Online Appendix for the paper
"Models of the Spiral-Down Effect in Revenue Management"
by William L. Cooper, Tito Homem-de-Mello and Anton J. Kleywegt

OA–1 Proofs for the Deterministic Example

Proposition 1. *Suppose that the probability distribution of the observed quantity is given by* (6) *with* $d < c$*, and that forecasts are made according to* (7)*. Then* $L^{k+1} \le L^k$ *for all* k*. Furthermore, there exists a* k^* *such that* $L^j = 0$ *and* $X^j = 0$ *for all* $j \ge k^*$*.*

Proof. Note that

$$X^{k+1} \;=\; [d-(c-L^k)^+]^+ \;\le\; [d-(c-L^k)]^+ \;\le\; L^k. \tag{OA–1}$$

In view of (7), we see that $\hat{H}^{k+1}(x) > \hat{H}^k(x)$ for all $x \ge X^{k+1}$ such that $\hat{H}^k(x) < 1$, and $\hat{H}^{k+1}(x) = 1$ for all $x \ge X^{k+1}$ such that $\hat{H}^k(x) = 1$. Therefore, $L^{k+1} \le L^k$ by (5), so the first part of the proposition is proved.

Let $\varepsilon := c - d > 0$. Notice that if $L^j \ge \varepsilon$ then $X^{j+1} \le L^j - \varepsilon$ by (OA–1). Moreover, (OA–1) also implies that if $0 \le L^j < \varepsilon$, then $X^{j+1} = 0$. Since we have already shown that the sequence of protection levels is non-increasing, it follows that if k is such that $L^k \ge \varepsilon$, then $X^{j+1} \le L^k - \varepsilon$ for all $j \ge k$.

Define

$$k' \;:=\; \min\left\{ j > k \,:\, \frac{k}{j}\hat{H}^k(L^k-\varepsilon) + \frac{j-k}{j} > \gamma \right\}. \tag{OA–2}$$

Observe that $k' < \infty$, because $\gamma < 1$. By (OA–2), we have that $\hat{H}^{k'}(L^k - \varepsilon) > \gamma$. Therefore, if $x \in (\hat{H}^{k'})^{-1}(\gamma)$ then $x \le L^k - \varepsilon$. Since $L^{k'} \in (\hat{H}^{k'})^{-1}(\gamma)$, it follows that $L^{k'} \le L^k - \varepsilon$.

Suppose now that $0 \le L^k < \varepsilon$. Then, (OA–1) implies that $X^{k+1} = 0$. An argument similar to that used above shows that there exists a $k^* > k$ such that $L^{k^*} = 0$. Since the sequence of protection levels is non-increasing, the second part of the proposition follows. □

Proposition 2. *Suppose that the probability distribution of the observed quantity is given by* (6) *with* $d > c$*, and that forecasts are made according to* (7)*. Suppose that* $L^0 \in [0, c]$*. Then* $L^{k+1} \ge L^k$ *for all* k*. Furthermore, there exists a* k° *such that* $L^j = d$ *and* $X^j = d$ *for all* $j \ge k^\circ$*.*

Proof. For the first part of the proposition, suppose that $L^k \in [0, c]$. Note that

$$X^{k+1} \;=\; d-(c-L^k) \;=\; L^k + \varepsilon. \tag{OA–3}$$

In view of (7), we see that $\hat{H}^{k+1}(x) \leq \hat{H}^k(x)$ for all $x < X^{k+1}$; in addition, $\hat{H}^{k+1}(x) < \hat{H}^k(x)$ for all $x < X^{k+1}$ such that $\hat{H}^k(x) > 0$. Therefore, $L^{k+1} \geq L^k$ by (5).

Recall that $\hat{H}^k(X^{k+1}-) := \lim_{x\uparrow X^{k+1}} \hat{H}^k(x)$ denotes the left limit of $\hat{H}^k$ at X^{k+1}. Consider any integer $j > k\hat{H}^k(X^{k+1}-)/\gamma$. Then one of two cases must hold: either there is an integer $k' \leq j$ such that $L^{k'} > c$, or $L^i \in [0, c]$ for all $i \leq j$. In the latter case, choose $k' = j$, and note that $\hat{H}^j(X^{k+1}-) = k\hat{H}^k(X^{k+1}-)/j < \gamma$, and thus $L^j := (\hat{H}^j)^{-1}(\gamma) \geq X^{k+1} = L^k + \varepsilon$. In summary, k' is such that $L^{k'} > c$ or $L^{k'} \geq L^k + \varepsilon$.

Next, note that if $L^k > c$, then $X^{k+1} = d$. An argument similar to that used above shows that there exists a $k^\circ \geq k$ such that $L^{k^\circ} = d$. Note that at the first time k' such that $L^{k'} > c$, it still holds that $L^{k'} \leq d$, because $X^k \leq d$ and thus $\hat{H}^k(d) = 1$ for all k, and hence $L^k \leq L^{k+1}$ also when $L^k > c$. For the same reason, given that $L^{k^\circ} = d$ then $L^k = d$ for all $k \geq k^\circ$, which is the second assertion of the proposition. □

OA–2 Proof of Proposition 17

Lemma OA–1. *Consider the metric space $(\mathcal{P}(\mathbb{R}), \lambda)$ of probability distributions on $\mathbb{R}$ endowed with the Lévy metric λ, defined as follows for $F, H \in \mathcal{P}(\mathbb{R})$:*

$$\lambda(F, H) \quad := \quad \inf\{\varepsilon > 0 \,:\, F(x-\varepsilon) - \varepsilon \leq H(x) \leq F(x+\varepsilon) + \varepsilon \;\forall\, x \in \mathbb{R}\}.$$

Let $\mathbb{N}$ denote the natural numbers, and let $\mathbb{Q}$ denote the rational numbers. Then for any $F, H \in \mathcal{P}(\mathbb{R})$ and any $r > 0$, $\lambda(F, H) < r$ if and only if there exists $m \in \mathbb{N}$ such that

$$F(x - r + 1/m) - r + 1/m \quad < \quad H(x) \quad < \quad F(x + r - 1/m) + r - 1/m$$

for all $x \in \mathbb{Q}$.

Proof. First, suppose that $\lambda(F, H) < r$. Then there exists $m \in \mathbb{N}$ such that $\lambda(F, H) < r - 1/m$, and it follows from F being nondecreasing that $F(x - r + 1/m) - r + 1/m < H(x) < F(x + r - 1/m) + r - 1/m$ for all $x \in \mathbb{R}$, and thus for all $x \in \mathbb{Q}$.

Next, suppose that there exists an $m \in \mathbb{N}$ such that $F(x - r + 1/m) - r + 1/m < H(x) < F(x + r - 1/m) + r - 1/m$ for all $x \in \mathbb{Q}$. Consider any $x \in \mathbb{R}$, and a sequence $\{x^n\} \subset \mathbb{Q}$ such that $x^n \downarrow x$. Then $F(x^n - r + 1/m) - r + 1/m < H(x^n) < F(x^n + r - 1/m) + r - 1/m$ for all n. It follows from the right continuity of F and H that $F(x - r + 1/m) - r + 1/m \leq H(x) \leq F(x + r - 1/m) + r - 1/m$. Hence $\lambda(F, H) := \inf\{\varepsilon > 0 \,:\, F(x - \varepsilon) - \varepsilon \leq H(x) \leq F(x + \varepsilon) + \varepsilon \;\forall\, x \in \mathbb{R}\} \leq r - 1/m < r$. □

Proposition 17. *Let B denote the Borel σ-algebra on $\mathbb{R}$. Consider the space $(\mathcal{P}(\mathbb{R}), \mathcal{B})$ of probability distributions on $\mathbb{R}$, endowed with the Borel σ-algebra $\mathcal{B}$ corresponding to the topology of weak convergence on $\mathcal{P}(\mathbb{R})$. Consider a measurable space $(\Omega, \mathcal{F})$. Let $\{H^k : \Omega \mapsto \mathcal{P}(\mathbb{R})\}$ be a sequence of $(\mathcal{F}, \mathcal{B})$-measurable functions.*

(i) *Consider a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and a filtration* $\{\mathcal{F}^k\}$. *Consider a random sequence* $\{Y^k\}$ *adapted to filtration* $\{\mathcal{F}^k\}$, *where* $Y^k : \Omega \mapsto \mathbb{R}$. *Let* $F^k : \Omega \mapsto \mathcal{P}(\mathbb{R})$ *be given by* $F^k(\omega, x) := \mathbb{P}[Y^{k+1} \leq x \,|\, \mathcal{F}^k]$, *that is,* F^k *is the conditional distribution of* Y^{k+1}. *Then* F^k *is* $(\mathcal{F}^k, \mathcal{B})$*-measurable.*

(ii) *The set* $\Omega^* := \{\omega \in \Omega \,:\, H^k(\omega, \cdot)$ *converges weakly as* $k \to \infty\}$ *is in* $\mathcal{F}$.

(iii) *Let* $\Omega^* := \{\omega \in \Omega \,:\, H^k(\omega, \cdot)$ *converges weakly as* $k \to \infty\}$, *and let* $\mathcal{F}^* := \{A \in \mathcal{F} \,:\, A \subset \Omega^*\}$. *For each* $\omega \in \Omega^*$, *let* $H^*(\omega, \cdot)$ *denote the weak limit of* $\{H^k(\omega, \cdot)\}$. *Then* $\mathcal{F}^*$ *is a* σ*-algebra on* Ω^*. *In addition,* H^* *is* $(\mathcal{F}^*, \mathcal{B})$*-measurable, and thus* H^* *is also* $(\mathcal{F}, \mathcal{B})$*-measurable.*

(iv) *For any* $(\mathcal{F}, \mathcal{B})$*-measurable* $F : \Omega \mapsto \mathcal{P}(\mathbb{R})$, *the set* $\left\{\omega \in \Omega \,:\, H^k(\omega, \cdot) \xrightarrow{w} F(\omega, \cdot)\right\}$ *is in* $\mathcal{F}$.

(v) *Let* $F : \Omega \mapsto \mathcal{P}(\mathbb{R})$ *be an* $(\mathcal{F}, \mathcal{B})$*-measurable function. For any* $x \in \mathbb{R}$, *let* $f_x : \Omega \mapsto \mathbb{R}$ *be defined as* $f_x(\omega) := F(\omega, x)$. *Then,* f_x *is* $(\mathcal{F}, B)$*-measurable. That is,* f_x *is a real-valued random variable.*

Proof.

(i) Fix k. For each $x \in \mathbb{R}$, define the function $\pi_x : \mathcal{P}(\mathbb{R}) \mapsto \mathbb{R}$ by $\pi_x(F) := F(x)$. Consider $\pi_x \circ F^k : \Omega \mapsto \mathbb{R}$. Note that $\pi_x(F^k(\omega, \cdot)) = F^k(\omega, x) := \mathbb{P}[Y^{k+1} \leq x \,|\, \mathcal{F}^k]$, and thus $\pi_x \circ F^k$ is $(\mathcal{F}^k, B)$-measurable.

Convergence in the Lévy metric λ, defined in Lemma OA–1, is equivalent to weak convergence of elements of $\mathcal{P}(\mathbb{R})$. Moreover, the space $\mathcal{P}(\mathbb{R})$, endowed with the Lévy metric λ, is complete and separable. For any $F \in \mathcal{P}(\mathbb{R})$ and $r > 0$, let $B(F, r) := \{H \in \mathcal{P}(\mathbb{R}) \,:\, \lambda(F, H) < r\}$ denote the ball with center F and radius r in $(\mathcal{P}(\mathbb{R}), \lambda)$. Since $(\mathcal{P}(\mathbb{R}), \lambda)$ is separable, its Borel sigma algebra $\mathcal{B}$ is generated by the countable collection of open balls $\{B(F, 1/m) : F \in D, m \in \mathbb{N}\}$, where D is a countable, dense subset of $\mathcal{P}(\mathbb{R})$. Therefore, to prove that F^k is $(\mathcal{F}^k, \mathcal{B})$-measurable, it suffices to show that $(F^k)^{-1}(B(F, r)) \in \mathcal{F}^k$ for all $F \in \mathcal{P}(\mathbb{R})$ and $r > 0$.

Consider any $F \in \mathcal{P}(\mathbb{R})$ and $r > 0$. For any $m \in \mathbb{N}$ and $x \in \mathbb{R}$, let $A_{m,x} := (F(x - r + 1/m) - r + 1/m, F(x + r - 1/m) + r - 1/m)$. It follows from Lemma OA–1 that $B(F, r) = \cup_{m\in\mathbb{N}} \cap_{x\in\mathbb{Q}} \pi_x^{-1}(A_{m,x})$.

Thus, for any $B(F, r)$,

$$\begin{aligned}
(F^k)^{-1}\big(B(F, r)\big) &= (F^k)^{-1}\left(\bigcup_{m\in\mathbb{N}} \bigcap_{x\in\mathbb{Q}} \pi_x^{-1}(A_{m,x})\right) \\
&= \bigcup_{m\in\mathbb{N}} \bigcap_{x\in\mathbb{Q}} (F^k)^{-1}\big(\pi_x^{-1}(A_{m,x})\big) \\
&= \bigcup_{m\in\mathbb{N}} \bigcap_{x\in\mathbb{Q}} (\pi_x \circ F^k)^{-1}(A_{m,x})
\end{aligned}$$

Recall that $\pi_x \circ F^k$ is $(\mathcal{F}^k, B)$-measurable. Thus $(\pi_x \circ F^k)^{-1}(A_{m,x}) \in \mathcal{F}^k$, and hence $(F^k)^{-1}\big(B(F,r)\big) \in \mathcal{F}^k$.

(ii) Completeness of $(\mathcal{P}(\mathbb{R}), \lambda)$ implies that

$$\big\{\omega \in \Omega : \lim_{k\to\infty} F^k(\omega,\cdot) \text{ exists}\big\} \quad = \quad \big\{\omega \in \Omega : \{F^k(\omega,\cdot)\} \text{ is Cauchy}\big\}. \tag{OA–4}$$

The event on the right above can be expressed as

$$\bigcap_{m\geq 1} \bigcup_{n\geq 1} \bigcap_{\{i:i\geq n\}} \bigcap_{\{j:j\geq n\}} \big\{\omega \in \Omega : \lambda(F^i(\omega,\cdot), F^j(\omega,\cdot)) < 1/m\big\} \tag{OA–5}$$

Separability of $(\mathcal{P}(\mathbb{R}), \lambda)$ implies that the mappings $\Lambda^{ij} : \Omega \mapsto \mathbb{R}$ defined by $\Lambda^{ij}(\omega) := \lambda(F^i(\omega,\cdot), F^j(\omega,\cdot))$ are all $(\mathcal{F}, B)$-measurable (Billingsley, 1968, p.25). Hence $\big\{\omega \in \Omega : \lambda(F^i(\omega,\cdot), F^j(\omega,\cdot)) < 1/m\big\} \in \mathcal{F}$ for all i, j, m, and therefore the set in (OA–5) is in $\mathcal{F}$.

(iii) It is easy to verify that $\mathcal{F}^*$ is a σ-algebra on Ω^*. It follows from Dudley (2002), Theorem 4.2.2, that F^* is $(\mathcal{F}^*, \mathcal{B})$-measurable. It follows immediately that F^* is also $(\mathcal{F}, \mathcal{B})$-measurable.

(iv) As before, separability of $(\mathcal{P}(\mathbb{R}), \lambda)$ implies that the mappings $\Lambda^k : \Omega \mapsto \mathbb{R}$ defined by $\Lambda^k(\omega) := \lambda(F^k(\omega,\cdot), F(\omega,\cdot))$ are all $(\mathcal{F}, B)$-measurable, and therefore so is $\limsup_{k\to\infty} \Lambda^k$. Since

$$\left\{\omega \in \Omega : F^k(\omega,\cdot) \xrightarrow{w} F(\omega,\cdot)\right\} \quad = \quad \left\{\omega \in \Omega : \limsup_{k\to\infty} \Lambda^k(\omega) = 0\right\},$$

and since the set on the right is in $\mathcal{F}$, it follows that the set on the left is in $\mathcal{F}$ as well.

(v) We follow closely an argument in Billingsley (1968), p.121. For each $x \in \mathbb{R}$, define the function $\pi_x : \mathcal{P}(\mathbb{R}) \mapsto \mathbb{R}$ by $\pi_x(F) := F(x)$. For each $\varepsilon > 0$, define the function $\pi_x^\varepsilon : \mathcal{P}(\mathbb{R}) \mapsto \mathbb{R}$ by $\pi_x^\varepsilon(F) := \varepsilon^{-1} \int_x^{x+\varepsilon} F(u)du$. We first show that π_x^ε is continuous on $\mathcal{P}(\mathbb{R})$ and thus measurable. Consider any sequence $\{H^k\} \subset \mathcal{P}(\mathbb{R})$ such that $H^k \xrightarrow{w} H$. It follows from a characterization of weak convergence of distribution functions on $\mathbb{R}$ that $H^k(u) \to H(u)$ for all u except on a countable set (the set of discontinuities of H). Since $H^k(u) \leq 1$ for all u, it follows by the bounded convergence theorem that $\int_x^{x+\varepsilon} H^k(u)du \to \int_x^{x+\varepsilon} H(u)du$, and thus $\pi_x^\varepsilon(H^k) \to \pi_x^\varepsilon(H)$. Hence, π_x^ε is continuous and thus measurable. Next, since H is right-continuous, it follows that $\pi_x(H) = \lim_{\varepsilon\downarrow 0} \pi_x^\varepsilon(H) = \lim_{m\to\infty} \pi_x^{1/m}(H)$. Thus π_x is the limit of a sequence of measurable functions and therefore is $(\mathcal{B}, B)$-measurable. It follows from the definition of π_x that $f_x(\omega) = \pi_x(F(\omega,\cdot))$. Therefore, f_x is $(\mathcal{F}, B)$-measurable. □

OA–3 Supporting Material for Proposition 3

Below, we use some notation from Section OA–2: λ is the Lévy metric on $\mathcal{P}(\mathbb{R})$, B is the Borel σ-algebra on $\mathbb{R}$, and $B(h, r)$ is the ball of radius r about $h \in \mathcal{P}(\mathbb{R})$.

Lemma OA–2. *Let $\psi : \mathbb{R} \mapsto \mathcal{P}(\mathbb{R})$ be given by*

$$\psi(y) \quad := \quad \mathbb{I}_{\{\,\cdot\,\geq y\}}\,. \tag{OA–6}$$

Then the mapping ψ is $(B, \mathcal{B})$-measurable.

Proof. For any $h \in \mathcal{P}(\mathbb{R})$ and $y \in \mathbb{R}$ we have

$$\begin{aligned}
\lambda(h, \psi(y)) &= \inf\left\{\varepsilon > 0 \;:\; \begin{array}{l} h(x-\varepsilon) - \varepsilon \leq 0 \leq h(x+\varepsilon) + \varepsilon \text{ for all } x \;:\; x < y \text{ and} \\ h(x-\varepsilon) - \varepsilon \leq 1 \leq h(x+\varepsilon) + \varepsilon \text{ for all } x \;:\; x \geq y \end{array}\right\} \\
&= \inf\left\{\varepsilon > 0 \;:\; \begin{array}{l} h(x-\varepsilon) \leq \varepsilon \text{ for all } x \;:\; x < y \text{ and} \\ h(x+\varepsilon) \geq 1 - \varepsilon \text{ for all } x \;:\; x \geq y \end{array}\right\} \\
&= \inf\{\varepsilon > 0 \;:\; \lim_{x \uparrow y} h(x-\varepsilon) \leq \varepsilon \text{ and } h(y+\varepsilon) \geq 1 - \varepsilon\}
\end{aligned} \tag{OA–7}$$

The space $(\mathcal{P}(\mathbb{R}), \lambda)$ is separable with countable base given by $\{B(h, r) \,:\, h \in D, r \in \mathbb{Q}\}$, where D is a countable dense subset of $\mathcal{P}(\mathbb{R})$. Hence, to show the $(B, \mathcal{B})$-measurability of ψ, it suffices to show that $\psi^{-1}(B(h, r)) \in B$ for all h and r.

To this end, for $h \in \mathcal{P}(\mathbb{R})$ and $\varepsilon > 0$, define $\psi^1_{h,\varepsilon}, \psi^2_{h,\varepsilon}, \psi_{h,\varepsilon} : \mathbb{R} \mapsto \mathbb{R}$ by

$$\psi^1_{h,\varepsilon}(y) \quad := \quad \varepsilon - \lim_{x \uparrow y} h(x-\varepsilon), \tag{OA–8}$$

$$\psi^2_{h,\varepsilon}(y) \quad := \quad h(y+\varepsilon) - 1 + \varepsilon, \tag{OA–9}$$

$$\psi_{h,\varepsilon}(y) \quad := \quad \min\{\psi^1_{h,\varepsilon}(y), \psi^2_{h,\varepsilon}(y)\} \tag{OA–10}$$

The functions in (OA–8)–(OA–10) above are (B, B)-measurable because $h \in \mathcal{P}(\mathbb{R})$. Moreover, by (OA–7) and (OA–8)–(OA–10), we have

$$\begin{aligned}
\psi^{-1}(B(h, r)) &= \{y \in \mathbb{R} \;:\; \lambda(h, \psi(y)) < r\} \\
&= \left\{y \in \mathbb{R} \;:\; \inf\{\varepsilon > 0 \;:\; \lim_{x \uparrow y} h(x-\varepsilon) \leq \varepsilon \text{ and } h(y+\varepsilon) \geq 1 - \varepsilon\} < r\right\} \\
&= \{y \in \mathbb{R} \;:\; \inf\{\varepsilon > 0 \;:\; \psi_{h,\varepsilon}(y) \geq 0\} < r\} \\
&= \bigcup_{n \,:\, n^{-1} < r} \{y \in \mathbb{R} \;:\; \psi_{h, r-1/n}(y) \geq 0\}.
\end{aligned}$$

In view of the measurability of $\psi_{h,\varepsilon}(\cdot)$, all the sets in the union in the final expression are in B, and hence the proof is complete. □

Lemma OA–3. *Suppose that $H_1, H_2 : \Omega \mapsto \mathcal{P}(\mathbb{R})$ are both $(\mathcal{F}, \mathcal{B})$-measurable mappings and that $\alpha \in [0, 1]$. The mapping $\xi_{\alpha,H_1,H_2} : \Omega \mapsto \mathcal{P}(\mathbb{R})$ given by*

$$\xi_{\alpha,H_1,H_2}(\omega, x) \quad := \quad \alpha H_1(\omega, x) + (1-\alpha) H_2(\omega, x), \qquad x \in \mathbb{R} \tag{OA–11}$$

is $(\mathcal{F}, \mathcal{B})$-measurable.

Proof. Note that ξ_{α,H_1,H_2} can be expressed as $\theta_\alpha \circ J_{H_1,H_2}$ where $J_{H_1,H_2} : \Omega \mapsto \mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R})$ is defined by $J_{H_1,H_2}(\omega) := (H_1(\omega), H_2(\omega))$, and $\theta_\alpha : \mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R}) \mapsto \mathcal{P}(\mathbb{R})$ is defined by

$$\theta_\alpha(h_1, h_2)(x) \quad := \quad \alpha h_1(x) + (1-\alpha)h_2(x), \qquad x \in \mathbb{R}\,.$$

The mapping J_{H_1,H_1} is $(\mathcal{F}, \mathcal{B} \times \mathcal{B})$-measurable, where $\mathcal{B} \times \mathcal{B}$ is defined as the σ-algebra generated by sets of the form $A_1 \times A_2$ with $A_1, A_2 \in \mathcal{B}$. So the lemma will be proved if we can show that θ_α is $(\mathcal{B} \times \mathcal{B}, \mathcal{B})$-measurable.

For this, consider the metric space $\mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R})$ with metric λ^* given by

$$\lambda^*((h_1, h_2), (h_1', h_2')) \quad := \quad \max\{\lambda(h_1, h_1'), \lambda(h_2, h_2')\};$$

see Billingsley (1968), p.225. From the definitions of λ, λ^*, and θ_α, it follows that $\lambda^*((h_1, h_2), (h_1', h_2')) \geq \lambda(\theta_\alpha(h_1, h_2), \theta_\alpha(h_1', h_2'))$. Therefore, θ_α is continuous. That is, for any open (in the topology metrized by λ) set $O \subset \mathcal{P}(\mathbb{R})$, $\theta_\alpha^{-1}(O)$ is an open set in the topology metrized by λ^*. The Borel sigma algebra on $(\mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R}), \lambda^*)$ is precisely $\mathcal{B} \times \mathcal{B}$ (Billingsley, 1968, p.225), so the open sets metrized by λ^* are in $\mathcal{B} \times \mathcal{B}$. Summarizing, the open sets metrized by λ generate $\mathcal{B}$, and the inverse image of any such open set under θ_α is in $\mathcal{B} \times \mathcal{B}$. Hence, θ_α is $(\mathcal{B} \times \mathcal{B}, \mathcal{B})$-measurable, which completes the proof. $\square$

Proposition 18. *Suppose that $\{Y^k : \Omega \mapsto \mathbb{R}\}$ are $(\mathcal{F}, B)$-measurable random variables. Then $\hat{H}^k$ defined in (12) is $(\mathcal{F}, \mathcal{B})$-measurable for all k.*

Proof. The proof is by induction. Let ψ be as defined in (OA–6). Note that

$$\hat{H}^k(\omega, \cdot) \quad = \quad \frac{1}{k}\sum_{n=1}^{k}(\psi \circ Y^n)(\omega, \cdot) \quad = \quad \frac{1}{k}\,(\psi \circ Y^k)(\omega, \cdot) + \frac{k-1}{k}\,\hat{H}^{k-1}(\omega, \cdot)$$

Lemma OA–2 and the assumptions on Y^n imply that $\psi \circ Y^n$ is $(\mathcal{F}, \mathcal{B})$-measurable for each n. Hence we immediately see that $\hat{H}^1 := \psi \circ Y^1$ is $(\mathcal{F}, \mathcal{B})$-measurable. Suppose that the result holds for $k-1$. With $\alpha = 1/k$, $H_1 = \psi \circ Y^k$, and $H_2 = \hat{H}^{k-1}$, we see that

$$\hat{H}^k(\omega, \cdot) \quad = \quad \xi_{1/k\,,\,\psi \circ Y^k,\,\hat{H}^{k-1}}(\omega, \cdot), \tag{OA–12}$$

where ξ_{α,H_1,H_2} is defined in (OA–11). The desired result now follows from (OA–12), the induction hypothesis, and Lemma OA–3. $\square$

Lemma OA–4. *Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a collection $\{A_i : i \in I\} \subset \mathcal{F}$ of events, where I is a countable index set. Suppose that $\mathbb{P}[A_i] \geq \varepsilon > 0$ for all $i \in I$, and that for any $n+1$ distinct indices $i_1, \ldots, i_{n+1} \in I$, it holds that $A_{i_1} \cap \cdots \cap A_{i_{n+1}} = \varnothing$. Then $|I| \leq n/\varepsilon$.*

Proof. Let $\{S_j : j \in J\} \subset 2^I$ denote the collection of all subsets of I such that $1 \leq |S_j| \leq n$ for all $j \in J$. Note that J is countable. For all $i \in I$,

$$A_i = \bigcup_{\{j \in J \,:\, i \in S_j\}} \bigcap_{i' \in S_j} A_{i'} \bigcap_{i' \in S_j^c} A_{i'}^c$$

and the sets $\{\cap_{i' \in S_j} A_{i'} \cap_{i' \in S_j^c} A_{i'}^c : j \in J\}$ are disjoint. Thus,

$$\mathbb{P}[A_i] = \sum_{\{j \in J \,:\, i \in S_j\}} \mathbb{P}\left[\bigcap_{i' \in S_j} A_{i'} \bigcap_{i' \in S_j^c} A_{i'}^c\right] \geq \varepsilon > 0$$

Also,

$$\bigcup_{i \in I} A_i = \bigcup_{j \in J} \bigcap_{i \in S_j} A_i \bigcap_{i \in S_j^c} A_i^c$$

and, as before, the sets $\{\cap_{i \in S_j} A_i \cap_{i \in S_j^c} A_i^c : j \in J\}$ are disjoint. Thus,

$$\sum_{j \in J} \mathbb{P}\left[\bigcap_{i \in S_j} A_i \bigcap_{i \in S_j^c} A_i^c\right] = \mathbb{P}\left[\bigcup_{j \in J} \bigcap_{i \in S_j} A_i \bigcap_{i \in S_j^c} A_i^c\right] = \mathbb{P}\left[\bigcup_{i \in I} A_i\right] \leq 1$$

Also,

$$\begin{aligned}
&\sum_{j \in J} \mathbb{P}\left[\bigcap_{i \in S_j} A_i \bigcap_{i \in S_j^c} A_i^c\right] \\
&\geq \inf\left\{\sum_{j \in J} x_j : \sum_{\{j \in J \,:\, i \in S_j\}} x_j = \mathbb{P}[A_i] \ \forall\, i \in I,\ x_j \geq 0\ \forall\, j \in J\right\} \\
&\geq \inf\left\{\sum_{j \in J} x_j : \sum_{\{j \in J \,:\, i \in S_j\}} x_j \geq \varepsilon\ \forall\, i \in I,\ x_j \geq 0\ \forall\, j \in J\right\} \\
&= \inf\left\{\sup\left\{\sum_{j \in J} x_j + \sum_{i \in I} y_i\left(\varepsilon - \sum_{\{j \in J \,:\, i \in S_j\}} x_j\right) : y_i \geq 0\ \forall\, i \in I\right\} : x_j \geq 0\ \forall\, j \in J\right\} \\
&\geq \sup\left\{\inf\left\{\sum_{j \in J} x_j + \sum_{i \in I} y_i\left(\varepsilon - \sum_{\{j \in J \,:\, i \in S_j\}} x_j\right) : x_j \geq 0\ \forall\, j \in J\right\} : y_i \geq 0\ \forall\, i \in I\right\} \\
&= \sup\left\{\inf\left\{\sum_{i \in I} \varepsilon y_i + \sum_{j \in J} x_j\left(1 - \sum_{i \in S_j} y_i\right) : x_j \geq 0\ \forall\, j \in J\right\} : y_i \geq 0\ \forall\, i \in I\right\} \\
&= \sup\left\{\sum_{i \in I} \varepsilon y_i : \sum_{i \in S_j} y_i \leq 1\ \forall\, j \in J, y_i \geq 0\ \forall\, i \in I\right\} \\
&\geq |I|\varepsilon/n
\end{aligned}$$

where the last inequality follows from the observation that $y_i = 1/n$ for all $i \in I$ satisfies $\sum_{i \in S_j} y_i \leq 1$ for all $j \in J$, because $|S_j| \leq n$ for all $j \in J$. Combining the results above, it follows that $|I|\varepsilon/n \leq 1$, and thus $|I| \leq n/\varepsilon$. $\square$

Lemma 3. *Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and the space $(\mathcal{P}(\mathbb{R}), \mathcal{B})$ of probability distributions on $\mathbb{R}$ endowed with the Borel σ-algebra $\mathcal{B}$ corresponding to the topology of weak convergence on $\mathcal{P}(\mathbb{R})$. Let $F : \Omega \mapsto \mathcal{P}(\mathbb{R})\}$ be a $(\mathcal{F}, \mathcal{B})$-measurable function. For each $\omega \in \Omega$, let $D(\omega) := \{x \in \mathbb{R} : F(\omega, x) > F(\omega, x-)\}$ denote the set of jump points of $F(\omega, \cdot)$. Then the set $\{x \in \mathbb{R} : \mathbb{P}[x \in D(\omega)] > 0\}$ is countable.*

Proof. For each $n \in \mathbb{N}$ and $x \in \mathbb{R}$, let $\Omega_x^n := \{\omega \in \Omega : F(\omega, x) - F(\omega, x-) > 1/(n+1)\}$. Then $\{\omega \in \Omega : x \in D(\omega)\} = \cup_{n \in \mathbb{N}} \Omega_x^n$. Thus $\mathbb{P}[x \in D(\omega)] = \mathbb{P}\left[\cup_{n \in \mathbb{N}} \Omega_x^n\right] \leq \sum_{n \in \mathbb{N}} \mathbb{P}\left[\Omega_x^n\right]$.

Consider any $n+1$ distinct points $x_1, \ldots, x_{n+1} \in \mathbb{R}$. Suppose that $\omega \in \cap_{i=1}^{n+1} \Omega_{x_i}^n$. Then $\sum_{i=1}^{n+1} [F(\omega, x_i) - F(\omega, x_i-)] > (n+1)/(n+1) = 1$. However, $\sum_{i=1}^{n+1} [F(\omega, x_i) - F(\omega, x_i-)] \leq \sum_{x \in D(\omega)} [F(\omega, x_i) - F(\omega, x_i-)] \leq 1$, and thus $\cap_{i=1}^{n+1} \Omega_{x_i}^n = \varnothing$.

For each $m, n \in \mathbb{N}$, let $D^{m,n} := \{x \in \mathbb{R} : \mathbb{P}\left[\Omega_x^n\right] \geq 1/m\}$. Then $\{x \in \mathbb{R} : \mathbb{P}[x \in D(\omega)] > 0\} = \cup_{m,n \in \mathbb{N}} D^{m,n}$. We show by contradiction that each set $D^{m,n}$ is finite. Suppose that $D^{m,n}$ is infinite; if $D^{m,n}$ is uncountable, choose a countably infinite subset of $D^{m,n}$ and denote the subset with $D^{m,n}$ as well. Consider the countably infinite collection of events $\{\Omega_x^n : x \in D^{m,n}\}$. Recall that for any $n+1$ distinct points $x_1, \ldots, x_{n+1} \in \mathbb{R}$, $\cap_{i=1}^{n+1} \Omega_{x_i}^n = \varnothing$. Also recall that $\mathbb{P}\left[\Omega_x^n\right] \geq 1/m$ for all $x \in D^{m,n}$. Thus it follows from Lemma OA–4 that $|D^{m,n}| \leq mn$. Hence each set $D^{m,n}$ is finite, and therefore $\{x \in \mathbb{R} : \mathbb{P}[x \in D(\omega)] > 0\} = \cup_{m,n \in \mathbb{N}} D^{m,n}$ is countable. $\square$

OA–4 Proof of Proposition 4

Proposition 4. *Consider a family of distributions $\{H(m, \cdot) : m \in \mathbb{M} \subset \mathbb{R}\}$, where $m = \int x H(m, dx)$ is the mean of $H(m, \cdot)$, $\mathbb{M}$ is closed, and $H(m, \cdot)$ is continuous in m with respect to the topology of weak convergence. Suppose that $\{Y^k\}$ and $\{F^k\}$ as in Definition 1 satisfy $F^k(\omega, \cdot) = H(U^k(\omega), \cdot)$ w.p.1, where $U^k := \mathbb{E}[Y^{k+1} \,|\, \mathcal{F}^k]$. Also suppose that $\sup_{k \geq 0} \mathbb{E}[(Y^{k+1})^2 \,|\, \mathcal{F}^k] < Z$ w.p.1, for some integrable random variable Z. Then $\{\hat{H}^k\}$ in* (13)–(14) *is a good forecasting method for $\{Y^k\}$.*

Proof. Note initially that, since H is continuous in the first argument and M^k is $(\mathcal{F}, B)$-measurable, it follows that $\hat{H}^k$ is $(\mathcal{F}, \mathcal{B})$-measurable for all k, i.e., $\hat{H}^k$ is a random distribution function.

Let

$$S^n := \sum_{k=1}^{n} (Y^k - U^{k-1}) .$$

Note that $\{S^n\}$ is a martingale with respect to $\{\mathcal{F}^n\}$, because $E|S^n| < \infty$ and $\mathbb{E}[S^n \,|\, \mathcal{F}^{n-1}] = \mathbb{E}[Y^n - U^{n-1} \,|\, \mathcal{F}^{n-1}] + \mathbb{E}[S^{n-1} \,|\, \mathcal{F}^{n-1}] = \mathbb{E}[Y^n \,|\, \mathcal{F}^{n-1}] - U^{n-1} + S^{n-1} = S^{n-1}$. In addition, $\mathbb{E}[(Y^k -$

$U^{k-1})^2] = \mathbb{E}[(Y^k)^2] - \mathbb{E}[(U^{k-1})^2] \le \mathbb{E}[(Y^k)^2] \le \mathbb{E}[Z]$, and consequently

$$\sum_{k=1}^{\infty} \frac{\mathbb{E}[(Y^k - U^{k-1})^2]}{k^2} \quad \le \quad \sum_{k=1}^{\infty} \frac{\mathbb{E}[Z]}{k^2} \quad < \quad \infty\,.$$

It follows from a strong law of large numbers for martingales (Chow 1967) that $\lim_{n\to\infty} S^n/n = 0$ w.p.1, that is, there is a set $\Omega'' \subset \Omega$ such that $\mathbb{P}[\Omega''] = 0$ and $\lim_{n\to\infty} S^n(\omega)/n = 0$ for all $\omega \in \Omega \backslash \Omega''$. Since $M^n := (1/n)\sum_{k=1}^{n} Y^k$, it follows that

$$M^n(\omega) - \frac{1}{n}\sum_{k=1}^{n} U^{k-1}(\omega) \quad \to \quad 0 \qquad \text{for all } \omega \in \Omega \setminus \Omega''. \tag{OA–13}$$

Let $\Omega''' := \cup_{k\ge1}\{\omega \in \Omega \,:\, F^k(\omega,\cdot) \ne H(U^k(\omega),\cdot)\} \cup \{\omega \in \Omega \,:\, \sup_{k\ge0} \mathbb{E}[(Y^{k+1})^2 \,|\, \mathcal{F}^k](\omega) \ge Z(\omega)\}$, and observe that $\mathbb{P}[\Omega'''] = 0$. Then, for all $\omega \in \Omega^* \setminus \Omega'''$, $H(U^k(\omega),\cdot) = F^k(\omega,\cdot) \stackrel{w}{\to} F^*(\omega,\cdot)$. In addition, for such ω, $\sup_{k\ge0} \int x^2 F^k(\omega, dx) < Z(\omega)$, and hence by Theorem 4.5.2 of Chung (1974),

$$U^k(\omega) \quad = \quad \int xH(U^k(\omega), dx) \quad = \quad \int xF^k(\omega, dx) \quad \to \quad \int xF^*(\omega, dx) \quad =: \quad U(\omega) \qquad \text{for all } \omega \in \Omega^* \setminus \Omega'''. \tag{OA–14}$$

Therefore, for all $\omega \in \Omega^* \setminus \Omega'''$, it holds that $F^*(\omega,\cdot) = H(U(\omega),\cdot)$, because $H(m,\cdot)$ is continuous in m.

Let $\Omega' = \Omega'' \cup \Omega'''$, and observe that $\mathbb{P}[\Omega'] = 0$. Then $M^k(\omega) \to U(\omega)$ for all $\omega \in \Omega^* \setminus \Omega'$ by (OA–13) and (OA–14). Again using the continuity of $H(m,\cdot)$ in m, it follows that $\hat{H}^k(\omega,\cdot) := H(M^k(\omega),\cdot) \stackrel{w}{\to} H(U(\omega),\cdot) = F^*(\omega,\cdot)$ for all $\omega \in \Omega^* \setminus \Omega'$, which proves that $\{\hat{H}^k\}$ is a good forecasting method for $\{Y^k\}$. □

OA–5 Remark Regarding Proposition 5

We briefly explain the difficulties in obtaining results for cases not covered by the proposition.

In the $\beta < 1$ case, note that $f^j > 1$ for all j. Thus, if $\alpha > 0$, then $g^k > \sum_{j=m}^{k} \alpha/j$ and hence $g^k \to \infty$ as $k \to \infty$. If $\alpha < 0$, then $g^k < \sum_{j=m}^{k} \alpha/j$ and hence $g^k \to -\infty$ as $k \to \infty$. Thus, if $\alpha \ne 0$, then even if we use the martingale convergence theorem to establish that, w.p.1, $f^k L^k - g^k \to A$, where A is a finite random variable, it does not establish the asymptotic behavior of L^k.

Next consider the case with $\beta > 1$. Note that $i/(i-1+\beta) \in (0,1)$ for all i, so $f^k \in (0,1)$ for all k. Let $a_i := i/(i-1+\beta)$. Then $\sum_{i=1}^{\infty}(1-a_i) = \sum_{i=1}^{\infty}(\beta-1)/(i-1+\beta) = \infty$. Thus

$f^k = \prod_{i=1}^k a_i \to 0$ as $k \to \infty$. Next, consider

$$\begin{aligned}
\log(f^k) &= \sum_{i=1}^k \log\left(\frac{i}{i-1+\beta}\right) = \sum_{i=1}^k \log\left(1 + \frac{1-\beta}{i-1+\beta}\right) \\
&\leq \sum_{i=1}^k \frac{1-\beta}{i-1+\beta} \leq -(\beta-1)\int_1^{k+1} \frac{1}{x-1+\beta}\,dx \\
&= -(\beta-1)\left[\log(k+\beta) - \log(\beta)\right] \\
&= \log\left[\left(\frac{\beta}{k+\beta}\right)^{\beta-1}\right].
\end{aligned}$$

It follows that $f^k \leq \left(\dfrac{\beta}{k+\beta}\right)^{\beta-1}$ and hence

$$\frac{g^k}{\alpha} = \sum_{j=1}^k \frac{1}{j} f^j \leq \sum_{j=1}^k \frac{1}{j}\left(\frac{\beta}{j+\beta}\right)^{\beta-1} \leq \sum_{j=1}^\infty \frac{1}{j}\left(\frac{\beta}{j+\beta}\right)^{\beta-1} < \infty.$$

In addition $\{g^k\}$ is non-decreasing, and thus $g^k \to \bar{g}$ as $k \to \infty$, where $|\bar{g}| < \infty$. Therefore, if $\sup_k \mathbb{E}|f^k L^k - g^k| < \infty$, then w.p.1, $f^k L^k - g^k \to A$ as $k \to \infty$, where A is a finite random variable. Then $f^k L^k \to B$ as $k \to \infty$, where B is a finite random variable. Recall that $f^k \in (0,1)$ for all k, and $f^k \to 0$ as $k \to \infty$. Thus, if $B(\omega) < 0$, then $L^k(\omega) \to -\infty$; and if $B(\omega) > 0$, then $L^k(\omega) \to \infty$. However, if $B(\omega) = 0$, then we need more information to determine the asymptotic behavior of L^k.

OA–6 Proof of Lemma 4

Lemma 4. *Consider a sequence of distribution functions $\{F^k\} \subset \mathcal{P}(\mathbb{R})$ such that $F^k \xrightarrow{w} F \in \mathcal{P}(\mathbb{R})$. For $\gamma \in (0,1)$, let $[q^k, Q^k] := (F^k)^{-1}(\gamma)$, that is, $[q^k, Q^k]$ denotes the set of γ-quantiles of F^k [cf. (2)], and let $[q, Q] := F^{-1}(\gamma)$. Then, $q \leq \liminf_{k\to\infty} q^k \leq \limsup_{k\to\infty} Q^k \leq Q$. That is, for any sequence $\{\xi^k\}$ of γ-quantiles of F^k, $d(\xi^k, F^{-1}(\gamma)) \to 0$ as $k \to \infty$.*

Proof. Consider any $q' < q$. We show that for all k sufficiently large, $q^k > q'$. Let $q^* \in (q', q)$ be a continuity point of F. Then $F(q^*) < \gamma$, and $F^k(q^*) \to F(q^*)$ as $k \to \infty$, and thus for all k sufficiently large, $F^k(q') \leq F^k(q^*) < \gamma$. Hence $q' < q^k$ for all k sufficiently large, and thus $q \leq \liminf_{k\to\infty} q^k$. It follows by a similar argument that $\limsup_{k\to\infty} Q^k \leq Q$. □

OA–7 More on Stochastic Approximation

In this section we show that, under appropriate assumptions, if the distribution of the observed quantity depends on the protection level and if L^k is updated according to (36), then $G(L^k, L^k)$ converges to γ. It follows that if L^k converges then it converges to a random variable L^* that satisfies $\mathbb{P}(L^* \in G^{-1}(L^*, \gamma)) = 1$. In this section we assume that $G(\ell, x) = 0$ for all $x < 0$ and all $\ell \in \mathbb{R}$, and therefore $X^k \geq 0$ w.p.1.

The following result on the convergence of stochastic approximation iterations is given in Proposition 4.1 of Bertsekas and Tsitsiklis (1996).

Proposition OA–1. *Consider the random sequences* $\{S^k\}_{k=1}^{\infty}$ *and* $\{L^k\}_{k=0}^{\infty}$ *in* $\mathbb{R}^n$ *that satisfy* $L^{k+1} = L^k + \xi_k S^{k+1}$*, where* $\{\xi_k\}_{k=0}^{\infty}$ *is a deterministic nonnegative step size sequence that satisfies* $\sum_{k=0}^{\infty} \xi_k = \infty$ *and* $\sum_{k=0}^{\infty} \xi_k^2 < \infty$*. Let* $\mathcal{F}^k$ *denote the* σ*-algebra generated by* $S^1, \ldots, S^k, L^0, \ldots, L^k$*. Consider a function* $V : \mathbb{R}^n \mapsto \mathbb{R}_+$ *with the following properties:*

1. ∇V *is Lipschitz continuous on* $\mathbb{R}^n$*.*
2. *There is a constant* $c > 0$ *such that, w.p.1,*
$$-\nabla V(L^k)^T \mathbb{E}[S^{k+1} \,|\, \mathcal{F}^k] \quad \geq \quad c \|\nabla V(L^k)\|^2$$
for all k*.*
3. *There exist constants* $K_1, K_2 > 0$ *such that, w.p.1,*
$$\mathbb{E}[\|S^{k+1}\|^2 \,|\, \mathcal{F}^k] \quad \leq \quad K_1 + K_2 \|\nabla V(L^k)\|^2$$
for all k*.*

Then the following hold w.p.1:

1. $V(L^k)$ *converges to a random variable* V^* *as* $k \to \infty$*.*
2. $\nabla V(L^k) \to 0$ *as* $k \to \infty$*.*
3. *Every limit point* L^* *of* $\{L^k\}$ *satisfies* $\nabla V(L^*) = 0$*.*

Next we construct a potential function V to study the convergence of (36). Note that by the assumptions we make on G in this section, we have that $F(\ell) = G(\ell, \ell) = 0$ if $\ell < 0$. We also make the following assumption:

Assumption (B2) The function F is Lipschitz continuous, i.e., there exists an $M > 0$ such that $|F(\ell_1) - F(\ell_2)| \leq M|\ell_1 - \ell_2|$ for all $\ell_1, \ell_2 \in \mathbb{R}$.

This essentially says that the rate of change of $G(\ell, \ell)$ with respect to ℓ is bounded for all ℓ. Assumption (B2) is satisfied, for instance, if

$$G(\ell, x) \quad = \quad 1 - e^{-x/m(\ell)}, \qquad x \geq 0, \tag{OA–15}$$

for $\ell \geq 0$, and $G(\ell, \cdot) = G(0, \cdot)$ for $\ell < 0$, i.e., negative protection levels have the same effect as $\ell = 0$. Here $m(\ell) > 0$ for all $\ell \geq 0$ and $r(\ell) := \ell/m(\ell)$ is Lipschitz continuous on $[0, \infty)$. Indeed, note that if $\ell_1, \ell_2 < 0$ then $|F(\ell_1) - F(\ell_2)| = 0$, and if $\ell_1 < 0 \leq \ell_2$ then $|F(\ell_1) - F(\ell_2)| = |F(0) - F(\ell_2)|$, so it suffices to check that F is Lipschitz continuous on $[0,\infty)$, which is indeed the case, because for $\ell_1, \ell_2 \geq 0$, we have $|F(\ell_1) - F(\ell_2)| = |e^{-r(\ell_2)} - e^{-r(\ell_1)}| \leq |r(\ell_2) - r(\ell_1)|$ since $r(\ell_1), r(\ell_2) \geq 0$.

One choice for $m(\ell)$ that satisfies the above conditions is

$$m(\ell) \;:=\; a_1 - a_2 e^{-a_3 \ell} \tag{OA–16}$$

where $a_1 > a_2 \geq 0$ and $a_3 \geq 0$. If the observed quantity X has distribution specified by (OA–15)–(OA–16), then it has properties that so-called "unconstrained demand" for high-price tickets could reasonably be expected to have (it is immaterial how this unconstraining is done — it only matters that it results in X). For instance, $m(\ell)$ increases in ℓ and approaches a constant as $\ell \to \infty$, which is an appealing property since one would not expect the mean demand to grow unboundedly with increasing protection levels.

To see that (OA–16) makes r Lipschitz continuous, note that

$$\begin{aligned}
|r'(\ell)| = \;& \left|\frac{a_1 - a_2 e^{-a_3\ell} - \ell(a_2 a_3 e^{-a_3\ell})}{(a_1 - a_2 e^{-a_3\ell})^2}\right| \;\leq\; \left|\frac{1}{a_1 - a_2 e^{-a_3\ell}}\right| + \left|\frac{\ell(a_2 a_3 e^{-a_3\ell})}{(a_1 - a_2 e^{-a_3\ell})^2}\right| \\
\leq \;& \left|\frac{1}{a_1 - a_2}\right| + \left|\frac{\ell(a_2 a_3 e^{-a_3\ell})}{(a_1 - a_2)^2}\right| \;\leq\; \frac{1}{a_1 - a_2} + \frac{a_2 e^{-1}}{(a_1 - a_2)^2}.
\end{aligned}$$

The final expression follows from the fact that $\ell e^{-a_3\ell}$ is maximized over $[0,\infty)$ at $\ell = 1/a_3$.

At this point we need the following assumption:

Assumption (B3) The quantity $\nu := \min_{\ell\in\mathbb{R}} \int_0^\ell [F(s) - \gamma]\, ds$ is finite.

When $\ell < 0$, we interpret the integral in the above expression for ν as $-\int_\ell^0$. Thus, for any $\ell < 0$, $\int_0^\ell [F(s) - \gamma] ds = -\int_\ell^0 [F(s) - \gamma] ds = -\int_\ell^0 [0 - \gamma] ds = -\ell\gamma > 0$. Hence, Assumption (B3) holds, for example, if there exists an $\ell_0 > 0$ such that $F(\ell) \geq \gamma$ for all $\ell \geq \ell_0$. For instance, this is the case when (OA–15)–(OA–16) specify the distribution of the observed quantity, since $F(\ell) \geq \gamma \Leftrightarrow \ln(1-\gamma) \geq -r(\ell) \Leftrightarrow -m(\ell)\ln(1-\gamma) \leq \ell$, which does indeed hold for ℓ sufficiently large. Under the assumptions of van Ryzin and McGill (2000), Assumptions (B2) and (B3) hold. Specifically, Assumption (B3) holds since it is always the case that $F(\ell) \geq \gamma$ for all ℓ large enough when G does not depend on ℓ.

Consider the function $V : \mathbb{R} \mapsto \mathbb{R}_+$ defined by

$$V(\ell) \;:=\; \int_0^\ell [F(s) - \gamma]\, ds \;-\; \nu. \tag{OA–17}$$

Next we verify that V satisfies the conditions in Proposition OA–1. Note that $V'(\ell) = F(\ell) - \gamma$.

1. V' is Lipschitz continuous, since by Assumption (B2) F is Lipschitz continuous.

2. Note from (36) that $S^{k+1} = \gamma - \mathbb{I}_{\{X^{k+1} \leq L^k\}}$. Thus

$$\mathbb{E}[S^{k+1} \,|\, \mathcal{F}^k] \;=\; \gamma - \mathbb{P}[X^{k+1} \leq L^k \,|\, L^k] \;=\; \gamma - G(L^k, L^k) \;=\; \gamma - F(L^k) \;=\; -V'(L^k).$$

3. Note that $S^{k+1} \in (-1,1)$ w.p.1, and thus there exist constants $K_1, K_2 > 0$ such that

$$\mathbb{E}[(S^{k+1})^2 \,|\, \mathcal{F}^k] \quad \leq \quad K_1 + K_2[V'(L^k)]^2$$

Recall that the stepsizes ξ_k satisfy $\sum_k \xi_k = \infty$ and $\sum_k \xi_k^2 < \infty$, and thus we obtain the conclusions of Proposition OA–1. Specifically, we have the following.

Proposition OA–2. *Suppose that Assumptions* (B2) *and* (B3) *hold and that the protection levels are updated according to* (36)*. Then $G(L^k, L^k) \to \gamma$ w.p.1, and every limit point L^* of $\{L^k\}$ satisfies $G(L^*, L^*) = \gamma$, that is, $L^* \in G^{-1}(L^*, \gamma)$.*

Note that Propositions 8 and 9 require the existence of a deterministic quantity ℓ^* that satisfies assumption 3 in Proposition 8 or Assumption (B1) respectively, and that convergence of L^k to this deterministic quantity ℓ^* is then established. In contrast, Propositions OA–1 and OA–2 do not require the existence of such a deterministic quantity, and do not establish convergence of L^k.

OA–8 Proofs for Stochastic Comparisons and Pathwise Comparisons

Lemma OA–5. *For any two $\mathcal{P}(\mathbb{R})$-valued random elements $H_1 \sim P_1$ and $H_2 \sim P_2$, $H_1 \preceq_{\rm st} H_2$ implies that $P_1[H_1(x) \geq \alpha] \geq P_2[H_2(x) \geq \alpha]$ for all $x, \alpha \in \mathbb{R}$.*

Proof. Fix any $x, \alpha \in \mathbb{R}$, and let $f : \mathcal{P}(\mathbb{R}) \mapsto \mathbb{R}$ be given by $f(h) := -\mathbb{I}_{\{h(x) \geq \alpha\}}$. Clearly f is bounded, and it follows from the characterization of $\leq_{\rm st}$ that f is nondecreasing. Moreover, by the argument in the proof of Proposition 17(v) we have that f is measurable.

Consider any two $\mathcal{P}(\mathbb{R})$-valued random elements $H_1 \preceq_{\rm st} H_2$. Then it follows that

$$P_1[H_1(x) \geq \alpha] \quad = \quad -\mathbb{E}_{P_1}[f(H_1)] \quad \geq \quad -\mathbb{E}_{P_2}[f(H_2)] \quad = \quad P_2[H_2(x) \geq \alpha].$$

□

To simplify the exposition below, suppose that L^k and $\underline{L}^k$ are chosen to be the smallest elements of the set of γ-quantiles of $\hat{H}^k$ and $\underline{\hat{H}}^k$ respectively, that is, $L^k \equiv \min\left\{x \in \mathbb{R} \,:\, \hat{H}^k(x) \geq \gamma\right\}$ and $\underline{L}^k \equiv \min\left\{x \in \mathbb{R} \,:\, \underline{\hat{H}}^k(x) \geq \gamma\right\}$.

Lemma OA–6. *Suppose that $\underline{G}(\underline{\ell}, \cdot) \leq_{\rm st} G(\ell, \cdot)$ for all $\underline{\ell} \leq \ell$, and that the empirical distribution is used for both $\hat{H}$ and $\underline{\hat{H}}$, that is $\hat{H}^k(x) := k^{-1} \sum_{j=1}^k \mathbb{I}_{\{X^j \leq x\}}$ and $\underline{\hat{H}}^k(x) := k^{-1} \sum_{j=1}^k \mathbb{I}_{\{\underline{X}^j \leq x\}}$. If $\underline{\hat{H}}^k \preceq_{\rm st} \hat{H}^k$, then*

$$\begin{aligned} \underline{L}^k &\leq_{\rm st} L^k \\ \underline{G}(\underline{L}^k, \cdot) &\preceq_{\rm st} G(L^k, \cdot) \\ \underline{X}^{k+1} &\leq_{\rm st} X^{k+1} \\ \underline{\hat{H}}^{k+1} &\preceq_{\rm st} \hat{H}^{k+1} \end{aligned}$$

Proof. Suppose $\{\underline{\hat{H}}^k, \underline{L}^k, \underline{X}^k\}$ is defined on probability space $(\underline{\Omega}, \underline{\mathcal{F}}, \underline{\mathbb{P}})$, and let $\underline{\mathbb{E}}$ denote expectation with respect to $\underline{\mathbb{P}}$. Suppose $\underline{\hat{H}}^k \preceq_{\text{st}} \hat{H}^k$. Then it follows from Lemma OA–5 that for all $x \in \mathbb{R}$,

$$\underline{\mathbb{P}}[\underline{L}^k \leq x] \quad = \quad \underline{\mathbb{P}}[\underline{\hat{H}}^k(x) \geq \gamma] \quad \geq \quad \mathbb{P}[\hat{H}^k(x) \geq \gamma] \quad = \quad \mathbb{P}[L^k \leq x].$$

That is, $\underline{L}^k \leq_{\text{st}} L^k$. By assumption, $\underline{G}(\underline{\ell}, \cdot) \leq_{\text{st}} G(\ell, \cdot)$ for all $\underline{\ell} \leq \ell$, and thus it follows easily from Kamae et al. (1977), Theorem 1 [in particular, the equivalence of (i) and (iv)], that $\underline{G}(\underline{L}^k, \cdot) \preceq_{\text{st}} G(L^k, \cdot)$. For $h \in \mathcal{P}(\mathbb{R})$, define $\ell(h) = \min\{x \in \mathbb{R} : h(x) \geq \gamma\}$. Then $\ell(\underline{h}) \leq \ell(h)$ for all $\underline{h} \leq_{\text{st}} h$. Hence, for $\underline{h} \leq_{\text{st}} h$ it holds that

$$\underline{\mathbb{P}}[\underline{X}^{k+1} \leq x | \underline{\hat{H}}^k = \underline{h}] \quad = \quad \underline{G}(\ell(\underline{h}), x) \quad \geq \quad G(\ell(h), x) \quad = \quad \mathbb{P}[X^{k+1} \leq x | \hat{H}^k = h].$$

Since $\underline{\hat{H}}^k \preceq_{\text{st}} \hat{H}^k$, it now follows from Proposition 1 of Kamae et al. (1977) that $\underline{X}^{k+1} \leq_{\text{st}} X^{k+1}$ and $(\underline{X}^{k+1}, \underline{\hat{H}}^k) \prec (X^{k+1}, \hat{H}^k)$ where $\prec$ denotes the usual stochastic order with the coordinate-wise partial ordering on $\mathbb{R} \times \mathcal{P}(\mathbb{R})$ — see page 901 of Kamae et al. (1977). Note that $\underline{\hat{H}}^{k+1} = \eta_k(\underline{X}^{k+1}, \underline{\hat{H}}^k)$ and $\hat{H}^{k+1} = \eta_k(X^{k+1}, \hat{H}^k)$ where $\eta_k : \mathbb{R} \times \mathcal{P}(\mathbb{R}) \mapsto \mathcal{P}(\mathbb{R})$ is defined by

$$\eta_k(x, h) \quad = \quad \frac{k}{k+1} h + \frac{1}{k+1} \mathbb{I}_{\{x \leq \ \cdot\}}$$

and observe that η_k is increasing on $\mathbb{R} \times \mathcal{P}(\mathbb{R})$; i.e., $\eta_k(\underline{x}, \underline{h}) \leq_{\text{st}} \eta(x, h)$ when $\underline{x} \leq x$ and $\underline{h} \leq_{\text{st}} h$. It follows that for bounded increasing $f : \mathcal{P}(\mathbb{R}) \mapsto \mathbb{R}$,

$$\underline{\mathbb{E}}[f(\underline{\hat{H}}^{k+1})] \quad = \quad \underline{\mathbb{E}}[(f \circ \eta_k)(\underline{X}^{k+1}, \underline{\hat{H}}^k)] \quad \leq \quad \mathbb{E}[(f \circ \eta_k)(X^{k+1}, \hat{H}^k)] \quad = \quad \mathbb{E}[f(\hat{H}^{k+1})],$$

where the inequality follows from the fact that $f \circ \eta_k$ is bounded and increasing on $\mathbb{R} \times \mathcal{P}(\mathbb{R})$ and $(\underline{X}^{k+1}, \underline{\hat{H}}^k) \prec (X^{k+1}, \hat{H}^k)$. Hence, $\underline{\hat{H}}^{k+1} \preceq_{\text{st}} \hat{H}^{k+1}$. □

Proposition 12 follows from Lemma OA–6.

Proposition 12 (Stochastic comparison with empirical distributions). *Suppose $\underline{G}(\underline{\ell}, \cdot) \leq_{\text{st}} G(\ell, \cdot)$ for all $\underline{\ell} \leq \ell$, and the empirical distribution is used for both $\hat{H}$ and $\underline{\hat{H}}$, that is, $\hat{H}^k(x) := k^{-1} \sum_{j=1}^k \mathbb{I}_{\{X^j \leq x\}}$ and $\underline{\hat{H}}^k(x) := k^{-1} \sum_{j=1}^k \mathbb{I}_{\{\underline{X}^j \leq x\}}$. If $\underline{L}^0 \leq_{\text{st}} L^0$, then*

$$\begin{aligned} \underline{G}(\underline{L}^k, \cdot) &\preceq_{\text{st}} G(L^k, \cdot) \\ \underline{X}^{k+1} &\leq_{\text{st}} X^{k+1} \\ \underline{\hat{H}}^{k+1} &\preceq_{\text{st}} \hat{H}^{k+1} \\ \underline{L}^{k+1} &\leq_{\text{st}} L^{k+1} \end{aligned}$$

for all $k = 0, 1, \ldots$.

Proposition 13 (Stochastic comparison with affine updates). *Suppose that* $\mu : \mathbb{R} \mapsto \mathbb{R}$ *satisfies* $\mu(\ell) \le \ell$ *for all* ℓ. *Suppose that* $\underline{G}(\underline{\ell}, \cdot) = G(\mu(\underline{\ell}), \cdot)$, *and that* $G(\underline{\ell}, \cdot) \le_{\mathrm{st}} G(\ell, \cdot)$ *for all* $\underline{\ell} \le \ell$. *Suppose that* $\hat{H}^k = G(M^k, \cdot)$ *and* $\underline{\hat{H}}^k = G(\underline{M}^k, \cdot)$, *where* $M^k = k^{-1} \sum_{j=1}^k X^j$ *and* $\underline{M}^k = k^{-1} \sum_{j=1}^k \underline{X}^j$. *If* $\underline{L}^0 \le_{\mathrm{st}} L^0$, *then*

$$\begin{aligned} \underline{G}(\underline{L}^k, \cdot) &\preceq_{\mathrm{st}} G(L^k, \cdot) && \text{(OA–18)} \\ \underline{X}^{k+1} &\le_{\mathrm{st}} X^{k+1} && \text{(OA–19)} \\ \underline{M}^{k+1} &\le_{\mathrm{st}} M^{k+1} && \text{(OA–20)} \\ \underline{\hat{H}}^{k+1} &\preceq_{\mathrm{st}} \hat{H}^{k+1} && \text{(OA–21)} \\ \underline{L}^{k+1} &\le_{\mathrm{st}} L^{k+1} && \text{(OA–22)} \end{aligned}$$

for all $k = 0, 1, \ldots$.

Proof. The proof is by induction; (OA–18)–(OA–22) hold for $k = 0$. For the inductive step, suppose that (OA–18)–(OA–22) hold for $k-1$ and consider a general k. Since $\underline{L}^k \le_{\mathrm{st}} L^k$, Theorem 1 of Kamae et al. (1977) implies that $\mu(\underline{L}^k) \le_{\mathrm{st}} L^k$ and $G(\mu(\underline{L}^k), \cdot) \preceq_{\mathrm{st}} G(L^k, \cdot)$. Hence, $\underline{G}(\underline{L}^k, \cdot) \preceq_{\mathrm{st}} G(L^k, \cdot)$. For $\underline{m} \le m$, we have

$$\mathbb{P}(\underline{X}^{k+1} \le x | \underline{M}^k = \underline{m}) \;=\; G(\,\ell(G(\underline{m}, \cdot))\,, x\,) \;\ge\; G(\,\ell(G(m, \cdot))\,, x\,) \;=\; \mathbb{P}(X^{k+1} \le x | M^k = m),$$

where $\ell(h) = \min\{x \in \mathbb{R} : h(x) \ge \gamma\}$ for $h \in \mathcal{P}(\mathbb{R})$. Proposition 1 of Kamae et al. (1977) implies that $\underline{X}^{k+1} \le_{\mathrm{st}} X^{k+1}$ and $(\underline{X}^{k+1}, \underline{M}^k) \prec (X^{k+1}, M^k)$, where $\prec$ here denotes the usual stochastic order on $\mathbb{R}^2$. Observe that $M^{k+1} = \varphi_k(X^{k+1}, M^k)$ and $\underline{M}^{k+1} = \varphi_k(\underline{X}^{k+1}, \underline{M}^k)$ where

$$\varphi_k(x, m) \;=\; \frac{k}{k+1} m + \frac{1}{k+1} x.$$

It follows that $\underline{M}^{k+1} \le_{\mathrm{st}} M^{k+1}$, and hence $\underline{\hat{H}}^{k+1} \preceq_{\mathrm{st}} \hat{H}^{k+1}$. Finally, $\mathbb{P}[\underline{L}^{k+1} \le x] = \mathbb{P}[\underline{\hat{H}}^{k+1}(x) \ge \gamma] \ge \mathbb{P}[\hat{H}^{k+1}(x) \ge \gamma] = \mathbb{P}[L^{k+1} \le x]$, so $\underline{L}^{k+1} \le_{\mathrm{st}} L^{k+1}$. $\square$

Proposition 14 (Pathwise comparison). *Consider any* $\omega \in \Omega$ *such that, for any* k, $\underline{L}^k(\omega) \le L^k(\omega)$ *implies that* $\underline{X}^{k+1}(\omega) \le X^{k+1}(\omega)$. *Suppose that the forecasting method used in both sequences satisfies the following condition for all* k: *If* $(\underline{X}^1(\omega), \ldots, \underline{X}^k(\omega)) \le (X^1(\omega), \ldots, X^k(\omega))$, *then* $\underline{\hat{H}}^k(\omega, \cdot) \le_{\mathrm{st}} \hat{H}^k(\omega, \cdot)$. *If* $\underline{L}^0(\omega) \le L^0(\omega)$, *then*

$$\begin{aligned} \underline{X}^k(\omega) &\le X^k(\omega) \\ \underline{\hat{H}}^k(\omega, \cdot) &\le_{\mathrm{st}} \hat{H}^k(\omega, \cdot) \\ \underline{L}^k(\omega) &\le L^k(\omega) \end{aligned}$$

for all $k = 1, 2, \ldots$.

Proof. The result follows from induction on k. □

References

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

Billingsley, P. 1968. *Convergence of Probability Measures*. John Wiley & Sons, New York.

Chow, Y. S. 1967. On a strong law of large numbers for martingales. *Ann. Math. Statist.* **38**(2) 610.

Chung, K. L. 1974. *A Course in Probability Theory*. 2nd ed. Academic Press, New York.

Dudley, R. M. 2002. *Real Analysis and Probability*. Cambridge Univ. Press, Cambridge, UK.

Kamae, T., U. Krengel, G. L. O'Brien. 1977. Stochastic inequalities on partially ordered spaces. *Ann. Probab.* **5**(6) 899–912.

van Ryzin, G., J. McGill. 2000. Revenue management without forecasting or optimization: An adaptive algorithm for determining airline seat protection levels. *Management Sci.* **46**(6) 760–775.