Electronic Companion

EC.1. Proof of Lemma 1

In what follows, we will only show the existence of a proper set $\mathcal{Q}$ under the single-product setting; the argument can be easily extended to the multiple-product setting. Let $F^t : \Omega_p \to [0,1]$ denote the CDF for pricing decision during period t under the optimal control π^*. Also, let $\bar{r}_j^t$ and $\bar{\lambda}_j^t$ denote the expected revenue and demand rate from location j during period t under π^* (since we only consider the single-product setting, there is no need to use subscript k), i.e.,

$$\bar{r}_j^t := \mathbb{E}^{\pi^*}[R_j^t(p^t)] = \int_{\Omega_p} r_j(p)\ dF^t(p) \quad \text{and} \quad \bar{\lambda}_j^t := \mathbb{E}^{\pi^*}[D_j^t(p^t)] = \int_{\Omega_p} \lambda_j(p)\ dF^t(p).$$

To prove Lemma 1, we first show that there exist weight vectors $\{\boldsymbol{\alpha}^t\}$ such that, for the uniform grid $\mathcal{Q}^u$ defined in Section 4 and some sufficiently small $\epsilon_r, \epsilon_\lambda > 0$, the following hold:

$$\left|\bar{r}_j^t - \sum_{m=1}^M \alpha_m^t r_j(q_m^u)\right| = \left|\int_{\Omega_p} p\lambda_j(p)\ dF^t(p) - \sum_{m=1}^M \alpha_m r_j(q_m^u)\right| \le \epsilon_r \quad \forall j, t, \tag{EC.1}$$

$$\left|\bar{\lambda}_j^t - \sum_{m=1}^M \alpha_m^t \lambda_j(q_m^u)\right| = \left|\int_{\Omega_p} \lambda_j(p)\ dF^t(p) - \sum_{m=1}^M \alpha_m^t \lambda_j(q_m^u)\right| \le \epsilon_\lambda \quad \forall j, t, \tag{EC.2}$$

$$\sum_{m=1}^M \alpha_m^t = 1, \qquad \alpha_m^t \ge 0, \quad \forall m, t. \tag{EC.3}$$

Define a uniform partition of the interval Ω_p as $\Omega_p = \cup_{m=1}^M \mathcal{P}_m := \left[\cup_{m=1}^{M-1} [p_\ell + (m-1)\Delta_q, p_\ell + m\Delta_q)\right] \cup [p_u - \Delta_q, p_u]$, where $\Delta_q := (p_u - p_\ell)/M$ is the length of the sub-intervals. Then the uniform price grid can be defined as $\mathcal{Q}^u := (p_\ell + (m - 1/2)\Delta_q)_{m=1}^M$. Consider a choice of weight vector $\alpha_m^t = \int_{\mathcal{P}_m} dF^t(p)$. Note that (EC.3) is satisfied immediately by definition. We now show that the combination of $\mathcal{Q}^u$ and $\boldsymbol{\alpha}^t$ defined above satisfy (EC.1) and (EC.2). By definition, for all $j \in [J]$, we have

$$\begin{aligned}\left|\bar{\lambda}_j^t - \sum_{m=1}^M \alpha_m^t \lambda_j(q_m^u)\right| &= \left|\int_{\Omega_p} \lambda_j(p)\ dF^t(p) - \sum_{m=1}^M \boldsymbol{\alpha}_m^t \lambda_j(q_m^u)\right| = \left|\sum_{m=1}^M \int_{\mathcal{P}_m} (\lambda_j(p) - \lambda_j(q_m^u))\ dF^t(p)\right| \\ &\le \sum_{m=1}^M \int_{\mathcal{P}_m} |\lambda_j(p) - \lambda_j(q_m^u)|\ dF^t(p) \le \lambda_u \Delta_q.\end{aligned}$$

where the first inequality follows from triangular inequality and the last inequality follows from Assumption A1 together with $\lambda_u := \max_{j\in[J], p\in\Omega_p} |\lambda_j'(p)|$. By similar argument, since $|r_j'(p)| \le |\lambda_j(p) + p\lambda_j'(p)| \le 1 + p_u\lambda_u$ for all $p \in \Omega_p$, it is not difficult to show that (EC.1) is satisfied for $\epsilon_r = (1 + p_u\lambda_u)\Delta_q$.

We now show that the choices of $\mathcal{Q}^u$ and $\boldsymbol{\alpha}^t$ above guarantees a good approximation. The fulfillment LP under the uniform discretization we construct is as follows:

$$\mathbf{FC}^A := \min_{\{x_{ij}^t\}} \left\{ \sum_{t=1}^T \sum_{i=0}^I \sum_{j=1}^J c_{ij} x_{ij}^t \ :\ \sum_{i=0}^I x_{ij}^t = \sum_{m=1}^M \alpha_m^t \lambda_j(q_m^u),\ \sum_{t=1}^T \sum_{j=1}^J x_{ij}^t \le C_i,\ 0 \le x_{ij}^t \le 1 \right\}.$$

On the other hand, the CDF of the fulfillment assignment under π^* can be solve by the following LP:

$$\mathbf{FC}^O := \min_{\{x_{ij}^t\}} \left\{ \sum_{t=1}^T \sum_{i=0}^I \sum_{j=1}^J c_{ij} x_{ij}^t \ :\ \sum_{i=0}^I x_{ij}^t = \bar{\lambda}_j^t,\ \sum_{t=1}^T \sum_{j=1}^J x_{ij}^t \le C_i,\ 0 \le x_{ij}^t \le 1 \right\}.$$

The only difference between $\mathbf{FC}^A$ and $\mathbf{FC}^O$ is on the RHS of fulfillment constraint. Note that both $\mathbf{FC}^A$ and $\mathbf{FC}^O$ have stationary optimal solution. Then given (EC.2) and the perturbation theory of the optimal objective value of LP (see e.g. Theorem 10.5 in Schrijver 1998), $\mathbf{FC}^A - \mathbf{FC}^O \leq IJT\lambda_u\Delta_q$. So the approximation error is bounded as follows:

$$\begin{aligned}
&\mathcal{J}^* - \mathcal{J}^{ALP} \\
&\leq \left[\sum_{t=1}^{T}\sum_{j=1}^{J}\bar{r}_j^t - \mathbf{FC}^O\right] - \left[\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^t r_j(q_m^u) - \mathbf{FC}^A\right] \leq \sum_{t=1}^{T}\sum_{j=1}^{J}\left|\bar{r}_j^t - \sum_{m=1}^{M}\alpha_m^t r_j(q_m^u)\right| + \left(\mathbf{FC}^A - \mathbf{FC}^O\right) \\
&\leq \left[JT(1+p_u\lambda_u) + IJT\lambda_u\right]\Delta_q \leq \frac{(p_u - p_\ell)JT\left[1+\lambda_u(p_u+I)\right]}{M}.
\end{aligned}$$

The proof is concluded by letting $M = \lceil (p_u - p_\ell)JT\left[1+\lambda_u(p_u+I)\right]/\epsilon\rceil$. For general K, the number of discrete prices required to reach an error of ϵ is at most $\lceil \left[(p_u - p_\ell)JT(1+K\Phi_1(p_u+IK))\right]^K/\epsilon^K\rceil$, $\Phi_1 = \max_{p\in\Omega_p,\ j\in[J],\ k,\ell\in[K]} |\partial\lambda_{jk}(\boldsymbol{p})/\partial p_\ell| > 0$ (it is finite by Assumption A1) □

EC.2. Proof of Theorem 1

Let $\mathcal{Q}^u$ be uniform grid defined in Section 4. Without loss of generality, we assume that $T=1$. We consider a variant of RPF (V-RPF) defined as follow: during period t, fulfill the order from location j according to $\sigma_k^t(j)$ regardless of the availability of the corresponding FC; if the FC runs out of inventory, the retailer incurs a penalty cost of $\bar{c} := 2\cdot \max_{j\in[J],k[K]} c_{0jk}$. In other words, V-RPF incurs the same revenue as RPF, yet no smaller fulfillment cost. Consequently, the loss can be bounded as follows:

$$\begin{aligned}
&\mathcal{J}^{ALP}(\theta) - \mathcal{J}^{RPF}(\theta) \leq \mathcal{J}^{ALP}(\theta) - \mathcal{J}^{V-RPF}(\theta) \\
&= \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^* r_j(\boldsymbol{q}_m^*) - \sum_{t=1}^{\theta}\sum_{j=1}^{J}(\boldsymbol{p}^t)^\top \boldsymbol{D}_j^t(\boldsymbol{p}^t)\right] + \ \bar{c}\ \mathbb{E}\left[\sum_{i=1}^{I}\sum_{k=1}^{K}\left(\sum_{t=1}^{\theta}\sum_{j=1}^{J}X_{ijk}^t - C_{ik}(\theta)\right)^+\right] \\
&\qquad + \ \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{j=1}^{J}\Delta R_j^t\right] + \bar{c}\ \mathbb{E}\left[\sum_{i=0}^{I}\sum_{k=1}^{K}\left(\sum_{t=1}^{\theta}\sum_{j=1}^{J}X_{ijk}^t - C_{ik}(\theta)\right)^+\right] + \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}\Delta X_{ijk}^t\right],
\end{aligned}$$

where $\Delta R_j^t := \sum_{m=1}^{M}\alpha_m^* r_j(\boldsymbol{q}_m^*) - (\boldsymbol{p}^t)^\top \boldsymbol{D}_j^t(\boldsymbol{p}^t)$, and $\Delta X_{ijk}^t := X_{ijk}^t - x_{ijk}^*$. By definition of RPF, $\mathbb{E}[\Delta R_j^t] = \mathbb{E}[\Delta X_{ijk}^t] = 0$. As for the last term, by triangular inequality,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=0}^{I}\sum_{k=1}^{K}\left(\sum_{t=1}^{\theta}\sum_{j=1}^{J}X_{ijk}^t - C_{ik}(\theta)\right)^+\right] \leq \mathbb{E}\left[\sum_{i=0}^{I}\sum_{k=1}^{K}\left(\sum_{t=1}^{\theta}\sum_{j=1}^{J}X_{ijk}^t - \theta\sum_{j=1}^{J}x_{ijk}^*\right)^+\right] + \mathbb{E}\left[\sum_{i=0}^{I}\sum_{k=1}^{K}\left(\theta\sum_{j=1}^{J}x_{ijk}^* - C_i(\theta)\right)^+\right] \\
&\leq \sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\mathbb{E}\left[\left(\sum_{t=1}^{\theta}X_{ijk}^t - x_{ijk}^*\right)^+\right] + 0 \leq \sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left[\mathbf{Var}\left(\sum_{t=1}^{\theta}\Delta X_{ijk}^t\right)\right]^{1/2} = O\left(\sqrt{\theta}\right),
\end{aligned}$$

where the second inequality follows from the inventory constraint in ALP, the last inequality follows because ΔX_{ijk}^t's are independent and bounded from above by $D_{jk}^t \leq 1$. This completes the proof. □

EC.3. Proof of Theorem 2

Let $T=1$. Per our discussion in Section 6, we can assume $\sum_{j=1}^{J} x^*_{ijk} = C_{ik}$ without loss of generality. Let $\boldsymbol{C}_i^t(\theta)$ be the on-hand inventory level in FC i at the beginning of period t for a problem with size θ. By definition, we have $\boldsymbol{C}_i^1(\theta) = \theta \boldsymbol{C}_i$. Fix $\theta > 0$. We divide our proof into several steps.

Step 1

In this step, we state and prove two key observations that are useful in helping us to express the evolution of pricing and fulfillment decisions over time. We call an $\text{FLP}^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ to be "balanced" if it satisfies (i) $\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha^*_m \lambda_{jk}(\boldsymbol{q}^t_m) = \sum_{i=0}^{I} C^t_{ik}/(T-t+1)$ for all k, and (ii) $C^t_{ik} > 0$ for all i,k. We make our first observation regarding the solution of a balanced FLP^t.

Observation EC.1. *The optimal solution $\boldsymbol{x}^t$ to a non-DR-degenerate balanced FLP$^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ has the following property: For every $k \in [K]$, there are exactly $I+J$ strictly positive components in $(x^t_{ijk})_{i\in\{0\}\cup[I], j\in[J]}$, with the other components equal to zero. Moreover, the inventory constraints are all binding.*

Proof. Note that $\text{FLP}^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ is separable over k, so solving $\text{FLP}^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ is equivalent to solving K sub-problems defined below:

$$\text{FLP}^t_k(\mathcal{Q}^t, C^t_k) := \left\{ \min_{x_{ijk}\geq 0} \sum_{i=0}^{I}\sum_{j=1}^{J} c_{ijk}x_{ijk} : \sum_{i=0}^{I} x_{ijk} = \sum_{m=1}^{M} \alpha^*_m \lambda_{jk}(\boldsymbol{q}^t_m), \sum_{j=1}^{J} x_{ijk} \leq C^t_{ik}/(T-t+1) \right\}.$$

Since $\text{FLP}^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ is balanced, all the inventory constraints in $\text{FLP}^t_k(\mathcal{Q}^t, \boldsymbol{C}^t_k)$ must be binding. Since $\text{FLP}^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ is non-DR-degenerate and separable over k, $\text{FLP}^t_k(\mathcal{Q}^t, \boldsymbol{C}^t_k)$ is also non-degenerate for each k. Thus, Observation EC.1 follows directly from the standard result on transportation LP (see Corollary 7.2 in Dantzig and Thapa 2006). $\square$

Let $\boldsymbol{x}_k = (x_{ijk})_{i\in\{0\}\cup[I], j\in[J]}$ and $\boldsymbol{c}_k = (c_{ijk})_{i\in\{0\}\cup[I], j\in[J]}$. Given our assumptions in the statement of Theorem 2 and at the beginning of this section, $\text{FLP}^1(\mathcal{Q}^u, \boldsymbol{C})$ is non-DR-degenerate and balanced. Thus, for all k, $\text{FLP}^1_k(\mathcal{Q}^u, \boldsymbol{C}_k)$ are non-degenerate and has $I+J$ non-zero components in $\boldsymbol{x}^*_k$ (since there are $I+J+1$ constraints with exactly one redundant). Let A_k and $\boldsymbol{V}_k$ denote the coefficient matrix and the RHS of inventory constraints in FLP^1_k. Let $\bar{A}_k$ be the matrix where we delete the $(J+1)^{th}$ row from A_k, i.e., the row corresponding to the inventory constraint on FC 0, and $\bar{\boldsymbol{V}}_k$ be the vector where we delete C_{0k}/θ from $\boldsymbol{V}_k$. This constraint is redundant, since any $\boldsymbol{x}_k$ satisfying the system of equations $\bar{A}_k\boldsymbol{x}_k = \bar{\boldsymbol{V}}_k$ automatically satisfies $\sum_{j=1}^{J} x^t_{0jk} = C_{0k}/\theta$ (the deleted constraint). Since the deleted constraint is redundant, FLP^1_k is equivalent to $\{\min\ \boldsymbol{c}_k^\top \boldsymbol{x}_k : \bar{A}_k\boldsymbol{x}_k = \bar{\boldsymbol{V}}_k,\ \boldsymbol{x} \succeq \boldsymbol{0}\}$; moreover, by Lemma 7.1 in Dantzig and Thapa (2006), $\bar{A}_k$ has linearly independent rows. Let $\mathcal{B}_k = \{(i,j) : 0 < x^*_{ijk} < 1\}$ and $\mathcal{N}_k = \{(i,j) : x^*_{ijk} = 0\}$ be the indices of the optimal basic and non-basic variables respectively. Without loss of generality, we assume that $\bar{A}_k$ is written as $[B_k, N_k]$ where B_k and N_k are the sub-matrices of $\bar{A}_k$ corresponding to the basic and non-basic indices in $\mathcal{B}_k$ and $\mathcal{N}_k$ respectively. Following the same decomposition, the optimal solution can be represented as $\boldsymbol{x}^*_k = [\boldsymbol{x}^*_{k,B}, \boldsymbol{x}^*_{k,N}]$, where $\boldsymbol{x}^*_{k,B} = B_k^{-1}\bar{\boldsymbol{V}}_k$ and $\boldsymbol{x}^*_{k,N} = \boldsymbol{0}$ (the invertibility of B_k is proved in Theorem 7.6 in Dantzig and Thapa 2006). Thus, the unique optimal solution to FLP^1 can be accordingly written as $\boldsymbol{x}^* = [\boldsymbol{x}^*_B; \boldsymbol{x}^*_N]$, where $\boldsymbol{x}^*_B = (\boldsymbol{x}^*_{k,B})_{k=1}^K$, $\boldsymbol{x}^*_N = (\boldsymbol{x}^*_{k,N})_{k=1}^K$. Note that if we define $B = \text{diag}(B_1, \ldots, B_K)$ as a block diagonal matrix with $(B_k)_{k=1}^K$ as its main diagonal blocks and zero matrices as off-diagonal blocks, and $\bar{\boldsymbol{V}} = [\bar{\boldsymbol{V}}_1; \ldots ; \bar{\boldsymbol{V}}_K]$, we can write $\boldsymbol{x}_B^* = B^{-1}\bar{\boldsymbol{V}}$. Let $\boldsymbol{V}_k^t$ be the RHS of FLP_k^t and $\bar{\boldsymbol{V}}_k^t$ be the vector where we delete $C_{0k}^t/(\theta - s)$ from $\boldsymbol{V}_t^k$. Define

$$\delta \boldsymbol{V}_k^t := \left(\left(\sum_{m=1}^M \alpha_m^* \lambda_{jk}(\boldsymbol{q}_m^t) - \sum_{m=1}^M \alpha_m^* \lambda_{jk}(\boldsymbol{q}_m^u) \right)_{j=1}^J , \ \left(- \sum_{s=1}^{t-1} \Delta C_{ik}^s/(\theta - s) \right)_{i=0}^I \right)$$

and let $\delta\bar{\boldsymbol{V}}_k^t$ be the vector where we delete $-\sum_{s=1}^{t-1} \Delta C_{0k}^s/(\theta - s)$ from $\delta\boldsymbol{V}_k^t$. Let $\delta\bar{\boldsymbol{V}}^t = (\delta\bar{\boldsymbol{V}}_k^t)_{k=1}^K$. Following the same decomposition, we will also write $\boldsymbol{c} = [\boldsymbol{c}_B; \boldsymbol{c}_N]$. Per our definition in Section 3, $\boldsymbol{\lambda}^{tot}(\boldsymbol{p})$ is the aggregated purchase probability given a price vector $p \in \Omega_p$. We make our second observation below:

Observation EC.2. *At period t, as long as the following conditions hold:*

$$\sum_{j=1}^J \boldsymbol{\lambda}_j(\boldsymbol{q}_m^t) = \hat{\boldsymbol{\lambda}}_m^t := \boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^u) - \frac{1}{M\alpha_m^*} \left(\sum_{i=0}^I \sum_{s=1}^{t-1} \frac{\Delta \boldsymbol{C}_i^s}{T - s} \right) \in [0,1]^K, \tag{EC.4}$$

$$C_{ik}^t(\theta) = \hat{C}_{ik}^t(\theta) := (\theta - t + 1) \left[C_{ik} - \sum_{s=1}^{t-1} \frac{\Delta C_{ik}^s}{\theta - s} \right] \geq 0, \tag{EC.5}$$

$$\boldsymbol{x}_{k,B}^* + B_k^{-1} \left(\delta\bar{\boldsymbol{V}}_k^t \right) \succeq \boldsymbol{0}, \tag{EC.6}$$

then the unique optimal solution to FLP^t is given by $\boldsymbol{x}_{k,B}^t = \boldsymbol{x}_{k,B}^* + B_k^{-1}\left(\delta\bar{\boldsymbol{V}}_k^t\right)$ and $\boldsymbol{x}_{k,N}^t = \boldsymbol{0}$ for all k.

Proof. Under condition (EC.4), FLP^t is balanced. This is so because, for all k,

$$\sum_{j=1}^J \sum_{m=1}^M \alpha_m^* \lambda_{jk}(\boldsymbol{q}_m^t) = \sum_{j=1}^J \sum_{m=1}^M \alpha_m^* \lambda_{jk}(\boldsymbol{q}_m^u) - \sum_{i=0}^I \sum_{s=1}^t \frac{\Delta C_{ik}^s}{T-s} = \sum_{i=0}^I C_{ik} - \sum_{i=0}^I \sum_{s=1}^{t-1} \frac{\Delta C_{ik}^s}{\theta - s} = \sum_{i=0}^I \frac{C_{ik}^t(\theta)}{\theta - t + 1},$$

where the second equality follows from our assumption in the beginning of this section, and the last equality follows from the definition of ΔC_{ik}^t. As a result, for all k, the inventory constraints in FLP_k^t are all binding. Notice that condition (EC.4) and (EC.5) implies that $\boldsymbol{V}_k^t = \boldsymbol{V}_k + \delta\boldsymbol{V}_k^t \succeq \boldsymbol{0}$, and thus FLP_k^t is equivalent to $\left\{ \min_{\mathbf{x}_k^t} \ \mathbf{c}_k^\top \mathbf{x}_k^t : \bar{A}_k \boldsymbol{x}_k^t = \bar{\boldsymbol{V}}_k + \delta\bar{\boldsymbol{V}}_k^t, \, \mathbf{x}_k^t \succeq \boldsymbol{0} \right\}$. The feasibility of the proposed optimal solution can be directly verified under condition (EC.6); its optimality follows from Karush-Kuhn-Tucker (KKT) conditions; and its uniqueness follows from the invertibility of B_k. □

Step 2

Define $\hat{\mathbf{x}}^t := (\hat{\boldsymbol{x}}_B^t, \boldsymbol{x}_N) = (\boldsymbol{x}_B^* + B^{-1}\left(\delta\bar{\boldsymbol{V}}^t\right), \boldsymbol{0})$. Let $\phi_x = \min_{k\in[K]} \min_{(i,j)\in\mathcal{B}_k} x_{ijk}^* > 0$ (by non-degeneracy assumption); $\Phi_1 = \max_{p\in\Omega_p, \ j\in[J], \ k,\ell\in[K]} |\partial\lambda_{jk}(\boldsymbol{p})/\partial p_\ell| > 0$ (it is finite by Assumption A1); $\Phi_2 = \max_{k\in[K]} ||B_k^{-1}||_\infty > 0$ (it is also finite by the invertibility of B_k); $\phi_\lambda := \max\{x > 0 : \boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^u) + x \cdot \mathbf{1} \in [0,1]^K, \ \forall m\} > 0$ (by Assumption A1 and the fact that $\boldsymbol{q}_m^u$ lies in the interior of Ω_p); and $v > 0$ denote the smallest absolute eigenvalue of $\mathcal{G}_{\boldsymbol{\lambda}^{tot}}$ (by Assumption A3). Without loss of generality, $\boldsymbol{\alpha}^* \succ \boldsymbol{0}$ since we can delete any α_m^* with zero value without changing anything else. We state a lemma.

Lemma EC.1. *Suppose that $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^s) = \hat{\boldsymbol{\lambda}}_m^s \in [0,1]^K$, $\mathbf{x}_s = \hat{\mathbf{x}}_s \succeq \boldsymbol{0}$ and $\boldsymbol{C}_i^s(\theta) = \hat{\boldsymbol{C}}_i^s(\theta) \succeq \boldsymbol{0}$ for all $s < t$. Then $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t) = \hat{\lambda}_m^t$, $\mathbf{x}^t = \hat{\mathbf{x}}^t$ and $\boldsymbol{C}_i^t(\theta) = \hat{\boldsymbol{C}}_i^t(\theta)$ hold if the following two conditions hold at time t*

$$(\dagger): \qquad \left| \sum_{i=1}^I \sum_{s=1}^{t-1} \frac{\Delta C_{ik}^s}{\theta - s} \right| \leq \min\left\{ \frac{\phi_x}{\Phi_2} \left(1 + \frac{K\Phi_1}{v} \right)^{-1}, \ \phi_\lambda M \cdot \min_{m\in[M]} \alpha_m^* \right\}, \ \forall k,$$

$$(\dagger\dagger): \qquad \left| \sum_{s=1}^{t-1} \frac{\Delta C_{ik}^s}{\theta - s} \right| \leq C_{ik}, \ \forall i,k,$$

Proof. We proceed by induction. The base case ($t=1$) is verified directly by definition. Now, consider $t>1$. Assume the identity holds for $s \le t-1$. Given condition (†) and the definition of ϕ_λ, it is not difficult to show that $\hat{\boldsymbol{\lambda}}_m^t \in [0,1]^K$. Since $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t)$ is simply the projection of $\hat{\boldsymbol{\lambda}}_m^t$ onto $[0,1]^K$ (see Step 2a), $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t) = \hat{\boldsymbol{\lambda}}_m^t$.

We now show that $C_{ik}^t(\theta) = \hat{C}_{ik}^t(\theta)$. Suppose that, in Step 2c of R^{2}PF, we sample m^t for some $m^t \in [M]$. Remember that, in period $t-1$, the probability of using FC i to fulfill the request of product k from location j conditioned on $D_{jk}=1$ is $y_{ijk}^{t-1} = x_{ijk}^{t-1} / \sum_{i=0}^I x_{ijk}^{t-1}$. Moreover, since conditions (EC.4) - (EC.6) are implied for all $s \le t$ by the inductive assumption, by Observation EC.2, the inventory constraints in FLP^{t-1} are binding. So, the remaining inventory at the beginning of period t satisfies:

$$\begin{aligned} C_{ik}^t(\theta) &= C_{ik}^{t-1}(\theta) - \sum_{j=1}^J X_{ijk}^{t-1} = C_{ik}^{t-1}(\theta) - \sum_{j=1}^J y_{ijk}^{t-1} \left(\sum_{m=1}^M \alpha_m^* \lambda_{jk}(\boldsymbol{q}_m^{t-1}) \right) - \Delta C_{ik}^{t-1} \\ &= C_{ik}^{t-1}(\theta) - \sum_{j=1}^J x_{ijk}^{t-1} - \Delta C_{ik}^{t-1} = C_{ik}^{t-1}(\theta) - \frac{C_{ik}^{t-1}(\theta)}{\theta - t + 2} - \Delta C_{ik}^{t-1} \\ &= (\theta - t + 2 - 1) \left[C_{ik}(\theta) - \sum_{s=1}^{t-2} \frac{\Delta C_{ik}^s}{\theta - s} \right] - \Delta C_{ik}^{t-1} = \hat{C}_{ik}^t(\theta), \end{aligned}$$

where the second equality follows from the definition of ΔC_{ik}^t; the third equality follows from the fulfillment constraint in FLP^t; the fourth constraint follows since the inventory constraints in FLP^{t-1} are binding; and, the fifth constraints follows from the inductive assumption.

At last, to show that $\boldsymbol{x}^t = \hat{\boldsymbol{x}}^t$, by Observation EC.2, it suffices to show conditions (EC.4) - (EC.6) are satisfied for period t. Condition (EC.4) is implied by $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t) = \hat{\boldsymbol{\lambda}}_m^t$. Since condition (††) implies $\hat{C}_{ik}^t(\theta) \ge 0$, and we have shown that $C_{ik}^t(\theta) = \hat{C}_{ik}^t(\theta)$, condition (EC.5) is satisfied. To check condition (EC.6), define $\delta \boldsymbol{q}_m^t = \boldsymbol{q}_m^t - \boldsymbol{q}_m^u$. By Assumption A1 and Mean Value Theorem, $\delta \boldsymbol{q}_m^t = [\mathcal{G}_{\boldsymbol{\lambda}^{tot}}(\boldsymbol{\xi}_m^t)]^{-1} \left(\sum_{i=0}^I \sum_{s=1}^{t-1} \Delta \boldsymbol{C}_i^s / (\theta - s) \right) / (M \alpha_m^*)$ for some $\boldsymbol{\xi}_m^t \in \Omega_p$. By Mean Value Theorem again, there exist $\boldsymbol{\zeta}_{mk}^t \in \Omega_p$ such that

$$\left| \sum_{m=1}^M \alpha_m^* \left[\lambda_{jk}(\boldsymbol{q}_m^t) - \lambda_{jk}(\boldsymbol{q}_m^u) \right] \right| = \left| \sum_{m=1}^M \frac{(\nabla \lambda_{jk}(\boldsymbol{\zeta}_{mk}^t))^\top [\mathcal{G}_{\boldsymbol{\lambda}^{tot}}(\boldsymbol{\xi}_m^t)]^{-1}}{M} \left(\sum_{i=0}^I \sum_{s=1}^{t-1} \frac{\Delta \boldsymbol{C}_i^s}{\theta - s} \right) \right| \le \frac{K \Phi_1}{v} \max_{k \in [K]} \left| \sum_{i=0}^I \sum_{s=1}^{t-1} \frac{\Delta C_{ik}^s}{\theta - s} \right|$$

where the inequality holds by Assumption A3 and the definition of Φ_1. So,

$$||B_k^{-1} (\delta \bar{\boldsymbol{V}}_k^t)||_\infty \le ||B_k^{-1}|| \cdot ||\delta \bar{\boldsymbol{V}}_k^t|| \le \Phi_2 \cdot \left(1 + \frac{K \Phi_1}{v} \right) \max_{k \in [K]} \left| \sum_{i=0}^I \sum_{s=1}^{t-1} \frac{\Delta C_{ik}^s}{\theta - s} \right| \le \phi_x,$$

where the last inequality follows from condition (†). This implies condition (EC.6). □

Step 3

In this step, we show that the conditions in Lemma EC.1 hold for the majority of the selling season. Define a stopping time $\tau(\theta)$ to be the first t such that either (†) or (††) is violated. According to Lemma EC.1, for any period before $\tau(\theta)$, we can explicitly characterize the evolution of price, fulfillment assignment, and inventory consumption. The following lemma provides a lower bound on the length of $\tau(\theta)$.

Lemma EC.2. *There exists a constant $\Psi_3 > 0$ independent of θ such that $\mathbb{E}[\theta - \tau(\theta)] \le \Psi_3(1 + \log \theta)$.*

Proof. Define $\tau_1(\theta)$ and $\tau_2(\theta)$ to be the first period t such that conditions (†) and (††) are violated, respectively. By definition $\tau(\theta)=\min_{i\in\{1,2\}}\tau_i(\theta)$. We only bound $\tau_1(\theta)$, since $\tau_2(\theta)$ can be bounded using a similar argument. Let Γ_k denote the RHS of the inequality in condition (†) in Lemma EC.1. The sequence

$$\left\{S_k^t=\sum_{i=0}^{I}\frac{\Delta C_{ik}^{t-1}}{\theta-(t-1)}+\sum_{i=0}^{I}\frac{\Delta C_{ik}^{t-2}}{\theta-(t-2)}+\cdots+\sum_{i=0}^{I}\frac{\Delta C_{ik}^{1}}{\theta-1}\right\}_{t\le\theta}$$

is a Martingle with respect to the natural filtration $\{\mathcal{H}^t\}$, where $\mathcal{H}^t$ is the history of all information up to the beginning of period t. This implies that the sequence $\{|S_k^t|\}_{t\le\theta}$ is a sub-Martingle. By Doob's submartingle inequality (see for example Williams 1991) and union bound,

$$\mathbb{P}(\tau_1(\theta)\le t)\;\le\;\mathbb{P}\left(|S_k^s|\ge\Gamma_k\text{ for some }s\le t,k\in[K]\right)\le\sum_{k=1}^{K}\mathbb{P}\left(\max_{s\le t}|S_k^s|\ge\Gamma_k\right)\;\le\;\sum_{k=1}^{K}\frac{\mathbb{E}\left[(S_k^t)^2\right]}{\Gamma_k^2}.$$

Note that ΔC_{ik}^s and ΔC_{jk}^t are independent for all $s\neq t$ and $i,j\in\{0\}\cup I$. So,

$$\mathbb{E}[(S_k^t)^2]=\mathbb{E}\left[\left(\sum_{s=1}^{t-1}\sum_{i=0}^{I}\frac{\Delta C_{ik}^s}{\theta-s}\right)^2\right]=\sum_{s=1}^{t-1}\frac{\mathbb{E}\left[\left(\sum_{i=0}^{I}\Delta C_{ik}^s\right)^2\right]}{(\theta-s)^2}=\sum_{s=1}^{t-1}\frac{\sum_{i,j\in\{0\}\cup[I]}\mathbb{E}\left[\Delta C_{ik}^s\Delta C_{jk}^s\right]}{(\theta-s)^2}\;=\;O\left(\frac{1}{\theta-t}\right),$$

where the last inequality follows from the boundedness of $\mathbb{E}\left[\Delta C_{ik}^s\Delta C_{jk}^s\right]$. The proof is complete by noting that $\mathbb{E}\left[\theta-\tau_1(\theta)\right]\;=\;\sum_{t=2}^{\theta}\mathbb{P}(\tau_1(\theta)\le t)\;=\;1+\sum_{t=2}^{\theta-1}O\left(\frac{1}{\theta-t}\right)\;=\;O(\log\theta)$. □

Step 4

We now bound the loss of R^2PF. First, note that we can decouple the loss into two terms as follows:

$$\begin{aligned}&\mathcal{J}^{ALP}(\theta)-\mathcal{J}^{R^2PF}(\theta)\\&=\mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^*r_j(\boldsymbol{q}_m^*)-\sum_{t=1}^{\theta}\sum_{j=1}^{J}(\boldsymbol{p}^t)^\top\boldsymbol{D}_j^t(\boldsymbol{p}^t)\right]+\mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t-\sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right].\end{aligned}$$

The two terms on the RHS of the equation above are the loss in revenue and the loss in fulfillment cost of R^2PF, respectively. We start with providing an upper bound for the loss in revenue:

$$\begin{aligned}&\mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^*r_j(\boldsymbol{q}_m^*)-\sum_{t=1}^{\theta}\sum_{j=1}^{J}(\boldsymbol{p}^t)^\top\boldsymbol{D}_j^t(\boldsymbol{p}^t)\right]\\
&\le\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^*r_j(\boldsymbol{q}_m^u)-\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}R_j^t(\boldsymbol{p}^t)\right]+\mathbb{E}\left[\sum_{t=\tau(\theta)}^{\theta}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^*r_j(\boldsymbol{q}_m^u)\right]\\
&=\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^*r_j(\boldsymbol{q}_m^u)-\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}R_j^t(\boldsymbol{p}^t)\right]+\mathbb{E}\left[(\theta-\tau(\theta)+1)\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^*r_j(\boldsymbol{q}_m^u)\right]\\
&\le\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^*r_j(\boldsymbol{q}_m^u)-\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}R_j^t(\boldsymbol{p}^t)\right]+Kp_u(1+\Psi_3+\Psi_3\log\theta),\end{aligned}\tag{EC.7}$$

where the last inequality follows from Lemma EC.2, the boundedness of price, and the assumption of at most one arrival per period. Let $\hat{\Delta}_j^t=\sum_{m=1}^{M}\alpha_m^*r_j(\boldsymbol{q}_m^t)-(\boldsymbol{p}^t)^\top D_j(\boldsymbol{p}^t)$. Define $r^{tot}(\boldsymbol{p})=\sum_{j=1}^{J}r_j(\boldsymbol{p})=\boldsymbol{p}^\top\boldsymbol{\lambda}^{tot}(\boldsymbol{p})$. By Assumption A1, there exists an inverse of $\boldsymbol{\lambda}^{tot}(\boldsymbol{p})$, which we will denote as $\boldsymbol{p}(\boldsymbol{\lambda}^{tot}):[0,1]^K\to\Omega_p$. With slight abuse of notation, we will use $r^{tot}(\boldsymbol{\lambda}^{tot})=(\boldsymbol{p}(\boldsymbol{\lambda}^{tot}))^\top\boldsymbol{\lambda}^{tot}$ to denote total revenue rate as a function of aggregate demand. Let $\boldsymbol{\lambda}_m^* = \boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^u)$, $\boldsymbol{\lambda}_m^t = \boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t)$, and $\boldsymbol{\epsilon}^t = \sum_{i=0}^{I}\sum_{s=1}^{t-1}\Delta \boldsymbol{C}_i^s/(\theta - s)$. For $t \le \tau(\theta)$, we know that $\boldsymbol{\lambda}_m^t = \boldsymbol{\lambda}_m^* - \boldsymbol{\epsilon}^t/(M\alpha_m^*)$. By Taylor's expansion at λ_m^*, we have

$$\begin{aligned} r^{tot}(\boldsymbol{q}_m^t) = r^{tot}(\boldsymbol{\lambda}_m^t) &= r^{tot}(\boldsymbol{\lambda}_m^*) - (\nabla r^{tot}(\boldsymbol{\lambda}_m^*))^\top \boldsymbol{\epsilon}_m^t/(M\alpha_m^*) + (\boldsymbol{\epsilon}^t)^\top \nabla^2 r^{tot}(\boldsymbol{\eta}^t)\boldsymbol{\epsilon}^t/(2M^2(\alpha_m^*)^2) \\ &= r^{tot}(\boldsymbol{q}_m^u) - (\nabla r^{tot}(\boldsymbol{\lambda}_{\boldsymbol{m}}^*))^\top \boldsymbol{\epsilon}_m^t/(M\alpha_m^*) + (\boldsymbol{\epsilon}^t)^\top \nabla^2 r^{tot}(\boldsymbol{\eta}^t)\,\boldsymbol{\epsilon}^t/(2M^2(\alpha_m^*)^2) \end{aligned}$$

for some $\boldsymbol{\eta}_m^t \in [0,1]^K \in \Omega_p$. So, the first term in (EC.7) can be bounded as follows:

$$\begin{aligned} &\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^* r_j(\boldsymbol{q}_m^u) - \sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J} R_j^t(\boldsymbol{p}^t)\right] \\ &= \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{m=1}^{M}\alpha_m^* r^{tot}(\boldsymbol{q}_m^u) - \sum_{t=1}^{\tau(\theta)-1}\sum_{m=1}^{M}\alpha_m^* r^{tot}(\boldsymbol{q}_m^t)\right] + \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^* r_j(\boldsymbol{q}_m^t) - \sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}(\boldsymbol{p}^t)^\top D_j(p^t)\right] \\ &\le \mathbb{E}\left[\sum_{t=2}^{\tau(\theta)-1}\sum_{m=1}^{M}\frac{(\nabla r^{tot}(\boldsymbol{\lambda}_m^*))^\top \boldsymbol{\epsilon}^t}{M}\right] - \mathbb{E}\left[\sum_{t=2}^{\tau(\theta)-1}\sum_{m=1}^{M}\frac{(\boldsymbol{\epsilon}^t)^\top \nabla^2 r^{tot}(\boldsymbol{\eta}_m^t)\,\boldsymbol{\epsilon}^t}{2M^2\min_{m\in[M]}\alpha_m^*}\right] + \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)}\sum_{j=1}^{J}\hat{\Delta}_j^t\right] + Kp_u, \qquad \text{(EC.8)} \end{aligned}$$

where the last inequality holds because $\mathbb{E}\,[\hat{\Delta}_j^{\tau(\theta)}] \le Kp_u$. Note that $\{\sum_{s=1}^{t}\sum_{j=1}^{J}\hat{\Delta}_j^s\}_{t\le\theta}$ is a Martingale with respect to $\{\mathcal{H}^t\}_{t\le\theta}$ and $\tau(\theta)$ is bounded. So, by stopping time theorem (Williams 1991), $\mathbb{E}\,[\sum_{t=1}^{\tau(\theta)}\sum_{j=1}^{J}\hat{\Delta}_j^t] = 0$. We are left to bound the first two terms in (EC.8). Note that $\mathbb{E}\,[\sum_{t=2}^{\tau(\theta)-1}\boldsymbol{\epsilon}^t] = \mathbb{E}\,[\sum_{t=2}^{\tau(\theta)}\boldsymbol{\epsilon}^t] - \mathbb{E}\,[\sum_{t=\tau(\theta)}^{\theta}\boldsymbol{\epsilon}^t] = -\mathbb{E}\,[\sum_{t=\tau(\theta)}^{\theta}\boldsymbol{\epsilon}^t]$. By stopping time theorem again, $\mathbb{E}[\boldsymbol{\epsilon}^{\tau(\theta)}] = \mathbf{0}$, and $\mathbb{E}[\boldsymbol{\epsilon}^t] = \mathbf{0}$ for all $t > \tau(\theta)$. Consequently, $\mathbb{E}[\sum_{t=2}^{\tau(\theta)-1}\sum_{m=1}^{M}(\nabla r^{tot}(\boldsymbol{\lambda}_m^*))^\top\boldsymbol{\epsilon}^t] = (\sum_{m=1}^{M}\nabla r^{tot}(\boldsymbol{\lambda}_m^*))^\top\mathbb{E}[\sum_{t=2}^{\tau(\theta)-1}\boldsymbol{\epsilon}^t] = 0$. As for the second term in (EC.8), let $\Phi_3 > 0$ be the largest absolute eigenvalue of $\nabla^2 r^{tot}$. By Assumption A3, Φ_3 is finite. We thus have

$$\begin{aligned} &\mathbb{E}\left[\sum_{t=2}^{\tau(\theta)-1}\sum_{m=1}^{M}(\boldsymbol{\epsilon}^t)^\top\nabla^2 r^{tot}(\boldsymbol{\eta}_m^t)\,\boldsymbol{\epsilon}^t\right] \le \Phi_3\,\mathbb{E}\left[\sum_{t=2}^{\tau(\theta)-1}\sum_{k=1}^{K}\left(\sum_{i=1}^{I}\sum_{s=1}^{t-1}\frac{\Delta C_{ik}^s}{\theta - s}\right)^2\right] \\ &\le \Phi_3\sum_{t=2}^{\theta}\sum_{k=1}^{K}\sum_{1\le s,v\le t-1}\frac{\mathbb{E}\left[\left(\sum_{i=1}^{I}\Delta C_{ik}^s\right)^2\left(\sum_{i=1}^{I}\Delta C_{ik}^v\right)^2\right]}{(\theta-s)(\theta-v)} = \Phi_3\sum_{t=2}^{\theta}\sum_{k=1}^{K}\sum_{s=1}^{t-1}\frac{\mathbb{E}\left[\left(\sum_{i=1}^{I}\Delta C_{ik}^s\right)^2\right]}{(\theta-s)^2} = O(\log\theta). \end{aligned}$$

At last we bound the loss of fulfillment cost. By Lemma EC.1, for $t < \tau(\theta)$, $\mathbf{x}^t = [x_B^* + B^{-1}\delta\bar{\boldsymbol{V}}^t; \mathbf{0}]$. By definition, $\bar{c}$ is larger than all unit shipping costs. So,

$$\begin{aligned} &\mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right] \\ &\le \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right] + \mathbb{E}\left[\sum_{t=\tau(\theta)}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t\right] \\ &\le \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right] + \bar{c}IJK\mathbb{E}[\theta - \tau(\theta) + 1] \\ &\le \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right] + \bar{c}IJK(1 + \Psi_3 + \Psi_3\log\theta). \qquad \text{(EC.9)} \end{aligned}$$

We are left to bound the first term in (EC.9). Let $\Delta x_{ijk}^t = X_{ijk}^t - x_{ijk}^t$. Since $\mathbf{x}^t = \hat{\mathbf{x}}^t$ for all $t < \tau(\theta)$, we have:

$$\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}\left(x_{ijk}^t - x_{ijk}^*\right)\right] + \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ij}\Delta x_{ijk}^t\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1} \boldsymbol{c}_B^\top B^{-1}\left(\delta \bar{\boldsymbol{V}}^t\right)\right] + \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}\Delta x_{ijk}^t\right] \\
&\leq \bar{c}(I+J)K\|B^{-1}\|_1 \left\{\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{m=1}^{M} \alpha_m^*\left(\lambda_{jk}(\boldsymbol{q}_m^t) - \lambda_{jk}(\boldsymbol{q}_m^u)\right) - \sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{k=1}^{K}\sum_{s=1}^{t}\frac{\Delta C_{ik}^s}{\theta - s}\right]\right\} \\
&\quad + \bar{c}\,\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\Delta x_{ijk}^t\right] \\
&= -2\bar{c}(I+J)K\|B^{-1}\|_1 \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{k=1}^{K}\sum_{s=1}^{t}\frac{\Delta C_{ik}^s}{\theta - s}\right] + \bar{c}\,\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\Delta x_{ijk}^t\right]
\end{aligned}
$$

where the second inequality follows from the definition of $\delta\bar{\boldsymbol{V}}^t$, the second equality follows from the definition of $\tau(\theta)$ and Lemma EC.1. Note that $\{\sum_{s=1}^{t}\Delta x_{ijk}^s\}_{t\leq\theta}$ is Martingale with respect to the filtration $\{\mathcal{H}^t\}_{t\leq\theta}$. Following a similar argument as in bounding the revenue loss, it is not difficult to see that the terms after the above equation can be bounded by a constant independent of θ. $\square$

EC.4. Remaining Details of Numerical Experiment

The Poisson process that models the arrival from location j has rate $\gamma_j = \textit{pois-rate} \times \textit{mkt-share}_j$, where *pois-rate* $\in (0,1]$ is the total arrival rate and $\textit{mkt-share}_j$ is the conditional probability that this arrival comes from region j. We set *pois-rate* to be 0.9 and $\textit{mkt-share}_j$ to be the ratio between the total population in the j^{th} largest MSA and the total population of all fifteen MSA. A customer arriving from location j makes a purchase with probability $\exp(A_j + B_j p)$. The parameters of purchasing probabilities are chosen as follows: We first set "baseline" demand parameters A_1 and B_1. For all $j \geq 2$, we then set $A_j = \frac{income_1}{income_j} \times A_1$ and $B_j = \frac{income_1}{income_j} \times B_1$, where $income_j$ represents the medium household income of the j^{th} largest MSA, as reported in U.S. Census Bureau (2014). Since we want $\exp(A_j + B_j p) \leq \mathbf{1}$ for all $p \in \Omega_p$, we set A_j's to be vectors with negative components, and B_j's to be diagonally dominated matrices with negative diagonal components. The baseline parameters shown below are generated to satisfy these constraints. The absolute magnitude of their entries depends on the price range, which, in our setting, depends on the shipping cost.

$$
A_1 = \begin{bmatrix} -1.0071 \\ -1.2603 \\ -1.3228 \\ -1.5005 \\ -1.4810 \end{bmatrix}, \quad B_1 = \begin{bmatrix} -9.5 & 1.1 & 1.1 & 1.1 & 1.7 \\ 1.9 & -10.5 & 2.0 & 1.4 & 1.0 \\ 1.1 & 1.6 & -11.5 & 2.0 & 1.9 \\ 2.0 & 2.0 & 1.5 & -12.6 & 2.0 \\ 1.6 & 2.0 & 1.8 & 2.0 & -12.1 \end{bmatrix} \times 10^{-3}
$$

The transportation cost is calculated using the cost equation estimated in Section EC.3 in Jasin and Sinha (2015) assuming that each product weighs exactly one pound. To be precise, $c_{ijk} = \bar{c}_k \cdot (9.182 + 0.000541 \cdot d_{ij})$, where d_{ij} is the distance in miles from FC i to demand location j, and $\bar{c}_k$ is uniformly distributed in $[0.9, 1.1]$. We set the inventory levels of the FCs to minimize the likelihood that we use FCs that are far away from the demand location even under a myopic fulfillment policy; this is to prevent the separate optimization heuristic from performing too bad. To do so, we first match between FCs and MSAs such that (1) each FC serves five MSA, (2) each MSA is served by 2 FCs, and (3) the total mileage between all the assigned

FC-MSA pairs is minimized. We then approximate an average purchase quantity from MSA j by $\hat{\lambda}_j = \textit{pois-rate} \times \textit{mkt-share}_j \times 0.9$. (The factor 0.9 means that the initial inventory levels are set to be slightly below the expected total arrivals; this reflects the common reality where firm stocks neither too low such that the induced demand has to be really scarce, nor too high as if there is no inventory constraint at all.) Each of the two FCs serving MSA_j fulfill a portion of the $\hat{\lambda}_j$, where the portion is decided by a random number drawn uniformly from $[0.4, 0.6]$. (Our results are robust with respect to perturbation in the numbers 0.9, 0.4, and 0.6.) The total initial inventory at each of the FC is then calculated as the sum of all the demand portions from the five MSAs it serves. At last, we distribute the initial inventory at each of the FC uniformly across all of the products. As a result, the initial inventory level is $C_{1k} = 0.0337$, $C_{2k} = 0.0218$, $C_{3k} = 0.0217$, $C_{4k} = 0.0276$ $C_{5k} = 0.0196$ and $C_{6k} = 0.0196$ for all $k = 1, \ldots, 5$. The fictitious FC is set to hold abundant initial inventories so that they will never be depleted. For a specific θ, we always round down θC_{ik}.

Table EC.1 reports the expected profits of all the heuristics implemented in Section 7. The coefficient of variations are consistently small (less than 0.5% for all instances); due to the space constraint, we will not report them in the paper.

θ	RPFC-2	RPF-5	RPF-8	R^2PF-2	R^2PF-5	R^2PF-8
200	9430.3	8654.1	8452.7	11188.2	11174.6	10944.4
400	21637.1	21301.9	20844.2	24584.6	24704.3	25130.5
600	34556.3	34326.4	34442.5	38487.0	38675.3	39717.1
800	47128.5	47354.4	47555.6	51096.7	52239.4	53663.8
1000	59421.7	60013.1	60280.1	63915.5	66210.0	66685.4
1300	78355.8	80807.3	80988.6	83891.0	86832.0	87803.4
1600	97707.0	100438.0	100663.0	104095.7	107811.5	109225.9
2000	123868.1	127255.7	128223.2	130583.3	136083.0	137391.0

θ	R^2PF-Ful-2	R^2PF-Ful-5	R^2PF-Ful-8	R^2PF-Pr-2	R^2PF-Pr-5	R^2PF-Pr-8
200	10685.9	10243.5	10293.8	9573.5	9625.1	9095.4
400	23772.8	23913.0	24415.4	22762.9	23170.0	21773.3
600	37361.6	38190.4	38684.6	35861.0	36447.4	35845.4
800	50420.3	51661.2	52378.8	49233.1	50155.6	49914.0
1000	63231.5	64542.2	65582.1	61764.5	63859.5	63264.4
1300	83296.8	85417.4	86609.7	81244.5	83893.7	84359.5
1600	103189.5	106475.0	108053.1	101556.2	104595.7	105571.4
2000	129244.4	134312.8	135841.8	127404.6	131924.8	132286.8

θ	ALP-Reopt-2	ALP-Reopt-5	Sep-reopt	DJPF-Reopt-1	DJPF-Reopt-10
200	11082.3	10875.5	6029.6	8409.3	11402.7
400	24509.0	25219.7	14857.1	20535.8	23762.8
600	38445.3	39409.5	24084.7	32610.1	35771.6
800	51269.1	53119.2	32596.9	44457.0	49576.2
1000	64106.5	66763.0	40625.2	57977.1	61993.9
1300	84184.8	88029.6	53726.0	77196.4	81677.3
1600	104209.7	108747.1	66782.8	96125.3	101167.5
2000	131023.8	136999.7	84005.0	122814.6	128769.5

Table EC.1 Expected profits of different heuristic with varying θ

Table EC.2 reports the running time of a single simulation for several different heuristics when $\theta = 2000$. The computation time for the last two heuristics is very long, therefore it is not feasible to implement them in practice. All simulations were implemented on a desktop computer with 3.40GHz Intel Core i7-3770 CPU and 8 GB of RAM.

R^2PFC-2	R^2PF-5	R^2PF-8	ALP-Reopt-2	ALP-Reopt-5	ALP-Reopt-8	DJPF-Reopt-θ
14.98	23.12	23.69	26.98	992.87	10814.08	58376.24

Table EC.2 Typical running time (in seconds) for a single simulation for selected heuristics

References

Dantzig, GB, MN Thapa. 2006. *Linear programming 1: Introduction*. Springer.

Jasin, S, A Sinha. 2015. An LP-based correlated rounding scheme for multi-item ecommerce order fulfillment. *Oper. Res.* **63**(6) 1336–1351.

Schrijver, A. 1998. *Theory of linear and integer programming*. John Wiley & Sons.

U.S. Census Bureau. 2014. Income and poverty in the United States: 2014. `http://bit.ly/2a35pzL`.

Williams, D. 1991. *Probability with martingales*. Cambridge university press.