_k]$ for all $k = 1, \ldots, \tau; i = 1, \ldots, n_k\}$. Note that $\mathbb{P}[\Omega'] = 1$, and that for $\omega \in \Omega'$ we have $\widehat{F}_k(\alpha_k-, \omega) = 0$ and $\widehat{F}_k(\beta_k, \omega) = 1$ for $k = 1, \ldots, \tau$. Consider a fixed $\omega \in \Omega'$ such that $\|F_k(\cdot) - \widehat{F}_k(\cdot, \omega)\| \le \gamma_k$ for each $k = 1, \ldots, \tau$ where $\{\gamma_k\}$ are defined as in the statement of part 1 of Theorem 1. For these values of $\{\gamma_k\}$, note that $\{\theta_k\}$ in the statement of Proposition S-1 are given by $\theta_k = \sum_{j=k}^{\tau} \epsilon'_j = \epsilon_k$. It follows from Proposition S-1 with $\{G_k(\cdot)\} = \{\widehat{F}_k(\cdot, \omega)\}$ that $\|V_k(\cdot) - \widehat{V}_k(\cdot, \omega)\| \le \|W_k(\cdot) - \widehat{W}_k(\cdot, \omega)\| \le \epsilon_k$ for $k = 1, \ldots, \tau$. We may now apply Proposition S-2 with $\{\eta_k\} = \{\epsilon_k\}$, $\{G_k(\cdot)\} = \{\widehat{F}_k(\cdot, \omega)\}$, and $\{V_k^{F,\varepsilon,G}(\cdot)\} = \{\widetilde{V}_k(\cdot, \omega)\}$. Doing so, we have that $0 \le \widetilde{V}_k(x, \omega) - V_k(x) \le 2\sum_{j=k}^{\tau} \epsilon_j + (\tau - k + 1)\varepsilon/\tau$ for all x and $k = 1, \ldots, \tau$.

To summarize, if $\omega \in \Omega'$ and $\|F_k(\cdot) - \widehat{F}_k(\cdot, \omega)\| \le \gamma_k$ for $k = 1, \ldots, \tau$, then $\|V_k(\cdot) - \widehat{V}_k(\cdot, \omega)\| \le \|W_k(\cdot) - \widehat{W}_k(\cdot, \omega)\| \le \epsilon_k$ for $k = 1, \ldots, \tau$ and $0 \le \widetilde{V}_k(x, \omega) - V_k(x) \le 2\sum_{j=k}^{\tau} \epsilon_j + (\tau - k + 1)\varepsilon/\tau$ for all x and $k = 1, \ldots, \tau$. Consequently, $A \supseteq A' := \Omega' \cap \big(\cap_{k=1,\ldots,\tau} \{\|F_k - \widehat{F}_k\| \le \gamma_k\}\big)$ and therefore

$$
\begin{aligned}
\mathbb{P}[A] \ge \mathbb{P}[A'] &= 1 - \mathbb{P}\big[\cup_{k=1,\ldots,\tau} \{\|F_k - \widehat{F}_k\| > \gamma_k\}\big] \\
&\ge 1 - \sum_{k=1}^{\tau} \mathbb{P}\big[\|F_k - \widehat{F}_k\| > \gamma_k\big] \\
&\ge 1 - \sum_{k=1}^{\tau} 2\exp(-2\gamma_k^2 n_k) \qquad \text{(S-2)}
\end{aligned}
$$

where (S-2) follows from Lemma S-1.

To prove part 2, we apply part 1 with $\epsilon'_k = (\epsilon - \varepsilon)/(\tau^2 + \tau)$, which case $\epsilon_k = (\tau - k + 1)(\epsilon - \varepsilon)/(\tau^2 + \tau)$, $\gamma_k = (\epsilon - \varepsilon)/[\lambda_k(\tau^2 + \tau)]$, and $2\sum_{j=k}^{\tau} \epsilon_j + (\tau - k + 1)\varepsilon/\tau = \xi_k(\epsilon, \varepsilon)$. If $n_k \ge [2(\epsilon - \varepsilon)^2]^{-1}(\tau^2 +$

$\tau)^2\lambda_k^2\log(2\tau/\delta)$, then $2\exp(-2\gamma_k^2 n_k) \le \delta/\tau$, and it follows that $1 - \sum_{k=1}^{\tau} 2\exp(-2\gamma_k^2 n_k) \ge 1 - \delta$, thereby completing the proof of part 2.

For part 3, if $\mathbb{P}\big[\min_{q\in\mathcal{A}(x)} \widehat{W}_t(x,q)$ exists for all $x \in \mathcal{X}$ and $t = 1,\ldots,\tau\big] = 1$, then the actions $\{\widehat{q}_t^0(x)\}$ exist $\mathbb{P}$-almost surely, in which case the arguments in parts 1 and 2 hold with $\varepsilon = 0$. $\square$

We close this section with a proposition that provides conditions for the application of Theorem 1. Checking the conditions in the proposition does not require knowledge of the disturbance distributions. Note also that the values of $c_t(x,q,\cdot)$ and $\psi_t(x,q,\cdot)$ do not matter off the support of F_t. Hence, for MDPs with finite $\mathcal{X}$ and $\mathcal{A}$ and finite support of each F_t, the suppositions of Proposition S-3 below and Theorem 1 can be made to hold by appropriately defining c_t and ψ_t off the support of F_t (doing so requires knowing the support of F_t but not F_t itself). Therefore, Theorem 1 applies to MDPs with finite state and action spaces and disturbance distributions with finite support.

Proposition S-3 *Suppose that $c_t(\cdot,q,z)$ is Lipschitz with constant L_t^c for all q,z for $t = 1,\ldots,\tau$ and that $\psi_t(\cdot,q,z)$ is Lipschitz with constant L_t^ψ for all q,z for $t = 1,\ldots,\tau-1$. Then V_t is Lipschitz with constant L_t for $t = 1,\ldots,\tau$ where $L_\tau = L_\tau^c$ and $L_t = L_t^c + L_{t+1}L_t^\psi$ for $t = 1,\ldots,\tau-1$. If in addition, $\psi_t(x,q,\cdot)$ is Lipschitz with constant ℓ_t^ψ for all x,q for $t = 1,\ldots,\tau-1$, then $V_{t+1}(\psi_t(x,q,\cdot))$ is Lipschitz with constant $H_{t+1} = \ell_t^\psi L_{t+1}$ for $t = 1,\ldots,\tau-1$.*

Proof. Observe that for arbitrary $x, x' \in \mathcal{X}$ and $q \in \mathcal{A}$ we have

$$\big|C_t(x,q) - C_t(x',q)\big| \le \int_z \big|c_t(x,q,z) - c_t(x',q,z)\big| dF_t(z) \le \int_z L_t^c \big|x - x'\big| dF_t(z) = L_t^c\big|x - x'\big|.$$

Hence, $C_t(\cdot,q)$ is Lipschitz with constant L_t^c for $t = 1,\ldots,\tau$.

We next prove that V_t is Lipschitz with constant L_t by backwards induction. For $t = \tau$ and $\varepsilon > 0$ we have

$$V_\tau(x) = \inf_{q\in\mathcal{A}(x)} C_\tau(x,q) \le C_\tau(x, q_\tau^\varepsilon(x')) \le C_\tau(x', q_\tau^\varepsilon(x')) + L_\tau^c\big|x - x'\big| \le V_\tau(x') + \varepsilon/\tau + L_\tau^c\big|x - x'\big|.$$

Letting $\varepsilon \downarrow 0$, we obtain $V_\tau(x) \le V_\tau(x') + L_\tau^c\big|x - x'\big|$, and hence $V_\tau(x) - V_\tau(x') \le L_\tau^c\big|x - x'\big|$. Reversing the roles of x and x', it follows that $\big|V_\tau(x) - V_\tau(x')\big| \le L_\tau^c\big|x - x'\big|$. That is, V_τ is Lipschitz with constant $L_\tau = L_\tau^c$.

For the inductive step we have

$$\begin{aligned}\big|W_t(x,q)-W_t(x',q)\big| &\le \big|C_t(x,q)-C_t(x',q)\big| + \int_z \big|V_{t+1}(\psi_t(x,q,z)) - V_{t+1}(\psi_t(x',q,z))\big| dF_t(z) \\ &\le L_t^c\big|x-x'\big| + \int_z L_{t+1}\big|\psi_t(x,q,z)-\psi_t(x',q,z)\big| dF_t(z) \\ &\le L_t^c\big|x-x'\big| + \int_z L_{t+1}L_t^{\psi}\big|x-x'\big| dF_t(z) \\ &= (L_t^c + L_{t+1}L_t^{\psi})\big|x-x'\big|,\end{aligned}$$

so $W_t(\cdot,q)$ is Lipschitz with constant $L_t = L_t^c + L_{t+1}L_t^{\psi}$. Consequently, V_t is Lipschitz with constant L_t by an argument identical to that given for $t=\tau$.

It follows immediately that if $\psi_t(x,q,\cdot)$ is also Lipschitz with constant ℓ_t^{ψ}, then $V_{t+1}(\psi_t(x,q,\cdot))$ is Lipschitz with constant $H_{t+1} = \ell_t^{\psi} L_{t+1}$. □

S-2 Proofs and Supplementary Material for Section 5

Proof of Lemma 2. The proof of the first statement is by backward induction on t. Suppose without loss of generality that $x_2 \ge x_1$ and define $\Delta = x_2 - x_1$. For x sufficiently large, we have $V_\tau(x) = K_\tau(x)$. Moreover, by the monotonicity and convexity of V_τ (see Lemma 1) we have $0 \le V_\tau(x_2) - V_\tau(x_1) \le \lim_{x\to\infty} V_\tau(x+\Delta) - V_\tau(x) = \lim_{x\to\infty} K_\tau(x+\Delta) - K_\tau(x) = h_\tau\Delta$. Hence, the result holds for $t=\tau$. For general $t<\tau$, suppose the result is true for $t+1$ and consider $x_2 \ge x_1$. By similar logic it follows that

$$\begin{aligned}0 &\le V_t(x_2) - V_t(x_1) \\ &= U_t(\min\{x_2, y_t^*\}) - U_t(\min\{x_1, y_t^*\}) \\ &\le \lim_{x\to\infty} U_t(x+\Delta) - U_t(x) \\ &= \lim_{x\to\infty} \{K_t(x+\Delta) - K_t(x)\} + \lim_{x\to\infty}\left\{\int \big[V_{t+1}(\phi(x+\Delta-z)) - V_{t+1}(\phi(x-z))\big] dF_t(z)\right\} \\ &\le \lim_{x\to\infty} \{K_t(x+\Delta) - K_t(x)\} + \lim_{x\to\infty}\left\{\int H_{t+1}\big[\phi(x+\Delta-z) - \phi(x-z)\big] dF_t(z)\right\} \\ &\le h_t\Delta + \int H_{t+1}\Delta dF_t(z) \;=\; [h_t + \cdots + h_\tau]\Delta \;=\; H_t\Delta\,,\end{aligned}$$

where the second-to-last equality follows from the inductive hypothesis and the monotonicity of ϕ, and the final inequality follows from the fact that ϕ is Lipschitz with constant 1.

The second statement in the lemma follows immediately from the first and the fact that ϕ is Lipschitz with constant 1. $\square$

In the remainder of this section, we prove Theorem 3. We follow an approach that is similar to that used for Theorem 1. That is, we begin by considering two separate inventory systems that are identical, except that one faces demand distributions $\{F_t\}$ and the other faces demand distributions $\{G_t\}$. We again indicate this dependence by appending a superscript F or G to our notation.

For $f : \mathbb{R} \to \mathbb{R}$, let f^r (respectively, f^l) denote the right-derivative (resp., left-derivative) of f.

Lemma S-3 *If* $\|F_t - G_t\| \leq \gamma_t$ *then* $\|(K_t^F)^r - (K_t^G)^r\| \leq (b_t + h_t)\gamma_t$.

Proof. For $y \in \mathbb{R}$, we have $(K_t^\Phi)^r(y) = \int_z k_t^r(y-z)\, d\Phi_t(z) = \int_{z>y} -b_t\, d\Phi_t(z) + \int_{z\leq y} h_t\, d\Phi_t(z) = (h_t+b_t)\Phi_t(y) - b_t$ for $\Phi = F, G$. Hence $|(K_t^F)^r(y) - (K_t^G)^r(y)| = (b_t+h_t)|F_t(y) - G_t(y)| \leq (b_t+h_t)\gamma_t$. $\square$

Given $\delta > 0$ and a convex function $f : \mathbb{R} \to \mathbb{R}$, we say that $y \in \mathbb{R}$ is a *δ-point* of f if there exists a subgradient $v \in \partial f(y)$ such that $|v| \leq \delta$; see Levi et al. (2007). We will use the following result due to Levi et al.

Lemma S-4 (Levi et al. 2007) *Suppose that* $f : \mathbb{R} \to \mathbb{R}$ *is convex with a minimum at* $y^* \in \mathbb{R}$. *Suppose also that* $h, b > 0$, $d \in \mathbb{R}$, *and* $f(y) \geq \bar{f}(y) := h(y-d)^+ + b(d-y)^+$ *for all* $y \in \mathbb{R}$. *For* $0 < \epsilon \leq 1$, *if* y' *is an* $(\epsilon \min\{b,h\}/3)$*-point of* f, *then* $f(y') \leq (1+\epsilon) f(y^*)$.

We will also use Theorem 24.1 of Rockafellar (1970) several times. We state it here for convenience as the following lemma.

Lemma S-5 *Suppose that* $f : \mathbb{R} \to \mathbb{R}$ *is convex. Let* $f^r(x)$ *and* $f^l(x)$ *be the right and left derivatives of* f *at point* x. *Then* $f^r(z_1) \leq f^l(x) \leq f^r(x) \leq f^l(z_2)$ *when* $z_1 < x < z_2$. *Moreover, for every* x *we have* $\lim_{z\downarrow x} f^r(z) = f^r(x)$, $\lim_{z\uparrow x} f^r(z) = f^l(x)$, $\lim_{z\downarrow x} f^l(z) = f^r(x)$, *and* $\lim_{z\uparrow x} f^l(z) = f^l(x)$.

Lemma S-5 implies that $(U_t^\Phi)^r$ [resp., $(U_t^\Phi)^l$] is a right-continuous [left-continuous] function, and $\lim_{z\uparrow x} (U_t^\Phi)^r(z) = (U_t^\Phi)^l(x)$ for $\Phi = F, G$. Since $y_t^\Phi = \inf\{y : (U_t^\Phi)^r(y) \geq 0\}$, it follows that $(U_t^\Phi)^r(y_t^\Phi) \geq 0$ by the right-continuity of $(U_t^\Phi)^r$. For any $y < y_t^\Phi$, we have $(U_t^\Phi)^r(y) < 0$, and hence $(U_t^\Phi)^l(y) \leq (U_t^\Phi)^r(y) < 0$. Therefore, $(U_t^\Phi)^l(y_t^\Phi) \leq 0$ by the left-continuity of $(U_t^\Phi)^l$.

Proposition S-4 *Suppose* $\|F_k - G_k\| \le \gamma_k$ *for* $k = t, \ldots, \tau$ *where* $\{\gamma_k\}$ *is a collection of non-negative numbers. Define* $\nu_k := \sum_{j=k}^{\tau} \zeta_j \gamma_j$ *for* $k = t, \ldots, \tau$, *where* $\zeta_j = b_j + H_j$ *and* $H_j = h_j + \cdots + h_\tau$ *is the Lipschitz constant of* V_j^F. *Then* $\|(V_k^F)^r - (V_k^G)^r\| \le \|(U_k^F)^r - (U_k^G)^r\| \le \nu_k$ *for* $k = t, \ldots, \tau$.

Proof. We begin by showing that $\|(V_k^F)^r - (V_k^G)^r\| \le \|(U_k^F)^r - (U_k^G)^r\|$ for all k. For $\Phi = F, G$, if $x \ge y_k^\Phi$ then $(V_k^\Phi)^r(x) = (U_k^\Phi)^r(x) \ge 0$, and if $x < y_k^\Phi$ then $0 = (V_k^\Phi)^r(x) > (U_k^\Phi)^r(x)$. Hence, if $x \ge \max\{y_k^F, y_k^G\}$, then $(V_k^G)^r(y_k^G) = (U_k^G)^r(y_k^G)$, and therefore $(V_k^F)^r(x) - (V_k^G)^r(x) = (U_k^F)^r(x) - (U_k^G)^r(x)$. If $y_k^F \le x < y_k^G$, then $0 = (V_k^G)^r(x) > (U_k^G)^r(x)$. Hence $0 \le (V_k^F)^r(x) - (V_k^G)^r(x) \le (U_k^F)^r(x) - (U_k^G)^r(x)$. Likewise, if $y_k^G \le x < y_k^F$, then $0 \le (V_k^G)^r(x) - (V_k^F)^r(x) \le (U_k^G)^r(x) - (U_k^F)^r(x)$. Finally, if $x < \min\{y_k^F, y_k^G\}$ then $|(V_k^F)^r(x) - (V_k^G)^r(x)| = 0 \le |(U_k^F)^r(x) - (U_k^G)^r(x)|$. Thus, we have proved that $\|(V_k^F)^r - (V_k^G)^r\| \le \|(U_k^F)^r - (U_k^G)^r\|$.

In the remainder of the proof, we show $\|(U_k^F)^r - (U_k^G)^r\| \le \nu_k$ for $k = t, \ldots, \tau$ by backwards induction on t. For $t = \tau$, $\|F_\tau - G_\tau\| \le \gamma_\tau$ implies that $\|(U_\tau^F)^r - (U_\tau^G)^r\| \le (b_\tau + h_\tau)\gamma_\tau = \nu_\tau$ by Lemma S-3, so the result holds for $t = \tau$. For the inductive hypothesis, suppose that the result is true for general $t > 1$.

Recall that $U_{t-1}^\Phi(y) = K_{t-1}^\Phi(y) + \int_z (V_t^\Phi \circ \phi)(y - z) d\Phi_{t-1}(z)$ for $1 < t \le \tau$. Hereafter, the notation $f \circ g$ represents the composition of a function f with a function g; i.e., $(f \circ g)(x) = f(g(x))$. By the dominated convergence theorem, we obtain

$$(U_{t-1}^\Phi)^r(y) = (K_{t-1}^\Phi)^r(y) + \int_z (V_t^\Phi \circ \phi)^r(y - z) d\Phi_{t-1}(z) \,. \tag{S-3}$$

By (S-3) it follows that

$$\begin{aligned} |(U_{t-1}^F)^r(y) - (U_{t-1}^G)^r(y)| \;\; &\le \;\; |(K_{t-1}^F)^r(y) - (K_{t-1}^G)^r(y)| \qquad \text{(S-4)} \\ &\quad + \left| \int_z (V_t^F \circ \phi)^r(y - z) dF_{t-1}(z) - \int_z (V_t^G \circ \phi)^r(y - z) dG_{t-1}(z) \right| \end{aligned}$$

We have

$$\begin{aligned} &\left| \int_z (V_t^F \circ \phi)^r(y - z) dF_{t-1}(z) - \int_z (V_t^G \circ \phi)^r(y - z) dG_{t-1}(z) \right| \\ &\quad \le \left| \int_z (V_t^F \circ \phi)^r(y - z) dF_{t-1}(z) - \int_z (V_t^F \circ \phi)^r(y - z) dG_{t-1}(z) \right| \qquad \text{(S-5)} \\ &\quad\quad + \left| \int_z (V_t^F \circ \phi)^r(y - z) dG_{t-1}(z) - \int_z (V_t^G \circ \phi)^r(y - z) dG_{t-1}(z) \right| . \end{aligned}$$

Let us first consider the second term on the right side of (S-5):

$$
\begin{aligned}
&\left|\int_z \left(V_t^F \circ \phi\right)^r (y-z) dG_{t-1}(z) - \int_z \left(V_t^G \circ \phi\right)^r (y-z) dG_{t-1}(z)\right| \\
&\qquad \le \int_z \left|\left(V_t^F \circ \phi\right)^r (y-z) - \left(V_t^G \circ \phi\right)^r (y-z)\right| dG_{t-1}(z) \\
&\qquad \le \sup_z \left|\left(V_t^F \circ \phi\right)^r (y-z) - \left(V_t^G \circ \phi\right)^r (y-z)\right| \\
&\qquad \le \sup_z \left|\left(V_t^F\right)^r (\phi(y-z)) - \left(V_t^G\right)^r (\phi(y-z))\right| \sup_z \left|\phi^r(y-z)\right| \\
&\qquad \le \nu_t \,, \qquad \text{(S-6)}
\end{aligned}
$$

where the final inequality follows from the inductive hypothesis and the fact that ϕ is Lipschitz with constant 1.

To deal with the first term on the right side of (S-5), assume $\left(V_t^F \circ \phi\right)^r(+\infty) > \left(V_t^F \circ \phi\right)^r(-\infty)$. Otherwise the term is zero, which is bounded above by $H_t\gamma_{t-1}$. Since ϕ is nondecreasing and convex,

$$
\Upsilon(z) := \frac{\left(V_t^F \circ \phi\right)^r(z) - \left(V_t^F \circ \phi\right)^r(-\infty)}{\left(V_t^F \circ \phi\right)^r(+\infty) - \left(V_t^F \circ \phi\right)^r(-\infty)}
$$

is nondecreasing and right-continuous (right continuity follows from Lemma S-5), with $\lim_{z\to\infty} \Upsilon(z) = 1$ and $\lim_{z\to-\infty} \Upsilon(z) = 0$. Hence, Υ is a distribution function, and therefore for $\Phi = F, G$, we have that $\int_z \Upsilon(y-z) d\Phi_{t-1}(z) = \int_z \Phi_{t-1}(y-z) d\Upsilon(z)$ by usual properties of convolutions; see, e.g., Billingsley (1995, p. 266). Therefore, the first term in inequality (S-5) becomes

$$
\begin{aligned}
&\left|\int_z \left(V_t^F \circ \phi\right)^r (y-z) dF_{t-1}(z) - \int_z \left(V_t^F \circ \phi\right)^r (y-z) dG_{t-1}(z)\right| \\
&\quad = \Big(\left(V_t^F \circ \phi\right)^r(+\infty) - \left(V_t^F \circ \phi\right)^r(-\infty)\Big) \left|\int_z \Upsilon(y-z) dF_{t-1}(z) - \int_z \Upsilon(y-z) dG_{t-1}(z)\right| \\
&\quad = \Big(\left(V_t^F \circ \phi\right)^r(+\infty) - \left(V_t^F \circ \phi\right)^r(-\infty)\Big) \left|\int_z \left[F_{t-1}(y-z) - G_{t-1}(y-z)\right] d\Upsilon(z)\right| \\
&\quad \le \Big(\left(V_t^F \circ \phi\right)^r(+\infty) - \left(V_t^F \circ \phi\right)^r(-\infty)\Big) \left\|F_{t-1} - G_{t-1}\right\| \int_z d\Upsilon(z) \\
&\quad \le H_t \gamma_{t-1}. \qquad \text{(S-7)}
\end{aligned}
$$

Recall that $\left\|F_{t-1} - G_{t-1}\right\| \le \gamma_{t-1}$ implies that $\left|\left(K_{t-1}^F\right)^r(y) - \left(K_{t-1}^G\right)^r(y)\right| \le (b_{t-1} + h_{t-1})\gamma_{t-1}$ by Lemma S-3. Hence by (S-4)–(S-7) we have

$$
\left|\left(U_{t-1}^F\right)^r(y) - \left(U_{t-1}^G\right)^r(y)\right| \;\le\; (b_{t-1} + h_{t-1})\gamma_{t-1} + \nu_t + H_t\gamma_{t-1} \;=\; \nu_{t-1}.
$$

Therefore, $\left\|\left(U_{t-1}^F\right)^r - \left(U_{t-1}^G\right)^r\right\| \le \nu_{t-1}$, which completes the proof. □

Let $U_\tau^{F,G}(y) = K_\tau^F(y) = U_\tau^F(y)$ and $U_t^{F,G}(y) = K_t^F(y) + \int_z V_{t+1}^{F,G}(\phi(y-z))dF_t(z)$, for $t = 1,\ldots,\tau-1$, where $V_t^{F,G}(x) = U_t^{F,G}(\max\{x, y_t^G\})$. Here, $V_t^{F,G}$ is the value, in a problem with demand distributions $\{F_t\}$, of following the policy specified by $\{y_t^G\}$.

Proposition S-5 *Suppose* $\left\|\left(U_k^F\right)^r - \left(U_k^G\right)^r\right\| \le \mu_k$ *for* $k = t,\ldots,\tau$, *where* $\mu_k \in (0, \min\{b_k, h_k\}/3]$ *for each* k. *Let* $P_{\tau+1} = 1$ *and* $P_k = \prod_{j=k}^{\tau}(1 + 3\mu_j/\min\{b_j, h_j\})$ *for* $k = t,\ldots,\tau$. *Then* $U_k^{F,G}(y) \le U_k^F(y)P_{k+1}$ *for all* $y \in \mathbb{R}$ *and* $V_k^{F,G}(x) \le V_k^F(x)P_k$ *for all* $x \in \mathbb{R}$ *for* $k = t,\ldots,\tau$.

Proof. The proof is by backward induction on t. For $t = \tau$, $U_\tau^{F,G}(y) = U_\tau^F(y) = U_\tau^F(y)P_{\tau+1}$. Below, for any t, we will establish that

(I) $\left\|\left(U_t^F\right)^r - \left(U_t^G\right)^r\right\| \le \mu_t$ implies $U_t^F(y_t^G) \le U_t^F(y_t^F)(1 + 3\mu_t/\min\{b_t, h_t\})$ and

(II) $U_t^F(y_t^G) \le U_t^F(y_t^F)(1 + 3\mu_t/\min\{b_t, h_t\})$ and $U_t^{F,G}(y) \le U_t^F(y)P_{t+1}$ for all $y \in \mathbb{R}$ together imply $V_t^{F,G}(x) \le V_t^F(x)P_t$ for all $x \in \mathbb{R}$.

Applying (I) and (II) with $t = \tau$, we see that the proposition holds for $t = \tau$.

For the inductive step, suppose that the proposition holds for general t; that is, suppose $U_k^{F,G}(y) \le U_k^F(y)P_{k+1}$ for all $y \in \mathbb{R}$ and $V_k^{F,G}(x) \le V_k^F(x)P_k$ for all $x \in \mathbb{R}$ for $k = t,\ldots,\tau$. For $t-1$ we have

$$U_{t-1}^{F,G}(y) = K_{t-1}^F(y) + \int_z V_t^{F,G}(\phi(y-z))dF_{t-1}(z) \le K_{t-1}^F(y) + P_t \int_z V_t^F(\phi(y-z))dF_{t-1}(z) \le U_{t-1}^F(y)P_t$$

for all $y \in \mathbb{R}$. By (I) and (II), we also have that $V_{t-1}^{F,G}(x) \le V_{t-1}^F(x)P_{t-1}$ for all $x \in \mathbb{R}$. This completes an inductive proof of the proposition; all that remains is to prove (I) and (II).

For any t, we have $U_t^F(y) \ge K_t^F(y) \ge b_t(\bar{z}_t - y)^+ + h_t(y - \bar{z}_t)^+$, where $\bar{z}_t = \int_z z dF_t(z)$ and the last inequality follows from Jensen's inequality. If we can prove that y_t^G is a μ_t-point of U_t^F, then Lemma S-4 implies that $U_t^F(y_t^G) \le U_t^F(y_t^F)(1 + 3\mu_t/\min\{b_t, h_t\})$.

If $y_t^F = y_t^G$ then y_t^G is a μ_t-point of U_t^F. It remains to consider $y_t^F \ne y_t^G$. Lemma S-5 implies $\left|\left(U_t^F\right)^l(y_t^G) - \left(U_t^G\right)^l(y_t^G)\right| = \lim_{z\uparrow y_t^G}\left|\left(U_t^F\right)^r(z) - \left(U_t^G\right)^r(z)\right| \le \mu_t$. Hence, if $y_t^F < y_t^G$, then $0 \le \left(U_t^F\right)^r(y_t^F) \le \left(U_t^F\right)^l(y_t^G) \le \left(U_t^G\right)^l(y_t^G) + \mu_t \le \mu_t$. The first two inequalities follow from Lemma S-5. Hence, y_t^G is a μ_t-point of U_t^F. Likewise from Lemma S-5, it follows that if $y_t^F > y_t^G$, then $0 \ge \left(U_t^F\right)^l(y_t^F) \ge \left(U_t^F\right)^r(y_t^G) \ge \left(U_t^G\right)^r(y_t^G) - \mu_t \ge -\mu_t$, and hence y_t^G is a μ_t-point of U_t^F. Therefore,

$$U_t^F(y_t^G) \le U_t^F(y_t^F)(1 + 3\mu_t/\min\{b_t, h_t\}) \tag{S-8}$$

by Lemma S-4. That is, (I) holds.

Turning to (II), suppose that (S-8) holds and that $U_t^{F,G}(y) \le U_t^F(y) P_{t+1}$ for all $y \in \mathbb{R}$. We will show that $V_t^{F,G}(x) \le V_t^F(x) P_t$ for all $x \in \mathbb{R}$. First, note that U_t^F is convex with a minimum at y_t^F. Hence for all x between y_t^F and y_t^G, irrespective of the ordering, $U_t^F(y_t^F) \le U_t^F(x) \le U_t^F(y_t^G)$. We can divide the proof into four cases.

Case 1: If $x > \max\{y_t^G, y_t^F\}$ then

$$V_t^{F,G}(x) = U_t^{F,G}(x) \le U_t^F(x) P_{t+1} = V_t^F(x) P_{t+1} \le V_t^F(x) P_t \,.$$

Case 2: If $y_t^G < x \le y_t^F$ then

$$V_t^{F,G}(x) = U_t^{F,G}(x) \le U_t^F(x) P_{t+1} \le U_t^F(y_t^G) P_{t+1} \le U_t^F(y_t^F)(1 + 3\mu_t / \min\{b_t, h_t\}) P_{t+1} = V_t^F(x) P_t \,.$$

Case 3: If $y_t^F < x \le y_t^G$, then

$$V_t^{F,G}(x) = U_t^{F,G}(y_t^G) \le U_t^F(y_t^G) P_{t+1} \le U_t^F(y_t^F)(1 + 3\mu_t / \min\{b_t, h_t\}) P_{t+1} \le U_t^F(x) P_t = V_t^F(x) P_t \,.$$

Case 4: If $x \le \min\{y_t^F, y_t^G\}$, then

$$V_t^{F,G}(x) = U_t^{F,G}(y_t^G) \le U_t^F(y_t^G) P_{t+1} \le U_t^F(y_t^F)(1 + 3\mu_t / \min\{b_t, h_t\}) P_{t+1} = V_t^F(x) P_t \,.$$

This completes the proof of (II) and of the proposition. □

Proof of Theorem 3. Consider a fixed $\omega \in \Omega$ such that $\|F_k(\cdot) - \widehat{F}_k(\cdot, \omega)\| \le \gamma_k$ for each $k = 1, \ldots, \tau$ where $\{\gamma_k\}$ are defined as in the statement of part 1 of Theorem 3. For these values of $\{\gamma_k\}$, note that $\{\nu_k\}$ in Proposition S-4 is given by $\nu_k = \sum_{j=k}^{\tau} \epsilon'_j$. It follows from Proposition S-4 with $\{G_k(\cdot)\} = \{\widehat{F}_k(\cdot, \omega)\}$ that $\|U_k^r(\cdot) - \widehat{U}_k^r(\cdot, \omega)\| \le \sum_{j=k}^{\tau} \epsilon'_j$ for $k = 1, \ldots, \tau$. Using Proposition S-5 with $t = 1$ and $\{\mu_k\} = \{\sum_{j=k}^{\tau} \epsilon'_j\}$, $\{G_k(\cdot)\} = \{\widehat{F}_k(\cdot, \omega)\}$, and $\{V_k^{F,G}(\cdot)\} = \{\widetilde{V}_k(\cdot, \omega)\}$, we have that $0 \le \widetilde{V}_k(x, \omega) \le V_k(x) \Pi_{j=k}^{\tau} [1 + 3(\sum_{i=j}^{\tau} \epsilon'_i) / \min\{b_j, h_j\}]$ for all x and $k = 1, \ldots, \tau$.

To summarize, if $\|F_k(\cdot) - \widehat{F}_k(\cdot, \omega)\| \le \gamma_k$ for $k = 1, \ldots, \tau$, then $\|U_k(\cdot) - \widehat{U}_k(\cdot, \omega)\| \le \sum_{j=k}^{\tau} \epsilon'_j$ for $k = 1, \ldots, \tau$ and $0 \le \widetilde{V}_k(x, \omega) \le V_k(x) \prod_{j=k}^{\tau} (1 + \epsilon_j)$ for all x and $k = 1, \ldots, \tau$, where we have defined $\epsilon_j \equiv 3(\sum_{i=j}^{\tau} \epsilon'_i) / \min\{b_j, h_j\}$. Therefore

$$\begin{aligned} \mathbb{P}\Big[\widetilde{V}_k(x) \le V_k(x) \Pi_{j=k}^{\tau} (1 + \epsilon_j) \text{ for all } x \in \mathbb{R};\, k = 1, \ldots, \tau\Big] &\ge \mathbb{P}\big[\cap_{k=1,\ldots,\tau} \{\|F_k - \widehat{F}_k\| \le \gamma_k\}\big] \\ &\ge 1 - \sum_{k=1}^{\tau} 2 \exp(-2\gamma_k^2 n_k) \qquad \text{(S-9)} \end{aligned}$$

where (S-9) follows from Lemma S-1 by the same argument used to obtain (S-2).

To prove part 2, we apply part 1 with $\epsilon'_k = \frac{c\epsilon}{3(\tau^2+\tau)}$ so that $\gamma_k := \epsilon'_k/\zeta_k = \frac{c\epsilon}{3\zeta_k(\tau^2+\tau)}$. Note that $\sum_{k=1}^{\tau} \epsilon'_k \leq c/3$ because $0 < \epsilon \leq 2\log 2 < 2 \leq \tau + 1$. If $n_k \geq \frac{9(\tau^2+\tau)^2\zeta_k^2}{2c^2\epsilon^2} \log(2\tau/\delta)$, then $2\exp(-2\gamma_k^2 n_k) \leq \delta/\tau$, and it follows that $1 - \sum_{k=1}^{\tau} 2\exp(-2\gamma_k^2 n_k) \geq 1 - \delta$. Hence, if $n_k \geq \frac{9(\tau^2+\tau)^2\zeta_k^2}{2c^2\epsilon^2} \log(2\tau/\delta)$ for all k, then

$$\mathbb{P}\Big[\widetilde{V}_k(x) \leq V_k(x) \prod_{j=k}^{\tau}(1+\epsilon_j) \text{ for all } x \in \mathbb{R};\ k = 1, \ldots, \tau\Big] \geq 1 - \delta. \tag{S-10}$$

Using the preceding choice of $\{\epsilon'_k\}$ and the assumption that $\epsilon \in (0, 2\log 2]$, it follows that

$$\sum_{j=k}^{\tau} \frac{3\sum_{i=j}^{\tau} \epsilon'_i}{\min\{b_j, h_j\}} \leq \sum_{j=k}^{\tau} \frac{3\sum_{i=j}^{\tau} \epsilon'_i}{c} \leq \sum_{j=1}^{\tau} \frac{3\sum_{i=j}^{\tau} \epsilon'_i}{c} = \frac{\epsilon}{2} \leq \log 2. \tag{S-11}$$

For $j = 1, \ldots, \tau$, if $a_j > 0$ then $\log(1 + a_j) \leq a_j$, which implies $\prod_{j=k}^{\tau}(1 + a_j) \leq \exp(\sum_{j=k}^{\tau} a_j)$. If $0 \leq \sum_{j=k}^{\tau} a_j \leq \log 2$, then $\exp(\sum_{j=k}^{\tau} a_j) \leq 1 + 2\sum_{j=k}^{\tau} a_j$. From the above facts, we have that if $\{a_j > 0\}$ are such that $\sum_{j=k}^{\tau} a_j \leq \log 2$, then $\prod_{j=k}^{\tau}(1 + a_j) \leq 1 + 2\sum_{j=k}^{\tau} a_j$. Hence, by (S-11) it follows that

$$\prod_{j=k}^{\tau}(1+\epsilon_j) = \prod_{j=k}^{\tau}\Big(1 + \frac{3\sum_{i=j}^{\tau} \epsilon'_i}{\min\{b_j, h_j\}}\Big) \leq 1 + 2\sum_{j=k}^{\tau} \frac{3\sum_{i=j}^{\tau} \epsilon'_j}{\min\{b_j, h_j\}} \leq \Xi_k(\epsilon). \tag{S-12}$$

Part 2 of the theorem now follows by (S-10) and (S-12). □

S-3 Comparison with Work of Shapiro and Shapiro et al.

Shapiro (2006) and Shapiro et al. (2009, Section 5.8.2) (SDR) develop performance guarantees similar to (12) for stochastic programs with three periods, and they indicate that their results can be extended to an arbitrary number τ of periods. In this section we provide such an extension, and subsequently make comparisons between the extension and our results.

We shall mostly use the notation of Shapiro and SDR, although we replace their $\mathcal{X}_1$ by $\mathcal{C}_1$ and $\{\mathcal{X}_t : t \geq 2\}$ by $\{\mathcal{A}_t : t \geq 2\}$. Likewise, we replace x_t by a_t. We do this to reduce confusion, because x_t in our notation refers to a state, but Shapiro and SDR use it to refer to an decision (i.e., an action). Shapiro and SDR do not have an explicit notion of state in their model. Note that SDR use the notation $f_k(\cdot)$ for the period-k cost function, whereas Shapiro uses $F_k(\cdot)$. Below, we follow SDR and use $f_k(\cdot)$.

The objective is to solve the following τ-period stochastic program:

$$\inf_{a_1 \in \mathcal{C}_1} f_1(a_1) + \mathbb{E}\left[\inf_{a_2 \in \mathcal{A}(a_1,\xi_2)} f_2(a_2,\xi_2) + \mathbb{E}\left[\cdots + \mathbb{E}\left[\inf_{a_\tau \in \mathcal{A}(a_{\tau-1},\xi_\tau)} f_\tau(a_\tau,\xi_\tau)\right]\right]\right] \tag{S-13}$$

Above, $\{\xi_t\}$ are independent d_t-dimensional random variables (disturbances), $\{a_t\}$ are l_t-dimensional decision variables (the actions), $f_1(\cdot),\ldots,f_\tau(\cdot)$ are continuous single-period cost functions, and $\mathcal{C}_1$ and $\mathcal{A}_2(\cdot),\ldots,\mathcal{A}_T(\cdot)$ are the action sets.

For $k = 2,\ldots,\tau+1$, let $Q_k(a_{k-1},\xi_k)$ be the optimal value of the problem with $\tau - k + 1$ periods remaining. We then have $Q_{\tau+1}(a_\tau,\xi_{\tau+1}) \equiv 0$ and for $\ell = 2,\ldots,\tau$ we have

$$Q_\ell(a_{\ell-1},\xi_\ell) = \inf_{a_\ell \in \mathcal{A}_\ell(a_{\ell-1},\xi_\ell)} \left\{ f_\ell(a_\ell,\xi_\ell) + \mathbb{E}[Q_{\ell+1}(a_\ell,\xi_{\ell+1})] \right\}.$$

Let

$$f(a_1) = f_1(a_1) + \mathbb{E}[Q_2(a_1,\xi_2)].$$

The problem (S-13) can now be re-expressed as:

$$\inf_{a_1 \in \mathcal{C}_1} f(a_1). \tag{S-14}$$

That is, the objective is to find the best time-1 action.

We impose the following assumptions [compare with assumptions (M′1)–(M′8) on pages 227–228 of SDR].

(B1) For every $a_1 \in \mathcal{C}_1$, the expectation $\mathbb{E}[Q_2(a_1,\xi_2)]$ is well-defined and finite-valued.

(B2) The random vectors $\xi_2,\ldots,\xi_\tau$ are independent.

(B3) The set $\mathcal{C}_1$ has a finite diameter D_1.

(B4) For each $k = 2,\ldots,\tau-1$, there is a set $\mathcal{C}_k$ with finite diameter D_k such that for every a_{k-1} and a.e. ξ_k, the set $\mathcal{A}_k(a_{k-1},\xi_k)$ is contained in $\mathcal{C}_k$.

(B5) For each $k = 1,\ldots,\tau-1$, there is a constant $L_k > 0$ such that

$$\left|Q_{k+1}(a'_k,\xi_{k+1}) - Q_{k+1}(a_k,\xi_{k+1})\right| \le L_k \left\|a'_k - a_k\right\|$$

for all $a'_k, a_k \in \mathcal{C}_k$ and a.e. ξ_{k+1}.

(B6) For each $k = 1,\ldots,\tau-1$, there is a constant $\sigma_k > 0$ such that for any $a_k \in \mathcal{A}_k(a_{k-1},\xi_k)$ and all $a_{k-1} \in \mathcal{C}_{k-1}$ and a.e. ξ_k it holds that

$$M_{k,a_k}(t) \le \exp\left(\sigma_k^2 t^2/2\right) \qquad \text{for all } t \in \mathbb{R}$$

where $M_{k,a_k}(t)$ is the moment-generating function of $Q_{k+1}(a_k, \xi_{k+1}) - \mathbb{E}[Q_{k+1}(a_k, \xi_{k+1})]$.

Consider N_k i.i.d. samples of ξ_{k+1} denoted by $\xi_{k+1}^1, \dots, \xi_{k+1}^{N_k}$. We focus on what SDR refer to on page 221 as *identical conditional sampling*. A similar analysis is possible for what they term *independent conditional sampling*.

For each fixed $k \in \{1, \dots, \tau - 1\}$, recursively define

$$\widehat{Q}_\ell^k(a_{\ell-1}, \xi_\ell) = \begin{cases} Q_\ell(a_{\ell-1}, \xi_\ell), & \ell = k+1, \dots, \tau; \\ \inf_{a_\ell \in \mathcal{A}_\ell(a_{\ell-1}, \xi_\ell)} \left\{ f_\ell(a_\ell, \xi_\ell) + \frac{1}{N_\ell} \sum_{i_\ell=1}^{N_\ell} \widehat{Q}_{\ell+1}^k(a_\ell, \xi_{\ell+1}^{i_\ell}) \right\}, & \ell = 2, \dots, k. \end{cases}$$

Note that

$$\widehat{Q}_k^k(a_{k-1}, \xi_k) = \inf_{a_k \in \mathcal{A}_k(a_{k-1}, \xi_k)} \left\{ f_k(a_k, \xi_k) + \frac{1}{N_k} \sum_{i_k=1}^{N_k} Q_{k+1}(a_k, \xi_{k+1}^{i_k}) \right\}.$$

Next, let $\hat{g}^0(a_1) = f(a_1)$ and

$$\hat{g}^k(a_1) = f_1(a_1) + \frac{1}{N_1} \sum_{i_1=1}^{N_1} \widehat{Q}_2^k(a_1, \xi_2^{i_1}), \qquad k = 1, \dots, \tau - 1.$$

The quantity $\hat{g}^k(a_1)$ is an approximation of the value (as a function of the action a_1) at the beginning of the time horizon (i.e., just before implementing the first action a_1) wherein the expectations with respect to the distributions of $\xi_2, \dots, \xi_{k+1}$ are approximated by sample averages.

If we extend the ideas of Shapiro and SDR to allow $\tau \geq 3$, then we would approximate (S-14) by solving

$$\inf_{a_1 \in \mathcal{C}_1} \hat{g}^{\tau-1}(a_1). \tag{S-15}$$

This is essentially the same as solving what we call the empirical MDP. To connect the above with the notation used in Shapiro and SDR, when $\tau = 3$ we have

$$\hat{g}^2(a_1) = \tilde{f}_{N_1,N_2}(a_1) \quad \text{and} \quad \hat{g}^1(a_1) = \hat{f}_{N_1}(a_1).$$

Note that to solve (S-15) we need access to the true single-period expected cost functions $\{f_k(\cdot)\}$. Such functions will generally depend upon the disturbance distributions, which we assume to be unknown. It is possible to further modify the approach of Shapiro and SDR to use the sample average approximation to estimate the single-period cost functions. We do not provide details.

If $\sup_{a_1 \in \mathcal{C}_1} |f(a_1) - \hat{g}^{\tau-1}(a_1)| \leq \epsilon/3$ and if $a_1^* \in \mathcal{C}_1$ is chosen to satisfy

$$\hat{g}^{\tau-1}(a_1^*) \leq \inf_{a_1 \in \mathcal{C}_1} \hat{g}^{\tau-1}(a_1) + \epsilon/3, \tag{S-16}$$

then $f(a_1^*) \le \inf_{a_1 \in \mathcal{C}_1} f(a_1) + \epsilon$. Hence, if we choose a_1^* as in (S-16) then

$$\begin{aligned}
\mathbb{P}\left[f(a_1^*) > \inf_{a_1 \in \mathcal{C}_1} f(a_1) + \epsilon\right] &\le \mathbb{P}\left[\sup_{a_1 \in \mathcal{C}_1} \left|f(a_1) - \hat{g}^{\tau-1}(a_1)\right| > \epsilon/3\right] \\
&= \mathbb{P}\left[\sup_{a_1 \in \mathcal{C}_1} \left|\sum_{k=1}^{\tau-1} \left(\hat{g}^{k-1}(a_1) - \hat{g}^{k}(a_1)\right)\right| > \epsilon/3\right] \\
&\le \sum_{k=1}^{\tau-1} \mathbb{P}\left[\sup_{a_1 \in \mathcal{C}_1} \left|\hat{g}^{k-1}(a_1) - \hat{g}^{k}(a_1)\right| > \epsilon/(3(\tau-1))\right] . \qquad \text{(S-17)}
\end{aligned}$$

Consider fixed $k \in \{2, \dots, \tau - 1\}$. We have

$$\begin{aligned}
&\sup_{a_1 \in \mathcal{C}_1} \left| \hat{g}^{k-1}(a_1) - \hat{g}^k(a_1) \right| \\
&\quad = \sup_{a_1 \in \mathcal{C}_1} \left| \left(f_1(a_1) + \frac{1}{N_1} \sum_{i_1=1}^{N_1} \widehat{Q}_2^{k-1}(a_1, \xi_2^{i_1}) \right) - \left(f_1(a_1) + \frac{1}{N_1} \sum_{i_1=1}^{N_1} \widehat{Q}_2^{k}(a_1, \xi_2^{i_1}) \right) \right| \\
&\quad = \sup_{a_1 \in \mathcal{C}_1} \left| \frac{1}{N_1} \sum_{i_1=1}^{N_1} \widehat{Q}_2^{k-1}(a_1, \xi_2^{i_1}) - \frac{1}{N_1} \sum_{i_1=1}^{N_1} \widehat{Q}_2^{k}(a_1, \xi_2^{i_1}) \right| \\
&\quad \leq \frac{1}{N_1} \sum_{i_1=1}^{N_1} \sup_{a_1 \in \mathcal{C}_1} \left| \widehat{Q}_2^{k-1}(a_1, \xi_2^{i_1}) - \widehat{Q}_2^{k}(a_1, \xi_2^{i_1}) \right| \\
&\quad = \frac{1}{N_1} \sum_{i_1=1}^{N_1} \sup_{a_1 \in \mathcal{C}_1} \left| \inf_{a_2 \in \mathcal{A}_2(a_1, \xi_2^{i_1})} \left\{ f_2(a_2, \xi_2^{i_1}) + \frac{1}{N_2} \sum_{i_2=1}^{N_2} \widehat{Q}_3^{k-1}(a_2, \xi_3^{i_2}) \right\} \right. \\
&\qquad\qquad \left. - \inf_{a_2 \in \mathcal{A}_2(a_1, \xi_2^{i_1})} \left\{ f_2(a_2, \xi_2^{i_1}) + \frac{1}{N_2} \sum_{i_2=1}^{N_2} \widehat{Q}_3^{k}(a_2, \xi_3^{i_2}) \right\} \right| \\
&\quad \leq \frac{1}{N_1} \sum_{i_1=1}^{N_1} \sup_{a_1 \in \mathcal{C}_1} \sup_{a_2 \in \mathcal{A}_2(a_1, \xi_2^{i_1})} \left| \frac{1}{N_2} \sum_{i_2=1}^{N_2} \left(\widehat{Q}_3^{k-1}(a_2, \xi_3^{i_2}) - \widehat{Q}_3^{k}(a_2, \xi_3^{i_2}) \right) \right| \\
&\quad \leq \frac{1}{N_1} \sum_{i_1=1}^{N_1} \sup_{a_1 \in \mathcal{C}_1} \sup_{a_2 \in \mathcal{C}_2} \left| \frac{1}{N_2} \sum_{i_2=1}^{N_2} \left(\widehat{Q}_3^{k-1}(a_2, \xi_3^{i_2}) - \widehat{Q}_3^{k}(a_2, \xi_3^{i_2}) \right) \right| \\
&\quad = \sup_{a_2 \in \mathcal{C}_2} \left| \frac{1}{N_2} \sum_{i_2=1}^{N_2} \widehat{Q}_3^{k-1}(a_2, \xi_3^{i_2}) - \frac{1}{N_2} \sum_{i_2=1}^{N_2} \widehat{Q}_3^{k}(a_2, \xi_3^{i_2}) \right| \\
&\qquad \vdots \\
&\qquad \vdots \\
&\quad \leq \sup_{a_{k-1} \in \mathcal{C}_{k-1}} \left| \frac{1}{N_{k-1}} \sum_{i_{k-1}=1}^{N_{k-1}} \widehat{Q}_k^{k-1}(a_{k-1}, \xi_k^{i_{k-1}}) - \frac{1}{N_{k-1}} \sum_{i_{k-1}=1}^{N_{k-1}} \widehat{Q}_k^{k}(a_{k-1}, \xi_k^{i_{k-1}}) \right| \\
&\quad = \sup_{a_{k-1} \in \mathcal{C}_{k-1}} \left| \frac{1}{N_{k-1}} \sum_{i_{k-1}=1}^{N_{k-1}} Q_k(a_{k-1}, \xi_k^{i_{k-1}}) - \frac{1}{N_{k-1}} \sum_{i_{k-1}=1}^{N_{k-1}} \widehat{Q}_k^{k}(a_{k-1}, \xi_k^{i_{k-1}}) \right| \\
&\qquad \vdots \\
&\quad \leq \sup_{a_k \in \mathcal{C}_k} \left| \mathbb{E}\left[Q_{k+1}(a_k, \xi_{k+1}) \right] - \frac{1}{N_k} \sum_{i_k=1}^{N_k} Q_{k+1}(a_k, \xi_{k+1}^{i_k}) \right| .
\end{aligned}$$

Hence, for $k \in \{2, \dots, \tau - 1\}$ we have

$$\sup_{a_1 \in \mathcal{C}_1} \left| \hat{g}^{k-1}(a_1) - \hat{g}^k(a_1) \right| \leq \sup_{a_k \in \mathcal{C}_k} \left| \mathbb{E}\left[Q_{k+1}(a_k, \xi_{k+1}) \right] - \frac{1}{N_k} \sum_{i_k=1}^{N_k} Q_{k+1}(a_k, \xi_{k+1}^{i_k}) \right| .$$

Similarly, it can be verified that

$$\sup_{a_1 \in \mathcal{A}} \left|\hat{g}^0(a_1) - \hat{g}^1(a_1)\right| \le \sup_{a_1 \in \mathcal{C}_1} \left| \mathbb{E}\left[Q_2(a_1, \xi_2)\right] - \frac{1}{N_1} \sum_{i_1=1}^{N_1} Q_2(a_1, \xi_2^{i_1}) \right| .$$

Therefore, for $k \in \{1, \dots, \tau - 1\}$ we have

$$\begin{aligned}
&\mathbb{P}\left[\sup_{a_1 \in \mathcal{C}_1} \left|\hat{g}^{k-1}(a_1) - \hat{g}^k(a_1)\right| > \epsilon/(3(\tau-1))\right] \\
&\le \mathbb{P}\left[\sup_{a_k \in \mathcal{C}_k} \left| \mathbb{E}\left[Q_{k+1}(a_k, \xi_{k+1})\right] - \frac{1}{N_k} \sum_{i_k=1}^{N_k} Q_{k+1}(a_k, \xi_{k+1}^{i_k}) \right| > \epsilon/(3(\tau-1))\right] \\
&\le O(1) \left(\frac{3D_k L_k(\tau-1)}{\epsilon}\right)^{n_k} \exp\left(-\frac{N_k \epsilon^2}{48(\tau-1)^2 \sigma_k^2}\right) , \qquad \text{(S-18)}
\end{aligned}$$

where the inequality (S-18) follows from Theorem 1 of Shapiro (2006).

Combining the above with (S-17) we see that

$$\mathbb{P}\left[f(a_1^*) > \inf_{a_1 \in \mathcal{C}_1} f(a_1) + \epsilon\right] \le \sum_{k=1}^{\tau-1} O(1) \left(\frac{3D_k L_k(\tau-1)}{\epsilon}\right)^{l_k} \exp\left(-\frac{N_k \epsilon^2}{48(\tau-1)^2 \sigma_k^2}\right) . \qquad \text{(S-19)}$$

Given any $\delta > 0$, we can make

$$\mathbb{P}\left[f(a_1^*) > \inf_{a_1 \in \mathcal{C}_1} f(a_1) + \epsilon\right] \le \delta \qquad \text{(S-20)}$$

by taking $N_1, \dots, N_{\tau-1}$ to satisfy

$$O(1) \left(\frac{3D_k L_k(\tau-1)}{\epsilon}\right)^{l_k} \exp\left(-\frac{N_k \epsilon^2}{48(\tau-1)^2 \sigma_k^2}\right) \le \delta/(\tau-1) .$$

Equivalently, to ensure (S-20) each N_k should satisfy

$$N_k \ge \frac{48(\tau-1)^2 \sigma_k^2}{\epsilon^2} \left[\log\left(\frac{\tau-1}{\delta}\right) + \log O(1) + l_k \log\left(\frac{3D_k L_k(\tau-1)}{\epsilon}\right)\right] . \qquad \text{(S-21)}$$

Observe that the bound on N_k in (S-21) is of order $\sigma_k^2 \tau^2 \log \tau$. The quantity σ_k^2 is related to the variability of the period-k value function; see assumption (B6). The quantity σ_k^2 depends upon τ.

Next, we will describe the application of the extended Shapiro and SDR result to inventory problems such as those covered in Section 5. As mentioned above, the extended result cannot be applied directly to such inventory problems because Shapiro and SDR assume access to the expected cost functions $f_k(\cdot)$, which depend upon the unknown demand distributions. It is a straightforward exercise to further extend their approach to also use sampling to approximate the cost functions.

We will ignore this issue, and simply allow the Shapiro and SDR method to have access to these functions.

To use the Shapiro and SDR framework for inventory problems, we interpret the period-k decision to be the inventory level before period-k demand ξ_{k+1} arrives. To make the connection with the notation through the rest of our paper, we have that $\xi_{k+1} = D_k$ for $k = 1, \ldots, \tau$. Assuming the system starts with initial inventory x, we have $\mathcal{C}_1 = [x, \infty)$ and $\mathcal{A}_k(a_{k-1}, \xi_k) = [a_{k-1} - \xi_k, \infty)$. The single-period cost functions are $f_1(a_1) = \mathbb{E}\left[b(a_1 - \xi_2)^+ + h(\xi_2 - a_1)^+\right]$ and $f_k(a_k, \xi_k) = f_k(a_k) = \mathbb{E}\left[b(a_k - \xi_{k+1})^+ + h(\xi_{k+1} - a_k)^+\right]$ for $k = 2, \ldots, \tau$ where b and h are single-period per-unit holding and backorder cost rates. Observe that the period-k cost is expressed as a function only of that period's chosen inventory level.

At this point, the preceding setup violates assumptions (B3) and (B4). However, if we have bounds on the support of the demand distributions (as we assume in our Theorem 2), then we can a priori bound the inventory levels. In particular, we would never order so that there would, with certainty, be inventory remaining at the end of the current period (unless it cannot be avoided). With this observation, in the setting of Corollary 1, such an a priori upper bound on a_k is $\bar{a}_k = \max\{x, \beta\}$. For simplicity, we shall hereafter assume $x \le \beta$. Likewise, we would never allow inventory to be negative after ordering but prior to realization of demand. From these observations, we can bound the inventory levels (with no loss of optimality) so that assumptions (B3) and (B4) hold.

Again, to make a connection between the notation of this section and that used in the rest of the paper, for period τ we have $f_\tau(a_\tau) = K_\tau(a_\tau) = U_\tau(a_\tau)$ and $Q_\tau(a_{\tau-1}, \xi_\tau) = V_\tau(\phi(a_{\tau-1} - \xi_\tau))$. Likewise, for periods $k = 1, \ldots, \tau - 1$ we have $f_k(a_k) = K_k(a_k)$ and $f_k(a_k) + \mathbb{E}[Q_{k+1}(a_k, \xi_{k+1})] = U_k(a_k)$, and for periods $k = 2, \ldots, \tau - 1$ we have $Q_k(a_{k-1}, \xi_k) = V_k(\phi(a_{k-1} - \xi_k))$. For period 1 the problem in (S-14) satisfies $\inf_{a_1 \in \mathcal{C}_1} f(a_1) = \inf_{a_1 \ge x} U_1(a_1) = V_1(x)$, where x is the fixed starting inventory (dependence upon x is not shown in the notation $f(\cdot)$ and $\mathcal{C}_1$ because x is fixed).

Fix $k \in \{1, \ldots, \tau - 1\}$ and let $m = h \cdot (\tau - k)$. If the support of demand ξ_{k+1} is bounded above by β and below by 0, then

$$\begin{aligned} M_{k,a_k}(t) &= \mathbb{E} \exp\Big(t\big(Q_{k+1}(a_k, \xi_{k+1}) - \mathbb{E}[Q_{k+1}(a_k, \xi_{k+1})]\big)\Big) \\ &\le \exp\left(\frac{(Q_{k+1}(a_k, 0) - Q_{k+1}(a_k, \beta))^2 t^2}{2} \right) \end{aligned} \tag{S-22}$$

where (S-22) follows from Theorem 7.63 of SDR. (The theorem states that if $\mathbb{E}[Y] = 0$ and $y' \le Y \le y''$ for constants y', y'', then $\mathbb{E}[\exp(tY)] \le \exp(t^2(y''-y')/8)$ for all t. Note that $\mathbb{E}\big[Q_{k+1}(a_k, \xi_{k+1}) - \mathbb{E}[Q_{k+1}(a_k, \xi_{k+1})]\big] = 0$ and $Q_{k+1}(a_k, \beta) - Q_{k+1}(a_k, 0) \le Q_{k+1}(a_k, \xi_{k+1}) - \mathbb{E}[Q_{k+1}(a_k, \xi_{k+1})] \le Q_{k+1}(a_k, 0) - Q_{k+1}(a_k, \beta)$.) From the monotonicity and convexity of the value function and Lemma 2 we have $0 \le Q_{k+1}(a_k, 0) - Q_{k+1}(a_k, \beta) \le m\beta$. So, we may take σ_k^2 in assumption (B6) to be

$$\sigma_k^2 := m^2\beta^2 = h^2(\tau - k)^2\beta^2 = O(\tau^2)\,. \tag{S-23}$$

Hence, the bound on N_k in (S-21) is of order $\tau^4 \log \tau$.

Is it possible to find a value for σ_k^2 that is smaller than (S-23) and that satisfies assumption (B6)? At this point, we cannot answer this question. However, the following shows that it is in general not possible to find a value for σ_k^2 smaller than $m^2\beta^2/4$.

Let

$$R_{k,a_k}(\xi_{k+1}) := Q_{k+1}(a_k, \xi_{k+1}) - Q_{k+1}(a_k, 0) + m\xi_{k+1} \quad \text{ for } \xi_{k+1} \in [0, \infty)$$

Then $R_{k,a_k}(\xi_{k+1}) \downarrow 0$ as $a_k \to \infty$ for each $\xi_{k+1} \in [0, \infty)$ and $0 \le R_{k,a_k}(\xi_{k+1}) \le m\xi_{k+1}$ for each $\xi_{k+1} \in [0, \infty)$. Moreover, $R_{k,a_k}(\xi_{k+1})$ is convex and increasing in ξ_{k+1} for each a_k.

Suppose as in the setting of Corollary 1 that the support of ξ_{k+1} is contained in $[0, \beta]$. A Taylor expansion gives us

$$\begin{aligned}
M_{k,a_k}(t) &= \mathbb{E}\exp\Big(t\big(Q_{k+1}(a_k, \xi_{k+1}) - \mathbb{E}[Q_{k+1}(a_k, \xi_{k+1})]\big)\Big) \\
&= 1 + \frac{t^2}{2}\mathbb{E}\Big[\big((Q_{k+1}(a_k, \xi_{k+1}) - \mathbb{E}[Q_{k+1}(a_k, \xi_{k+1})]\big)^2\Big] + o(t^2) \\
&= 1 + \frac{t^2}{2}\Big[m^2\mathrm{Var}(\xi_{k+1}) + \mathrm{Var}(R_{k,a_k}(\xi_{k+1})) \\
&\qquad + 2m\mathbb{E}\big[\{\xi_{k+1} - \mathbb{E}[\xi_{k+1}]\}\{R_{a_k}(\xi_{k+1}) - \mathbb{E}[R_{k,a_k}(\xi_{k+1})]\}\big]\Big] + o(t^2)
\end{aligned}$$

where $\lim_{t\to 0} o(t^2)/t^2 = 0$.

The Cauchy-Schwarz Inequality implies

$$\Big|\mathbb{E}\big[\{\xi_{k+1} - \mathbb{E}[\xi_{k+1}]\}\{R_{k,a_k}(\xi_{k+1}) - \mathbb{E}[R_{k,a_k}(\xi_{k+1})]\}\big]\Big| \le \sqrt{\mathrm{Var}(\xi_{k+1})\mathrm{Var}(R_{k,a_k}(\xi_{k+1}))}\,.$$

Moreover, $\mathrm{Var}(R_{k,a_k}(\xi_{k+1})) \to 0$ as $a_k \to \infty$ by the Dominated Convergence Theorem. Consequently, for any $\epsilon > 0$, there exists $a_k(\epsilon)$ such that

$$\mathrm{Var}(R_{k,a_k(\epsilon)}(\xi_{k+1})) + 2m\mathbb{E}\big[\{\xi_{k+1} - \mathbb{E}[\xi_{k+1}]\}\{R_{k,a_k(\epsilon)}(\xi_{k+1}) - \mathbb{E}[R_{k,a_k(\epsilon)}(\xi_{k+1})]\}\big] > -\epsilon\,,$$

and therefore

$$M_{k,a_k(\epsilon)}(t) \geq 1 + \frac{t^2}{2}\Big[m^2 \mathrm{Var}(\xi_{k+1}) - \epsilon\Big] + o(t^2)\,. \tag{S-24}$$

Also by a Taylor expansion, we have that

$$\exp\left(\frac{\sigma_k^2 t^2}{2}\right) = 1 + \frac{\sigma_k^2 t^2}{2} + o(t^2)\,. \tag{S-25}$$

If we know only that the support of ξ_{k+1} is in $[0,\beta]$, then the variance of ξ_{k+1} can be as large as $\beta^2/4$ (when the distribution of ξ_{k+1} places mass 1/2 at 0 and at β). Hence, in view of (S-24) and (S-25), we must have

$$\sigma_k^2 \geq \frac{m^2\beta^2}{4} - \epsilon \tag{S-26}$$

in order for assumption (B6) to hold. (If (S-26) does not hold, then we will have $\sup_{a_k} M_{k,a_k}(t) > \exp\left(\sigma_k^2 t^2/2\right)$ for t in a neighborhood of 0.) Moreover, $\epsilon > 0$ is arbitrary, and consequently we need $\sigma_k^2 \geq m^2\beta^2/4 = h^2(\tau - k)^2\beta^2/4 = O(\tau^2)$. Therefore, if we assume only that ξ_{k+1} has bounded support, then direct application of the extension of the result of Shapiro and SDR cannot improve on the asymptotic growth rate of the sufficient numbers of samples provided by Corollary 1. In particular, the bounds in (S-21) are of order $\tau^4 \log \tau$. Our Corollary 1 provides bounds on the number of samples that are of order $\tau^4 \log \tau$.

One might possibly be able to improve the bounds in (S-21) by identifying and a priori eliminating suboptimal actions (aside from those above $\bar{a}_k$ and below 0). It is not apparent how to do this without imposing additional conditions on the true problem. Similar statements apply to Corollary 1 as well.

S-4 Computation for Inventory Models

In this section we explain how to solve exactly (without any truncation or discretization of the state or action spaces) the empirical MDP for inventory problems such those described in Section 5. This is of interest because the state and action spaces of the true MDP are both uncountably infinite. At a high level, we are able to do this because the functions $\widehat{K}_t(\cdot)$, $\widehat{U}_t(\cdot)$, and $\widehat{V}_t(\cdot)$ are piecewise linear and convex with a finite number of "kinks" (a kink is a point at which the left and right derivatives are different). The kinks as well as the left and right derivatives can be determined from the finite set of demand samples. The exact method described below will not be practical when the number of distinct demand samples is very large.

Let f^r (f^l) denote the right (left) derivative of a function f. For sets A and B, let $A \oplus B$ denote set addition; i.e., $A \oplus B = \{a + b : a \in A, b \in B\}$. For problems with backorders, the algorithm below uses the basic relations

$$\widehat{U}_{t-1}(y) = \widehat{K}_{t-1}(y) + \frac{1}{n_{t-1}} \sum_{i=1}^{n_{t-1}} \widehat{V}_t(y - Z_{t-1}^i)$$

$$\widehat{y}_{t-1} = \min\{y : \widehat{U}_{t-1}^r(y) \geq 0\}$$

$$\widehat{V}_{t-1}(x) = \widehat{U}_{t-1}(\max\{\widehat{y}_{t-1}, x\})$$

for $t = 2, \ldots, \tau + 1$ [with $\widehat{V}_{\tau+1}(\cdot) = 0$]. See equations (15), (16), and (19) and Lemma 1.

Note that we can express $\widehat{K}(\cdot)$ as

$$\widehat{K}_{t-1}(y) = \int_z k_{t-1}(y - z) d\widehat{F}_{t-1}(z) = \frac{1}{n_{t-1}} \sum_{i=1}^{n_{t-1}} b_{t-1}(Z_{t-1}^i - y)^+ + h_{t-1}(y - Z_{t-1}^i)^+ . \tag{S-27}$$

We make the following observations:

- $\widehat{K}_{t-1}$ is a piecewise linear convex function with a finite number of kinks. The kinks occur at the realized demand samples. Let $\mathcal{K} = \{m_1, \ldots, m_p\}$ be the distinct ordered kinks of $\widehat{K}_{t-1}$; there are no repetitions in $\mathcal{K}$ and $m_1 < m_2 < \cdots < m_p$. Note that we do not show the dependence on time to avoid cumbersome notation. Also, let α_i denote the number of times m_i occurs in the samples $\{Z_{t-1}^j\}$ from time period $t-1$; i.e., $\alpha_i = |\{j : Z_{t-1}^j = m_i\}|$. For $i = 1, \ldots, p$, let d_i be the right derivative of $\widehat{K}_{t-1}$ at m_i and d_0 be the left derivative at m_1; i.e., $d_i = \widehat{K}_{t-1}^r(m_i)$ and $d_0 = \widehat{K}_{t-1}^l(m_1)$. Then by (S-27) we have

$$d_i = \frac{n_{t-1} - |\{i : Z_{t-1}^j \leq m_i\}|}{n_{t-1}} (-b_{t-1}) + \frac{|\{i : Z_{t-1}^j \leq m_i\}|}{n_{t-1}} h_{t-1} \tag{S-28}$$

  for $i = 1, \ldots, p$ and $d_0 = -b_{t-1}$.

- We can rewrite the expression for $\widehat{U}_{t-1}(y)$ as

$$\widehat{U}_{t-1}(y) := \widehat{K}_{t-1}(y) + \frac{1}{n_{t-1}} \sum_{i=1}^{p} \alpha_i \widehat{V}_t(y - m_i). \tag{S-29}$$

  By inductive arguments, it is possible to show that $\widehat{V}_t$ and $\widehat{U}_{t-1}$ are also piecewise linear convex functions with a finite number of kinks. Let $\mathcal{V} = \{v_1, \ldots, v_q\}$ be the ordered sequence of distinct kinks of $\widehat{V}_t$. For $j = 1, \ldots, q$, let f_j be the right derivative of $\widehat{V}_t$ at v_j and f_0 denote the left derivative at v_1.

Let $\mathcal{U} = \{u_1, \ldots, u_r\}$ be the ordered sequence of distinct kinks for $\widehat{U}_{t-1}$. For $i = 1, \ldots, r$, let g_i be the right derivative of $\widehat{U}_{t-1}$ at u_i and g_0 be the left derivative at u_1. From (S-29) we see that every kink of $\widehat{U}_{t-1}$ will either be a kink of $\widehat{K}_{t-1}$ or a kink of $w_i(\cdot) = \widehat{V}_t(\cdot - m_i)$ for some $i = 1, \ldots, p$. The kinks of w_i are $\{m_i + v_1, \ldots, m_i + v_q\}$. Hence, $\mathcal{U}$ is given by the set $\mathcal{K} \cup (\mathcal{K} \oplus \mathcal{V})$ with its (distinct) elements ordered smallest to largest.

From the preceding observations we get the following algorithm for solving the empirical inventory MDP. In the algorithm we amend the above notation to show dependence on time.

Algorithm 1.

1. Set $t = \tau$ and $\mathcal{U}_\tau = \mathcal{K}_\tau$.

2. Let $\mathcal{V}_t = \{u_{i,t} \in \mathcal{U}_t : g_{i,t} \geq 0\}$ and $\widehat{y}_t = \min \mathcal{V}_t$.

3. Let $\mathcal{U}_{t-1}$ be the ordered elements of $\mathcal{K}_{t-1} \cup (\mathcal{K}_{t-1} \oplus \mathcal{V}_t)$. Obtain $\{g_{0,t-1}, g_{1,t-1}, \ldots, g_{r_{t-1},t-1}\}$ with (S-28) and (S-29).

4. If $t \geq 2$ let $t = t - 1$ and go to step 2. Otherwise, stop.

The above yields a finite algorithm for solving the inventory problem. Next we provide some comments on how to efficiently perform step 3 of the above algorithm. Below, we again suppress dependence on time in the interest of notational simplicity.

From (S-29) the right derivative of $\widehat{U}_{t-1}$ at u_k is

$$g_k = \widehat{U}^r_{t-1}(u_k) = \widehat{K}^r_{t-1}(u_k) + \frac{1}{n_{t-1}} \sum_{i=1}^{p} \alpha_i \widehat{V}^r_t(u_k - m_i).$$

Fix k. To compute g_k, we do the following. For each $i = 1, \ldots, p$, let $h(i)$ be such that $v_{h(i)} \leq u_k - m_i < v_{h(i)+1}$ if $v_1 \leq u_k - m_i < v_q$, otherwise $h(i) = 0$ if $u_k - m_i < v_1$ and $h(i) = q$ if $u_k - m_i \geq v_q$. Similarly, let $h(0)$ be such that $m_{h(0)} \leq u_k < m_{h(0)+1}$ if $m_1 \leq u_k < m_p$ else $h(0) = 0$ if $u_k < m_1$ and $h(0) = p$ if $u_k \geq m_p$. For $i = 1, \ldots, p$, one can think of u_k as being in the interval $[m_i + v_{h(i)}, m_i + v_{h(i)+1})$ where the function $\widehat{V}_t(y - m_i)$ has a constant right derivative of $f_{h(i)}$. Similarly u_k is in the interval $[m_{h(0)}, m_{h(0)+1})$ where the function $\widehat{K}_{t-1}$ has a constant right derivative of $d_{h(0)}$.

Then, the right derivative at u_k is given by

$$g_k = d_{h(0)} + \frac{1}{n_{t-1}} \sum_{i=1}^{p} \alpha_i f_{h(i)}.$$

Based upon this, the following algorithm performs Step 3 of Algorithm 1.

Algorithm 2.

1. Set $h(i) = 0$ for all $i = 1, \ldots, p$ and $h(0) = 0$. Set $k = 1$, $g_0 = d_0 + \frac{1}{n_{t-1}} \sum_{i=1}^{p} \alpha_i f_0 = d_0 + f_0$. Set $L \leftarrow \{m_1, m_i + v_1, i = 1, \ldots, p\}$.

2. Pick all the smallest element(s) in L. For each smallest element $m_i + v_j$, set $u_k = m_i + v_j$ and $h(i) = j$. Remove $m_i + v_{h(i)}$ from L. Add $m_i + v_{h(i)+1}$ to L if $h(i) + 1 \leq q$. For each smallest element m_i, set $u_k = m_i$ and $h(0) = i$. Remove $m_{h(0)}$ from L. Add $m_{h(0)+1}$ to L if $h(0) + 1 \leq p$.

3. Set right derivative at the kink u_k to be $g_k = d_{h(0)} + \frac{1}{n_{t-1}} \sum_{i=1}^{p} \alpha_i f_{h(i)}$.

4. Set $k \leftarrow k + 1$ and repeat Step 2.

The following remarks show that $\{u_k\}$ computed by Algorithm 2 is a strictly monotone increasing sequence and that $\{u_k\} = \mathcal{K} \cup (\mathcal{K} \oplus \mathcal{V})$.

1 The elements of $\mathcal{K}$ are distinct as are those of $\mathcal{V}$. Hence, we can make the following observations. For any $i = 1, \ldots, p$ and $j \neq j'$, we have $m_i + v_j \neq m_i + v_{j'}$. For any $i \neq i'$ we have $m_i \neq m_{i'}$.

2 The smallest elements removed from the list L are replaced by elements that are strictly larger because for $i = 1, \ldots, p$, $m_i + v_{h(i)} < m_i + v_{h(i)+1}$ and $m_{h(0)} < m_{h(0)+1}$. As a result, the smallest element of L in the $(k+1)$-th iteration of Algorithm 2 will be strictly larger than the smallest element of L in the k-th iteration.

3 At every iteration either $h(i)$ goes up by 1 for some value(s) of $i = 1, \ldots, p$ and/or $h(0)$ goes up by 1. Hence

[a] Since $h(i), i = 1, \ldots, p$ and $h(0)$ start at 0, every element of $\mathcal{K} \cup (\mathcal{K} \oplus \mathcal{V})$ will be included in the list L in some iteration.

[b] Since $h(i), i = 1, \ldots, p$ and $h(0)$ cannot exceed q and p respectively, the algorithm terminates in finite time.

The algorithm's utility goes beyond implementing Step 2 of Algorithm 1 as the following illustrate:

1. It is possible to implement Step 2 of the Algorithm 1 within Algorithm 2, simply by adding one line of code in Step 3 of Algorithm 2. We can then compute $\widehat{y}_t$ and the kinks and right derivatives of $\widehat{V}_{t-1}(y)$, which we denote by $\{u'_{k'}\}$ and $\{g'_{k'}\}$ respectively. (Note that the left derivative of $\widehat{V}_{t-1}$ at u'_1 is $g'_0 = 0$.) We can do this by initializing $k' = 1$, $\hat{y} = +\infty$, and including the following in Step 3 of Algorithm 2:

   If $g_k \geq 0$ then $u'_{k'} = u_k$, $g'_{k'} = g_k$, $\hat{y} = \min\{\hat{y}, g_k\}$ and $k' \leftarrow k' + 1$.

2. Building on the inclusion of Step 2 of Algorithm 1 within Algorithm 2, we can address the case of lost sales models with the minor addition (also in Step 3 of Algorithm 2) to the code listed below. In the lost-sales case we have $\widehat{U}_{t-1}(y) := \widehat{K}_{t-1}(y) + \frac{1}{n_{t-1}} \sum_{i=1}^{p} \alpha_i \widehat{V}_t((y - m_i)^+)$. Hence in the lost-sales case, we need to maintain the kinks of $\widehat{V}_t(y^+)$ which are directly derived from the kinks of $\widehat{V}_t(y)$. (Note that the functions $\widehat{U}_t$ and $\widehat{V}_t$ in the lost sales case are different than the identically named functions in the backorders case.) For $\widehat{V}_t(y^+)$ we denote the kinks and the derivatives by $\{u''_{k''}\}$ and $\{d''_{k''}\}$ respectively.

   Initialize $k'' = 1$ and $g''_0 = 0$.

   If $g_k \geq 0$

   Set $u'_{k'} = u_k$, $g'_{k'} = g_k$, $\hat{y} = \min\{\hat{y}, g_k\}$ and $k' \leftarrow k' + 1$. Do exactly one of the following.

   (A.1) If $u'_{k'} = 0$ then set $u''_{k''} = u'_{k'}$, $g''_{k''} = g'_{k'}$, $k'' \leftarrow k'' + 1$.

   (A.2) If $u'_{k'} > 0$ and $k'' = 1$ then $u''_1 = 0$, $g''_1 = g'_{k'-1}$, $u''_2 = u'_{k'}$, $g''_2 = g'_{k'}$, $k'' \leftarrow k'' + 2$.

   (A.3) If $u'_{k'} > 0$ and $k'' > 1$ then set $u''_{k''} = u'_{k'}$, $g''_{k''} = g'_{k'}$, $k'' \leftarrow k'' + 1$.

   end (If $g_k \geq 0$).

Observe that for a $\widehat{V}_t(y^+)$ function whose first kink is at zero the conditions (A.2) will never hold. For a $\widehat{V}_t(y^+)$ function whose first kink is strictly positive the conditions (A.1) will never hold.

S-5 Parametric Estimation for Inventory Models

To compare the method based upon empirical distribution functions with an approach based upon parametric estimation we considered cases in which a Poisson distribution was fitted to the empirical data, and then order quantities were selected based upon the estimated Poisson demand distributions. This is a natural procedure if one believes that demand is in fact Poisson distributed. Let $\Phi^m(\cdot)$ denote the distribution function of a Poisson random variable with mean (parameter) m. To estimate the parameter from data, we use the empirical mean, which is both the maximum likelihood estimator and the method of moments estimator of m.

Let $M_t = n_t^{-1} \sum_{i=1}^{n_t} Z_t^i$ be the empirical average of period-t demand realizations. We then estimate the period-t demand distribution by $\Phi^{M_t}(\cdot)$, and then solve

$$V_t^{\text{pois}}(x) = \min_{q:q\geq 0} U_t^{\text{pois}}(x+q) \tag{S-30}$$

$$U_t^{\text{pois}}(y) = K_t^{\text{pois}}(y) + \int V_{t+1}^{\text{pois}}(\phi(y-z))d\Phi^{M_t}(z) \qquad t = 1, \dots, \tau \tag{S-31}$$

where $K_t^{\text{pois}}(y) = \int k_t(y-z)d\Phi^{M_t}(z)$ and $V_\tau^{\text{pois}}(\cdot) = 0$ to obtain base-stock levels $\{y_t^{\text{pois}}\}$ where y_t^{pois} is a minimizer of $U_t^{\text{pois}}(\cdot)$. Note that (S-30)–(S-31) are of the same form as (15)–(16), but with Poisson distribution functions with estimated means in place of the empirical distribution functions. As before, we will be interested in evaluating the true performance of the policy that uses base-stock levels $\{y_t^{\text{pois}}(x)\}$; that is, we are interested in evaluating $\widetilde{V}_\tau^{\text{pois}}(x) = K_\tau(\max\{x, y_\tau^{\text{pois}}\})$ and

$$\widetilde{V}_t^{\text{pois}}(x) = K_t(\max\{x, y_t^{\text{pois}}\}) + \int \widetilde{V}_{t+1}^{\text{pois}}(\phi(\max\{x, y_t^{\text{pois}}\} - z))dF_t(z)\,. \tag{S-32}$$

If the true distributions are indeed Poisson [i.e., if $F_t(\cdot) = \Phi^{m_t}(\cdot)$ for $t = 1, \dots, \tau$ for some $\{m_t\}$], then M_t and Φ^{M_t} will be roughly equal to m_t and F_t respectively when n_t is large. These statements can be made precise.

For a problem in which the actual distributions of demand in periods $t = 1, \dots, 5$ are Poisson with means $1, 2, 6, 10, 1$ and $h = 1$ and $b = 10$, the row labeled "Poisson" in Table S-1 shows the mean and standard deviation of $R^{\text{pois}} = \max_x \big(\widetilde{V}_1^{\text{pois}}(x) - V_1(x)\big)/V_1(x)$ over 10,000 replications as well as the fraction of replications for which R^{pois} was no greater than 0.1 [which is given by $E_n^{\text{pois}}(0.1)$], and the 90%-quantile of the empirical distribution of the 10,000 simulated values of R^{pois}. As before, the columns $n = 5$, $n = 20$, and $n = 100$ correspond to having 5, 20, and 100 samples for each time period in each replication.

Not surprisingly, the parametric method that uses the fact demand is actually Poisson generally performs better in this example than the non-parametric method that uses empirical distribution functions. From the table, we see that the parametric method is, on average 13.70%, 3.13%, and 0.62% above optimal (for $n = 5, 20, 100$) compared to the empirical distribution method which is 24.58%, 6.52%, and 1.22% above optimal (see Table 2). So, the parametric approach is roughly half as far from optimal on average as is the non-parametric approach.

One of the reasons for using the non-parametric approach is that it does not require knowledge of the actual form of the demand distributions. Consequently, it is not susceptible to *misspecification*. In our context, misspecification occurs when a decision maker makes decisions by fitting a particular parametric family of distributions to data generated by actual demand distributions that do not have that parametric form. Misspecification occurs in the Poisson parametric context above if $F_t(\cdot) \notin \{\Phi^m(\cdot) : m \in [0, \infty)\}$; i.e., if the true demand distribution is not Poisson. There have been many studies of parameter estimation in the presence of misspecification; see, for example, White (1994). In practice, one might perform goodness-of-fit tests on the realized demand data in an attempt to avoid misspecification. However, doing so cannot completely eliminate the issue. Discussion of such procedures is beyond the scope of this paper.

The rows labeled $k = 2$, $k = 4$, and so on in Table S-1 show the performance of the method (S-30)–(S-31) when it is misspecified and the actual demand distributions are negative binomial. (The explanation of k is given in Section 6 in our discussion of Table 2.) Perhaps surprisingly, the misspecified Poisson approach performs better than the non-parametric approach when $n = 5$. This can be attributed to the empirical distribution "overfitting" the data for such a small n. This is particularly noticeable for large values of k (recall, such values correspond to high variability in demand). For large values of k, the actual negative binomial demand distributions become heavily concentrated at zero, and also assign mass to very large demand values. More precisely, as $k \to \infty$, the negative binomial distribution converges weakly to a point mass at zero, although its variance increases to ∞. For large k, a negative binomial distribution can, in a sense, be well-approximated by a Poisson distribution with very small mean (which is also "close" to being a point mass at zero). In addition, as k grows the optimal base stock levels decrease. Gallego et al. (2007) prove for single period problems with negative binomial demand, that if we fix the mean and scale the variance up, then in the limit as the variance grows to infinity, the optimal base-stock level converges to zero. When k is large, there infrequently occur realizations of extremely large demand values. For a replication in which there is such a large realization, the empirical distribution places a mass of size 1/5 at that large value when $n = 5$. This causes the empirical distribution method to select a base-stock level that is much too large. The approach based upon fitting a Poisson does not do this, and to some degree smooths out the data. Note that the non-parametric approach outperforms the Poisson parametric approach for all values of k when $n = 100$ and all values of k except $k = 64$ when $n = 20$. Hence, as more data is accumulated, the risk of overfitting decreases, and the non-parametric approach does better. It should be expected that the non-parametric method dominates as the number of samples n grows, because for large n, the empirical distribution function is very likely to closely approximate the actual distribution. In the presence of misspecification, this is not the case for the parametric approach.

It is notable that further increases in n need not improve the performance of the parametric approach in the presence of misspecification. On the other hand, increases in n do improve the performance of the approach based upon the empirical distribution function. For example, we ran the $k = 2$ case from Table S-1 with $n = 200$, 300, 400, 1000, 2000, and 5000 samples in each time period. Over 10,000 replications for each of these six additional settings the average values of R^{pois} were, respectively 0.0588, 0.0603, 0.0615, 0.0678, 0.0675, and 0.0776. As n grows even larger, the means of R^{pois} approach 0.0810, which corresponds to $M_t = \mathbb{E}Z_t^i$ in (S-30)–(S-32). This occurs because $M_t \to \mathbb{E}Z_t^i$ as $n \to \infty$ with probability one. These values should be compared with mean of 0.0602 in Table S-1 for $n = 100$. We also ran the approach based upon the empirical distributions with $n = 200$, 300, 400, 1000, 2000, and 5000 samples in each time period. The average values of R were 0.0094, 0.0061, 0.0044, 0.0013, 0.0005, and 5.3×10^{-5}. The non-parametric approach continues to improve with additional data collection whereas the misspecified parametric method does not. Theorems 2 and 3 make precise the idea that greater numbers of samples (larger n) ensure better performance of the empirical distribution approach for the inventory problem.

We also considered parametric estimation for negative binomial distributions, and used the method of moments to estimate the parameters a and m for each time t using the demand data $Z_t^1, \ldots, Z_t^{n_t}$. Let $\Psi^{(a,m)}(\cdot)$ denote the distribution function of a negative binomial random variable with parameters a and m in (22). The method of moments estimators for m and a are respectively $n_t^{-1} \sum_{i=1}^{n_t} Z_t^i$ and $(S_t^2 - n_t^{-1} \sum_{i=1}^{n_t} Z_t^i)/(n_t^{-1} \sum_{i=1}^{n_t} Z_t^i)^2$ where S_t^2 is the sample variance of $Z_t^1, \ldots, Z_t^{n_t}$; see Clark and Perry (1989). We modified these slightly to yield estimators $M_t = \max\{0.01, n_t^{-1} \sum_{i=1}^{n_t} Z_t^i\}$ and $A_t = \max\{0.01, (S_t^2 - M_t)/M_t^2\}$ for m and a. This was done to ensure that $\Psi^{(A_t,M_t)}(\cdot)$ is a valid distribution function; see Clark and Perry (1989) and references therein for discussion of difficulties associated with estimating the parameters of a negative binomial distribution.

As in the Poisson-based approach above, we used the parametric estimates of demand distributions to obtain base-stock levels. Specifically, we solved $V_t^{\text{nb}}(x) = \min_{q:q\geq 0} U_t^{\text{nb}}(x+q)$, $U_t^{\text{nb}}(y) = K_t^{\text{nb}}(y) + \int V_{t+1}^{\text{nb}}(\phi(y-z))d\Psi^{(A_t,M_t)}(z)$ where $K_t^{\text{nb}}(y) = \int k_t(y-z)d\Psi^{(A_t,M_t)}(z)$ and $V_\tau^{\text{nb}}(\cdot) = 0$. This yielded base-stock levels $\{y_t^{\text{nb}}\}$ where y_t^{nb} is a minimizer of $U_t^{\text{nb}}(\cdot)$. The true performance of the policy that uses base-stock levels $\{y_t^{\text{nb}}(x)\}$ was obtained by evaluating $\widetilde{V}_\tau^{\text{nb}}(x) = K_\tau(\max\{x, y_\tau^{\text{nb}}\})$ and $\widetilde{V}_t^{\text{nb}}(x) = K_t(\max\{x, y_t^{\text{nb}}\}) + \int \widetilde{V}_{t+1}^{\text{nb}}(\phi(\max\{x, y_t^{\text{nb}}\} - z))dF_t(z)$.

Table S-2 shows the results for this method, again based upon 10,000 simulation replications. We append "nb" as a superscript to indicate quantities associated with this method. The true demand distributions $\{F_t\}$ and other problem parameters are the same as in Tables 2 and S-1. Note that in Table S-2 there is no misspecification in the rows labeled $k = 2$ through $k = 64$, because for those rows, the actual demand distributions are negative binomial. In the first row, actual demand is Poisson. Here the parametric approach based upon the negative binomial does essentially as well the parametric approach based upon the Poisson distribution (compare with Table S-1). This occurs because the Poisson distribution is a limiting case of the negative binomial. In particular, the negative binomial distribution will converge weakly to a Poisson if we let the parameter $a \downarrow 0$.

Comparing Tables 2, S-1, and S-2, we see that if the actual demand distributions are negative binomial, then the parametric approach based upon the negative binomial usually outperforms the approaches based upon empirical distributions and the parametric approach based upon the Poisson. This is to be expected. It is surprising to see that the Poisson approach outperforms the negative binomial approach when the actual demand is negative binomial with $k = 64$ and $n = 5$ or $n = 20$ and also with $k = 32$ and $n = 5$. In these cases, we obtain better inventory policies by fitting misspecified demand distributions. The explanation for this seems again to be overfitting. For large k, there are occasional very large demand realizations. Fitting a negative binomial to those realizations yields base-stock levels that are much too high, whereas fitting a Poisson does not seem to suffer from the same effect. With more demand realizations ($n = 100$) the negative binomial approach does better, as individual large demand realizations have less effect.

Finally, we also studied the performance of both the Poisson and negative binomial parametric approaches when applied to problems in which the actual demand was scaled Bernoulli. We again used the true demand distributions that were used in Table 3. The results of 10,000 simulation replications are shown in Tables S-3 and S-4. Note that the parametric methods perform quite poorly, because for the shown values of k the scaled Bernoulli cannot be approximated well by a Poisson or a negative binomial. The performance was worst for $k = 2$. In contrast, the non-parametric approach based upon empirical distributions works very well in these settings, as shown in Table 3. As k increases, the parametric approaches do better because the scaled Bernoulli becomes closer to a point mass at zero (see the discussion above). Observe that in some cases (e.g., $k = 3$ with the Poisson approach), the number n of samples has little affect on the average performance of the derived policy. This occurs because even with larger values of n, estimates of the misspecified demand distribution do not approach the actual demand distribution.

References

P. Billingsley. *Probability and Measure*. John Wiley & Sons, New York, 3rd edition, 1995.

S. J. Clark and J. N. Perry. Estimation of the negative binomial parameter κ by maximum quasi-likelihood. *Biometrics*, 45(1):309–316, 1989.

G. B. Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, New York, 2nd edition, 1999.

G. Gallego, K. Katircioglu, and B. Ramachandran. Inventory management under highly uncertain demand. *Operations Research Letters*, 35(3):281–289, 2007.

R. Levi, R. O. Roundy, and D. B. Shmoys. Provably near-optimal sampling-based policies for stochastic inventory control models. *Mathematics of Operations Research*, 32(4), 2007.

P. Massart. The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *The Annals of Probability*, 18(3):1269–1283, 1990.

Actual demand distributions		$n = 5$	$n = 20$	$n = 100$
	mean	0.1370	0.0313	0.0062
Poisson	std. dev.	0.1190	0.0318	0.0068
c.o.v.'s: $(1.0, 0.7, 0.4, 0.3, 1.0)$	$E_n^{\text{pois}}(0.1)$	0.4715	0.9629	1.0000
	$(E_n^{\text{pois}})^{-1}(0.9)$	0.2864	0.0711	0.0115
	mean	0.2035	0.0897	0.0602
NB, $k = 2$	std. dev.	0.1491	0.0568	0.0240
c.o.v.'s: $(1.4, 1.0, 0.6, 0.5, 1.4)$	$E_n^{\text{pois}}(0.1)$	0.2793	0.6454	0.9400
	$(E_n^{\text{pois}})^{-1}(0.9)$	0.4118	0.1669	0.0810
	mean	0.2492	0.1463	0.1151
NB, $k = 4$	std. dev.	0.1618	0.0715	0.0323
c.o.v.'s: $(2.0, 1.4, 0.8, 0.6, 2.0)$	$E_n^{\text{pois}}(0.1)$	0.1670	0.2879	0.3271
	$(E_n^{\text{pois}})^{-1}(0.9)$	0.4701	0.2434	0.1621
	mean	0.2404	0.1590	0.1335
NB, $k = 8$	std. dev.	0.1478	0.0722	0.0329
c.o.v.'s: $(2.8, 2.0, 1.2, 0.9, 2.8)$	$E_n^{\text{pois}}(0.1)$	0.1577	0.2180	0.1502
	$(E_n^{\text{pois}})^{-1}(0.9)$	0.4427	0.2574	0.1741
	mean	0.1907	0.1291	0.1100
NB, $k = 16$	std. dev.	0.1183	0.0606	0.0278
c.o.v.'s: $(4.0, 2.8, 1.6, 1.3, 4.0)$	$E_n^{\text{pois}}(0.1)$	0.2509	0.3523	0.3640
	$(E_n^{\text{pois}})^{-1}(0.9)$	0.3561	0.2119	0.1475
	mean	0.1292	0.0804	0.0672
NB, $k = 32$	std. dev.	0.0844	0.0431	0.0207
c.o.v.'s: $(5.7, 4.0, 2.3, 1.8, 5.7)$	$E_n^{\text{pois}}(0.1)$	0.4308	0.7114	0.9366
	$(E_n^{\text{pois}})^{-1}(0.9)$	0.2440	0.1392	0.0953
	mean	0.0878	0.0417	0.0294
NB, $k = 64$	std. dev.	0.0845	0.0254	0.0117
c.o.v.'s: $(8.0, 5.7, 3.3, 2.5, 8.0)$	$E_n^{\text{pois}}(0.1)$	0.6916	0.9775	1.0000
	$(E_n^{\text{pois}})^{-1}(0.9)$	0.1643	0.0753	0.0451

Table S-1: Parametric Estimation of Poisson Distribution Functions. Sample means of R^{pois}, sample standard deviations of R^{pois}, fractions of replications that produced a policy within 10% of optimal, and 90% quantiles of $E_n^{\text{pois}}(\cdot)$. True demand distributions are negative binomial (except for the row labeled "Poisson"), with means 1, 2, 6, 10, 1 in periods $t = 1, 2, 3, 4, 5$.

Actual demand distributions		$n = 5$	$n = 20$	$n = 100$
Poisson c.o.v.'s: $(1.0, 0.7, 0.4, 0.3, 1.0)$	mean	0.1484	0.0356	0.0086
	std. dev.	0.1132	0.0317	0.0074
	$E_n^{\text{nb}}(0.1)$	0.3873	0.9535	1.0000
	$(E_n^{\text{nb}})^{-1}(0.9)$	0.2873	0.0750	0.0136
NB, $k = 2$ c.o.v.'s: $(1.4, 1.0, 0.6, 0.5, 1.4)$	mean	0.1827	0.0513	0.0129
	std. dev.	0.1295	0.0391	0.0102
	$E_n^{\text{nb}}(0.1)$	0.2898	0.8868	1.0000
	$(E_n^{\text{nb}})^{-1}(0.9)$	0.3485	0.1048	0.0262
NB, $k = 4$ c.o.v.'s: $(2.0, 1.4, 0.8, 0.6, 2.0)$	mean	0.1903	0.0553	0.0114
	std. dev.	0.1292	0.0395	0.0093
	$E_n^{\text{nb}}(0.1)$	0.26	0.8756	1.0000
	$(E_n^{\text{nb}})^{-1}(0.9)$	0.3596	0.1067	0.0239
NB, $k = 8$ c.o.v.'s: $(2.8, 2.0, 1.2, 0.9, 2.8)$	mean	0.1805	0.0540	0.0119
	std. dev.	0.1267	0.0390	0.0091
	$E_n^{\text{nb}}(0.1)$	0.2929	0.8894	1.0000
	$(E_n^{\text{nb}})^{-1}(0.9)$	0.3529	0.1048	0.0237
NB, $k = 16$ c.o.v.'s: $(4.0, 2.8, 1.6, 1.3, 4.0)$	mean	0.1592	0.0506	0.0117
	std. dev.	0.1203	0.0383	0.0091
	$E_n^{\text{nb}}(0.1)$	0.382	0.8952	1.0000
	$(E_n^{\text{nb}})^{-1}(0.9)$	0.3196	0.1016	0.0234
NB, $k = 32$ c.o.v.'s: $(5.7, 4.0, 2.3, 1.8, 5.7)$	mean	0.1426	0.0487	0.0122
	std. dev.	0.1217	0.0400	0.0096
	$E_n^{\text{nb}}(0.1)$	0.4574	0.9027	1.0000
	$(E_n^{\text{nb}})^{-1}(0.9)$	0.292	0.0990	0.0247
NB, $k = 64$ c.o.v.'s: $(8.0, 5.7, 3.3, 2.5, 8.0)$	mean	0.1301	0.0478	0.0127
	std. dev.	0.1629	0.0446	0.0108
	$E_n^{\text{nb}}(0.1)$	0.5773	0.9044	1.0000
	$(E_n^{\text{nb}})^{-1}(0.9)$	0.2733	0.0980	0.0271

Table S-2: Parametric Estimation of Negative Binomial Distribution Functions. Sample means of R^{nb}, sample standard deviations of R^{nb}, fractions of replications that produced a policy within 10% of optimal, and 90% quantiles of $E_n^{\text{nb}}(\cdot)$. True demand distributions are negative binomial (except for the row labeled "Poisson"), with means 1, 2, 6, 10, 1 in periods $t = 1, 2, 3, 4, 5$.

Actual demand distributions		$n = 5$	$n = 20$	$n = 100$
	mean	1.5783	1.3903	1.3439
SB, $k = 2$	std. dev.	0.7690	0.4554	0.2130
c.o.v.'s: 1.0	$E_n^{\mathrm{pois}}(0.1)$	0	0.0009	0
	$(E_n^{\mathrm{pois}})^{-1}(0.9)$	2.6034	2.0000	1.6034
	mean	1.1119	1.0965	1.1039
SB, $k = 3$	std. dev.	0.3499	0.1752	0.0819
c.o.v.'s: 1.4	$E_n^{\mathrm{pois}}(0.1)$	0	0	0
	$(E_n^{\mathrm{pois}})^{-1}(0.9)$	1.5670	1.3257	1.2107
	mean	0.7271	0.7205	0.7230
SB, $k = 4$	std. dev.	0.1805	0.0841	0.0386
c.o.v.'s: 1.7	$E_n^{\mathrm{pois}}(0.1)$	0.0001	0	0
	$(E_n^{\mathrm{pois}})^{-1}(0.9)$	0.9547	0.8262	0.7723

Table S-3: Parametric Estimation of Poisson Distribution Functions. Sample means of R^{pois}, sample standard deviations of R^{pois}, fractions of replications that produced a policy within 10% of optimal, and 90% quantiles of $E_n^{\mathrm{pois}}(\cdot)$. True demand distributions are scaled Bernoulli with means 1, 2, 6, 10, 1 in periods $t = 1, 2, 3, 4, 5$.

Actual demand distributions		$n = 5$	$n = 20$	$n = 100$
	mean	0.9690	0.5100	0.4046
SB, $k = 2$	std. dev.	0.6808	0.2958	0.1810
c.o.v.'s: 1.0	$E_n^{\mathrm{nb}}(0.1)$	0	0.0080	0.0123
	$(E_n^{\mathrm{nb}})^{-1}(0.9)$	1.8276	0.9138	0.6724
	mean	0.7681	0.5424	0.4502
SB, $k = 3$	std. dev.	0.4212	0.2324	0.1007
c.o.v.'s: 1.4	$E_n^{\mathrm{nb}}(0.1)$	0.0092	0.0082	0
	$(E_n^{\mathrm{nb}})^{-1}(0.9)$	1.3812	0.8736	0.5862
	mean	0.5550	0.4479	0.4074
SB, $k = 4$	std. dev.	0.2467	0.1508	0.0687
c.o.v.'s: 1.7	$E_n^{\mathrm{nb}}(0.1)$	0.0048	0.0008	0
	$(E_n^{\mathrm{nb}})^{-1}(0.9)$	0.9145	0.6544	0.4964

Table S-4: Parametric Estimation of Negative Binomial Distribution Functions. Sample means of R^{nb}, sample standard deviations of R^{nb}, fractions of replications that produced a policy within 10% of optimal, and 90% quantiles of $E_n^{\mathrm{nb}}(\cdot)$. True demand distributions are scaled Bernoulli with means 1, 2, 6, 10, 1 in periods $t = 1, 2, 3, 4, 5$.

T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

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

A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory*. SIAM and MPS, Philadelphia, 2009.

H. White. *Estimation, Inference, and Specification Analysis*. Cambridge University Press, New York, 1994.