Appendix: Proof of Lemma 1

In this section we provide the proof of the lemma that the linear program objective function converges to the cost of the perfect hindsight solution as demand and supply are scaled up.

Recall the integer optimization problem:

$$\begin{aligned} v_{PH}(D',G') \equiv \min_x \quad & \sum_{i,j} c_{ij}x_{ij} \\ s.t. \quad & \sum_i x_{ij} \geq D_j' \quad \forall\ j \\ & \sum_j x_{ij} \leq X_i \quad \forall\ i \\ & x_{ij} \leq G_{ij}' \quad \forall i,j \\ & x_{ij} \in \mathbb{Z}_{\geq 0} \quad \forall\ i,j \end{aligned} \tag{1}$$

and the linear programing formulation:

$$\begin{aligned} v_{LP} \equiv \min_x \quad & \sum_{i,j} c_{ij}x_{ij} \\ s.t. \quad & \sum_i x_{ij} \geq \mu_j \quad \forall\ j \\ & \sum_j x_{ij} \leq X_i \quad \forall\ i \\ & x_{ij} \leq \mu_j \rho_i \quad \forall i,j \\ & x_{ij} \geq 0 \quad \forall\ i,j \end{aligned} \tag{2}$$

The random demand process $D^{(\theta)}$ was defined in such a way that if θ is a scaling parameter, then:

$$\begin{aligned} \mu^{(\theta)} &= \mu \cdot \theta \\ \sigma_j^{D,(\theta)} &= \sigma_j^D \cdot \sqrt{\theta} \\ \sigma_j^{G,(\theta)} &= \sigma_j^G \cdot \sqrt{\theta} \end{aligned} \tag{3}$$

Lemma: *When inventory and demand are scaled up, the transportation linear program objective value approaches that of the expected value of the perfect hindsight optimization in ratio. Or* $\lim_{\theta \to \infty} E[v_{PH}^{(\theta)}] / v_{LP}^{(\theta)} = 1$.

Proof of Lemma:

Recall that for given realizations *D'* and *G'* of random processes *D* and *G* respectively, we defined $v_{PH}(D',G')$ as the objective value of an integer optimization problem outlined above in formulation (1). If we relax the integrality constraint from formulation (1), we can define $v_{PH,relax}(D',G')$

as the objective function of the resulting formulation. Because this relaxed integer optimization problem can be formulated as a network flow problem, then $v_{PH}(D',G') = v_{PH,relax}(D',G')$.

Using Wald's equation and the assumptions from section 7 of the main paper, we can redefine the expected value of G.

$$
\begin{aligned}
G_{ij} &\equiv \sum_{k \in K_j : D_k^{\{0,1\}}=1} \widetilde{Z}_{ik} \\
E[G_{ij}] &= E\left[\sum_{k \in K_j : D_k^{\{0,1\}}=1} \widetilde{Z}_{ik} \right] \\
&= E\left[D_j\right] \cdot E\left[\widetilde{Z}_{ik}\right] \\
&= E\left[D_j\right] \cdot P(\text{FC } i \text{ has other items in stock for customer } k) \\
&= \mu_j \cdot P(\text{FC } i \text{ has other items in stock for a random customer}) \\
&= \mu_j \cdot \rho_i
\end{aligned} \tag{4}
$$

Therefore, if we substitute the expected value of the random variables for the specific realizations, we obtain the following formulation:

$$
\begin{aligned}
v_{PH,relax}(\mu,\mu\rho) \quad \equiv \quad \min_x \quad & \sum_{i,j} c_{ij} x_{ij} \\
s.t. \quad & \sum_i x_{ij} \ge \mu_j \quad \forall\ j \\
& \sum_j x_{ij} \le X_i \quad \forall\ i \\
& x_{ij} \le \mu_j \rho_i \quad \forall\, i,j \\
& x_{ij} \ge 0 \quad \forall\ i,j
\end{aligned} \tag{5}
$$

Recall, though, that this is the same formulation of the linear program outlined in (2) from above. Thus, $v_{LP} = v_{PH,relax}(\mu,\mu\rho)$. We also note that minimization linear optimization problems are piecewise linear convex in the right hand side. Therefore, by Jensen's inequality,

$$
v_{LP} = v_{PH,relax}(\mu,\mu\rho) \le E_{D,G}[v_{PH,relax}(D,G)] \tag{6}
$$

This inequality is useful for obtaining the lower bound on the ratio, and will be utilized again towards the end of the proof.

We now turn our attention to obtaining the upper bound of the ratio needed to prove the lemma. The value $v_{PH,relax}(D',G')$ that was obtained by relaxing the integrality constraint of formulation (1) can alternatively be formulated as such:

$$
\begin{aligned}
v_{PH,relax}(D',G') \quad = \quad & \min_{x} \quad \sum_{i,j} c_{ij} x_{ij} \\
& s.t. \quad \sum_{i} x_{ij} \geq \mu_j + \left(D_j^{'} - \mu_j\right) \quad \forall\ j \\
& \qquad \sum_{j} x_{ij} \leq X_i \qquad \forall\ i \\
& \qquad x_{ij} \leq \mu_j \rho_i - \left(\mu_j \rho_i - G_{ij}^{'}\right) \quad \forall i, j \\
& \qquad x_{ij} \geq 0 \qquad \forall\ i, j
\end{aligned}
\tag{7}
$$

The following formulation includes expensive backward arcs $\overleftarrow{x}_{ji}$. Flow on these arcs has cost c_N. There are no backward arcs leading from demand nodes to the dummy node N. The formulation below has the equivalent objective value as the one above, with all flow $\overleftarrow{x}_{ji}$ on these backward arcs equal to zero. It can be shown that if some backward arc did have flow, an alternate formulation exists with a lower overall cost and no flow on that backward arc. Thus, adding the backward arcs leads to a formulation with an equal objective value. Additionally, though, the $()^+$ notation is added to the first and third set of constraints. Due to this change, the polyhedron defined by the constraints can only become smaller, which leads to an objective value that is not smaller than the previous one. The polyhedron is smaller because in the first set of constraints (which includes a greater-than-or-equal sign), the right hand side becomes larger if $D_j^{'} < \mu_j$, and in the third set of constraints (which has a less-than-or-equal sign), the right hand side becomes smaller if $\mu_j \rho_i < G_{ij}^{'}$. Both of these adjustments on the constraints tighten the formulation.

$$v_{PH,relax}(D',G') \leq \begin{array}{lll} \min\limits_{x} & \sum\limits_{i,j} c_{ij}x_{ij} + \sum\limits_{i,j} c_N \overleftarrow{x}_{ji} & \\ s.t. & \sum\limits_{i} x_{ij} - \sum\limits_{i} \overleftarrow{x}_{ji} \geq \mu_j + \left(D_j^{'} - \mu_j\right)^{+} & \forall\ j \\ & \sum\limits_{j} x_{ij} - \sum\limits_{j} \overleftarrow{x}_{ji} \leq X_i & \forall\ i \\ & x_{ij} - \overleftarrow{x}_{ji} \leq \mu_j\rho_i - \left(\mu_j\rho_i - G_{ij}^{'}\right)^{+} & \forall i,j \\ & x_{ij}, \overleftarrow{x}_{ji} \geq 0 & \forall\ i,j \end{array} \tag{8}$$

The following formulation has an additional constraint, which, again, cannot reduce the objective value.

$$\leq \begin{array}{lll} \min\limits_{x} & \sum\limits_{i,j} c_{ij}x_{ij} + \sum\limits_{i,j} c_N \overleftarrow{x}_{ji} & \\ s.t. & \sum\limits_{i} x_{ij} - \sum\limits_{i} \overleftarrow{x}_{ji} \geq \mu_j + \left(D_j^{'} - \mu_j\right)^{+} & \forall\ j \\ & \sum\limits_{j} x_{ij} - \sum\limits_{j} \overleftarrow{x}_{ji} \leq X_i & \forall\ i \\ & x_{ij} - \overleftarrow{x}_{ji} \leq \mu_j\rho_i - \left(\mu_j\rho_i - G_{ij}^{'}\right)^{+} & \forall i,j \\ & \overleftarrow{x}_{ji} = \left(\mu_j\rho_i - G_{ij}^{'}\right)^{+} & \\ & x_{ij}, \overleftarrow{x}_{ji} \geq 0 & \forall\ i,j \end{array} \tag{9}$$

Using algebra and substituting $\left(\mu_j\rho_i - G_{ij}^{'}\right)^{+}$ for $\overleftarrow{x}_{ji}$:

$$= \begin{array}{lll} \min\limits_{x} & \sum\limits_{i,j} c_{ij}x_{ij} + \sum\limits_{i,j} c_N \left(\mu_j\rho_i - G_{ij}^{'}\right)^{+} & \\ s.t. & \sum\limits_{i} x_{ij} - \sum\limits_{i} \left(\mu_j\rho_i - G_{ij}^{'}\right)^{+} \geq \mu_j + \left(D_j^{'} - \mu_j\right)^{+} & \forall\ j \\ & \sum\limits_{j} x_{ij} - \sum\limits_{j} \left(\mu_j\rho_i - G_{ij}^{'}\right)^{+} \leq X_i & \forall\ i \\ & x_{ij} \leq \mu_j\rho_i & \forall i,j \\ & \overleftarrow{x}_{ji} = \left(\mu_j\rho_i - G_{ij}^{'}\right)^{+} & \\ & x_{ij}, \overleftarrow{x}_{ji} \geq 0 & \forall\ i,j \end{array} \tag{10}$$

Rearranging terms, and removing the $\overleftarrow{x}_{ji}$'s from the formulation yields:

$$
= \quad \begin{aligned}
\min_{x} \quad & \sum_{i,j} c_{ij} x_{ij} + \sum_{i,j} c_N \left(\mu_j \rho_i - G_{ij}^{'} \right)^+ \\
s.t. \quad & \sum_i x_{ij} \geq \mu_j + \left(D_j^{'} - \mu_j \right)^+ + \sum_i \left(\mu_j \rho_i - G_{ij}^{'} \right)^+ \quad && \forall\ j \\
& \sum_j x_{ij} \leq X_i + \sum_j \left(\mu_j \rho_i - G_{ij}^{'} \right)^+ && \forall\ i \\
& x_{ij} \leq \mu_j \rho_i && \forall\, i, j \\
& x_{ij} \geq 0 && \forall\ i, j
\end{aligned} \tag{11}
$$

The polyhedron from the following formulation is not larger than the previous one because the right hand side of the second set of constraints is reduced. This again leads to an objective value that is not smaller than the previous one:

$$
\leq \quad \begin{aligned}
\min_{x} \quad & \sum_{i,j} c_{ij} x_{ij} + \sum_{i,j} c_N \left(\mu_j \rho_i - G_{ij}^{'} \right)^+ \\
s.t. \quad & \sum_i x_{ij} \geq \mu_j + \left(D_j^{'} - \mu_j \right)^+ + \sum_i \left(\mu_j \rho_i - G_{ij}^{'} \right)^+ \quad && \forall\ j \\
& \sum_j x_{ij} \leq X_i && \forall\ i \\
& x_{ij} \leq \mu_j \rho_i && \forall\, i, j \\
& x_{ij} \geq 0 && \forall\ i, j
\end{aligned} \tag{12}
$$

The following formulation explicitly separates out the flow from dummy supply node *N'* to each demand node. This has no effect on the objective value. Recall though that dummy supply node *N* is still included in the indexing of *i*.

$$
= \quad \begin{aligned}
\min_{x} \quad & \sum_{i \neq N^{'}, j} c_{ij} x_{ij} + \sum_{j} c_N x_{N^{'},j} + \sum_{i,j} c_{N^{'}} \left(\mu_j \rho_i - G_{ij}^{'} \right)^+ \\
s.t. \quad & \sum_{i \neq N'} x_{ij} + x_{N^{'},j} \geq \mu_j + \left(D_j^{'} - \mu_j \right)^+ + \sum_i \left(\mu_j \rho_i - G_{ij}^{'} \right)^+ \quad && \forall\ j \\
& \sum_j x_{ij} \leq X_i && \forall\ i \\
& x_{ij} \leq \mu_j \rho_i && \forall\, i, j \\
& x_{ij} \geq 0 && \forall\ i, j
\end{aligned} \tag{13}
$$

We then add a set of equality constraints which forces a certain value of flow along the arcs from dummy fulfillment center node *N'* to each demand node. Feasibility is maintained because there are no such restrictions on the arcs from the dummy node *N* to each demand node. (In fact, this is the reason we created two dummy fulfillment centers rather than one.) These equality constraints yield a possibly smaller polyhedron with an objective value not smaller than the previous one:

$$
\leq \quad \begin{aligned}
\min_x \quad & \sum_{i \neq N', j} c_{ij} x_{ij} + \sum_j c_N x_{N',j} + \sum_{i,j} c_N \left(\mu_j \rho_i - G'_{ij} \right)^+ \\
s.t. \quad & \sum_{i \neq N'} x_{ij} + x_{N',j} \geq \mu_j + \left(D'_j - \mu_j \right)^+ + \sum_i \left(\mu_j \rho_i - G'_{ij} \right)^+ \quad \forall\ j \\
& x_{N',j} = \left(D'_j - \mu_j \right)^+ + \sum_i \left(\mu_j \rho_i - G'_{ij} \right)^+ \quad \forall\ j \\
& \sum_j x_{ij} \leq X_i \quad \forall\ i \\
& x_{ij} \leq \mu_j \rho_i \quad \forall i, j \\
& x_{ij} \geq 0 \quad \forall\ i, j
\end{aligned} \tag{14}
$$

Substituting in for $x_{N',j}$ and adjusting the indexing so that *N'* is now *not* included in the indexing of *i* yields the following, which we term the perfect hindsight, relaxed, penalized solution:

$$
\leq \quad v_{PH,relaxed,penalized}\left(D',G'\right) \equiv \begin{aligned}
\min_x \quad & \sum_{i,j} c_{ij} x_{ij} + \sum_{i,j} c_N \left(\mu_j \rho_i - G'_{ij} \right)^+ \\
& + \sum_j c_N \left\{ \left(D'_j - \mu_j \right)^+ + \sum_i \left(\mu_j \rho_i - G'_{ij} \right)^+ \right\} \\
s.t. \quad & \sum_i x_{ij} \geq \mu_j \quad \forall\ j \\
& \sum_j x_{ij} \leq X_i \quad \forall\ i \\
& x_{ij} \leq \mu_j \rho_i \quad \forall i, j \\
& x_{ij} \geq 0 \quad \forall\ i, j
\end{aligned} \tag{15}
$$

The cost of this policy is at least as expensive as that of the relaxed prefect hindsight policy due to the above logic: $v_{PH,relaxed}\left(D',G'\right) \leq v_{PH,relaxed,penalized}\left(D',G'\right)$. The penalized and relaxed perfect hindsight formulation can be decomposed because two terms in the objective function are constants:

$$
v_{PH,relaxed,penalized}\left(D',G'\right) = v_{LP} + 2\sum_{i,j} c_N \left(\mu_j \rho_i - G'_{ij} \right)^+ + \sum_j c_N \left(D'_j - \mu_j \right)^+ \tag{16}
$$

Due to a result from the inventory management literature (Gallego 1992):

$$
E[(Y - E[Y])^+] \leq \frac{\sigma_Y}{2} \tag{17}
$$

Therefore, we can write the relations:

$$
\begin{aligned}
v_{LP} &\leq E[v_{PH,relax}(D,G)] && \text{by (6)}\\
&\leq E[v_{PH,relax,penalized}(D,G)] && \text{by (7)-(15)}\\
&= v_{LP} + E\left[2\sum_{i,j} c_N\left(\mu_j\rho_i - G_{ij}\right)^+ + \sum_j c_N\left(D_j - \mu_j\right)^+\right] && \text{by (16)}\\
&= v_{LP} + c_N \cdot E\left[2\sum_{i,j}\left(\mu_j\rho_i - G_{ij}\right)^+ + \sum_j\left(D_j - \mu_j\right)^+\right] && \text{bring } c_N \text{ out of expectation}\\
&\leq v_{LP} + c_N \cdot \sum_{i,j}\sigma_j^G + c_N \cdot \sum_j \frac{\sigma_j^D}{2} && \text{by (17)}\\
&= v_{LP} + c_N\,|I|\sum_j \sigma_j^G + c_N \sum_j \frac{\sigma_j^D}{2} && \text{by removing summation over } i\\
&\leq v_{LP} + c_N \cdot |I| \cdot |J| \cdot \sigma^{G,MAX} + \frac{c_N \cdot |J| \cdot \sigma^{D,MAX}}{2}
\end{aligned}
$$

where in the last inequality, $\sigma^{G,MAX} \equiv \max_j \ \sigma_j^G$. Therefore, because of the above logic and equation

Error! Reference source not found. which states $v_{PH}(D',G') = v_{PH,relax}(D',G')$:

$$
v_{LP} \leq E[v_{PH}(D,G)] \leq v_{LP} + c_N \cdot |J| \cdot \left(|I| \cdot \sigma^{G,MAX} + \frac{\sigma^{D,MAX}}{2}\right) \tag{18}
$$

If we scale up inventory and demand, we can show that the LP transportation problem approaches the value of the expectation of the perfect hindsight policy's cost in ratio.

Let θ be a scaling parameter, and let $D^{(\theta)}$ be the random demand process such that $\mu^{(\theta)} = \mu \cdot \theta$, $\sigma_j^{D,(\theta)} = \sigma_j^D \cdot \sqrt{\theta}$, and $\sigma_j^{G,(\theta)} = \sigma_j^G \cdot \sqrt{\theta}$. This can be achieved, for instance, if scaled demand in a given region can be viewed as the sum of independent random variables, which might be the case if demand from day to day were independent and if we viewed the process over a longer time horizon. Define $X^{(\theta)}$ similarly, i.e., $X_j^{(\theta)} \equiv \theta(X_j)$ *for all* j. Also, let the objective values $v^{(\theta)}$ be defined as using the adjusted demand processes and inventory positions. For notational convenience, let us write $E[v_{PH}^{(\theta)}]$ for $E[v_{PH}^{(\theta)}(D,G)]$. Thus:

$$
v_{LP}^{(\theta)} \leq E[v_{PH}^{(\theta)}] \leq v_{LP}^{(\theta)} + c_N \cdot |J| \cdot \left(|I| \cdot \sigma^{G,(\theta),MAX} + \frac{\sigma^{D,(\theta),MAX}}{2}\right) \tag{19}
$$

and, dividing the three terms by $v^{(\theta)}_{LP}$:

$$\frac{v_{LP}^{(\theta)}}{v_{LP}^{(\theta)}} \le \frac{E[v_{PH}^{(\theta)}]}{v_{LP}^{(\theta)}} \le \frac{v_{LP}^{(\theta)}}{v_{LP}^{(\theta)}} + \frac{c_N \cdot |J| \left(|I| \cdot \sigma^{G,(\theta),MAX} + \frac{\sigma^{D,(\theta),MAX}}{2} \right)}{v_{LP}^{(\theta)}}$$
$$1 \le \frac{E[v_{PH}^{(\theta)}]}{v_{LP}^{(\theta)}} \le 1 + \frac{c_N \cdot |J| \left(|I| \cdot \sigma^{G,(\theta),MAX} + \frac{\sigma^{D,(\theta),MAX}}{2} \right)}{v_{LP}^{(\theta)}} \tag{20}$$

We first note that the effect of scaling constraints by the same factor for a linear program increases the objective value by the same factor: $w_{LP}^{(\theta)} = \theta \cdot w_{LP}$. By our definition of the random demand process for both *G* and *D*: $\sigma^{(\theta),MAX} = \sqrt{\theta} \cdot \sigma^{MAX}$. Therefore:

$$\lim_{\theta \to \infty} \frac{c_N \cdot |J| \left(|I| \cdot \sigma^{G,(\theta),MAX} + \frac{\sigma^{D,(\theta),MAX}}{2} \right)}{w_{LP}^{(\theta)}} = \lim_{\theta \to \infty} \frac{c_N \cdot |J| \left(|I| \cdot \sigma^{G,MAX} + \frac{\sigma^{D,MAX}}{2} \right)}{w_{LP}} \frac{\sqrt{\theta}}{\theta} = 0 \tag{21}$$

because the denominator grows as θ, while the numerator grows only as the square root of θ. Therefore,

$$\lim_{\theta \to \infty} \frac{E[v_{PH}^{(\theta)}]}{v_{LP}^{(\theta)}} = 1 \tag{22}$$

□

Thus, we have shown that in the limiting case, the objective function of the linear program approaches that of a perfect hindsight optimization for a single period model, justifying its use as an estimate of cost-to-go.

Bibligography

Gallego, G. 1992. A minmax distribution free procedure for the (Q, R) inventory model. *Operations Research Letters* 11 (1):55–60.