Global Dual Sourcing and Order Smoothing

Robert N. Boute Jan A. Van Mieghem
robert.boute@vlerick.com *vanmieghem@northwestern.edu*

Oct 19, 2011

Abstract

Motivated by the recent trend of offshoring manufacturing facilities to seek lower production cost, while keeping some costly local facilities to better respond to changes in market demand, this paper studies a dual sourcing policy and characterizes the average sourcing allocation. We consider the total landed cost from origin (combining a cheap but slow source and a fast but expensive source) to destination (finished goods warehouse), including sourcing, capacity and inventory related costs. We adopt a linear control rule that allows an exact and analytically-tractable analysis of a dual sourcing or mixed-mode transportation policy that is easy to implement. Another distinguishing feature of our model is that it captures the leadtime difference between both sources. We provide an exact lower bound formula on the strategic base or offshore allocation when this leadtime difference equals one review period, and a close approximation for longer leadtime differences. This formula provides structural insight on the impact of financial, operational and demand parameters, and a starting point for data-driven decision making. We demonstrate the robustness of our results by showing that this dual sourcing smoothing policy is near optimal in terms of total cost minimization compared to existing single sourcing and dual sourcing models in the literature. We describe under which practical conditions the dual sourcing smoothing policy is advisable and when it can even outperform the other policies.

Keywords: inventory, capacity, dual sourcing, production smoothing, mixed-mode transportation.

1 Introduction

After a decade of rapid globalization, there is increased interest to bring offshored production closer to home. A key driver behind global sourcing has been its ability to increase margins by using cheap labor and materials. However, those benefits do not come for free: offshoring suffers from higher transaction costs, long and complicated logistics that are sensitive to increased volatilities in demand, supply, currency exchange rates and oil prices, and political and environmental criticism (Maher and Tita 2010). To hedge against these risks, offshoring requires substantial safety inventory, the holding cost of which can outweigh the labor and materials cost advantage. In addition, responding to customers' new-product requests, shorter delivery times, and swift corrections to improve designs and quality has magnified the need for responsive and agile supply chains. Local sourcing therefore is receiving increasing attention.

A complete reversal to local sourcing, however, is unlikely and ill-advised. Indeed, the concepts of global and local sourcing are not mutually exclusive. Rather, the combined use of multiple supply sources, each of which is different and possesses unique advantages, might be better than any single sourcing strategy. Admittedly, multi-sourcing involves higher coordination costs, but a strategically configured portfolio of suppliers with complementary skills can often perform better than any individual supplier.

In this paper we analyze a global dual sourcing policy for companies that have access to two suppliers with complementary competencies: a local supplier that is responsive but more expensive, and a global supplier that is (globally) most cost-efficient but with a longer leadtime. (Supplier competencies can correspond to transportation modes so that our analysis also applies to balancing mixed-mode transportation.) While our policy can be used for two local suppliers, it is more likely that the low-cost supplier is offshore and far away from the local responsive supplier. The main contribution of this paper is to provide the first exact and analytically-tractable analysis of a dual sourcing policy that is easy to implement. This policy allows us to design an ordering policy that allocates the order volume to both sources so as to optimally trade-off cost and responsiveness.

We present a tight approximation for the optimal volume fraction ordered from the slow but cheap source, which we refer to as the strategic base or offshoring allocation, and its corresponding total landed cost. The strategic allocation is characterized by a simple formula that provides structural insight on the impact of financial, operational and demand parameters, and a starting point for data-driven decision making:

$$\text{Strategic base or offshoring allocation} = \left(1 - \left(\frac{\kappa_I}{\frac{L\Delta}{\text{COV}} + \kappa^l + \frac{1}{2}\kappa^g}\right)^{2/3}\right)^{L/2} \tag{1}$$

where Δ and L denote the unit-cost and leadtime differences between both sources, COV the coefficient of variation in demand, and κ_I, κ^l and κ^g measure the cost of resp. inventory, local and global capacity (which will be defined precisely later in the paper on pages 11 and 13). Formula (1) shows how the strategic allocation increases in the global source's financial cost advantage Δ but decreases in the demand variability COV. Since offshoring smooths the order volumes, the global order allocation increases with higher capacity costs κ^l or κ^g but decreases when inventory costs κ_I are important. Finally, the formula quantifies how a longer leadtime difference L decreases the order allocation to the global source. We prove that this approximation is also a tight lower bound when the leadtime difference between both locations equals one period.

The remainder of this section is organized as follows. We first present a precise mathematical formulation of our dual sourcing problem in section 1.1. In section 1.2 we review the literature related to our problem definition, and in section 1.3 we position our contribution relative to the existing literature.

1.1 Model

Consider a periodic-review inventory system that can be replenished from two sources. Time is discrete and the sequence of events at each time $t = 0, 1, ..., T$ is as follows: First the demand D_t is observed and satisfied (unfilled demand is backordered). Demand is stationary and i.i.d. with $\mathbb{E}(D_t) = \mu$, $Var(D_t) = \sigma^2$, and distribution Φ. Let ϕ_N and Φ_N denote the standard normal density and distribution and $I_N(z) = \phi_N(z) - z(1 - \Phi_N(z))$ the unit normal loss function. Then, the net inventory I_t, which is the inventory on hand minus backorders, is observed and replenishment orders are placed. The analysis is simplified by letting q_t^i denote the order quantity *received* in period t from source i with $i \in \{l, g\}$ for resp. the local and global source, and let $q_t = q_t^l + q_t^g$ denote the total order quantity received in period t. The order from the local source is received in time to fill next period's demand; this is equivalent to saying that the local replenishment is received by the end of the period in which its order is placed. In contrast, the global order faces a delay of L periods so that the quantity q_t^g that is received at time t from the global source must be ordered in period $t - L$ (and thus depend only on quantities observed up to $t - L$). Following Zipkin (2000, p. 404), we say that the local source has a risk period or total leadtime of 1 period while the global source has leadtime $L + 1 \geq 2$. (This risk period includes the one period review. One may alternatively say the local source has physical leadtime 0 and the global L. The essence is that the sources have a leadtime difference of $L \geq 1$.)

Given this sequence of events where we first satisfy demand, then observe inventory and finally place and receive orders, we have the following inventory dynamics for $t = 1, ..., T$:

$$I_t = I_{t-1} + q_{t-1} - D_t. \tag{2}$$

The dynamics for the first period are $I_0 = I_{-1} - D_0$. The initial inventory I_{-1}, which is a constant, can be decomposed into $I_{-1} = \mu + I_s$, where I_s denotes the safety-stock.

Let the total cost per period be the total landed cost incurred from origin (one of either source) to destination (finished goods warehouse). Orders to source i incur sourcing cost c^i per unit, with the faster source being more expensive: $c^l > c^g$. This sourcing cost reflects the standard variable cost component in the total cost for units coming from location i. This component includes direct material as well as any labor cost that can directly be attributed to the order size.

The capacity investment cost at location i reflects the standard fixed cost component in the total cost for units coming from location i. The cost of capacity K^i includes capital, labor, and other overhead costs that remain unchanged over in the medium-long term (i.e., during the time horizon $[0, T]$). Orders q_t^i can be produced up to the capacity K^i that is installed at time 0. In the natural regime, local capacity is more expensive than global capacity, although the model works without those conditions. The installed capacity incurs a cost per period of $C(K^i) = k^i K^i$, with k^i the constant, marginal investment cost rate to add one unit of capacity at source i. Any excess order $\left(q^i - K^i\right)^+$ requires overtime capacity at extra cost o^i per unit. (Obviously, $k^i < o^i$, otherwise it would never be optimal to invest in capacity.) Overtime reflects excess cost not covered in regular capacity costs nor standard direct labor. We could additionally include a fixed cost component, which is independent of the size of the capacity over a reasonable range (e.g., real estate and administrative overhead). The inclusion of any fixed cost will impact whether single or dual sourcing is optimal (and simply requires comparing both costs) but not the interior solution.

As such, our cost model is sufficiently flexible to capture location-differences in sourcing, capacity and overtime costs. We assume sourcing and capacity costs are incurred at receipt (although it doesn't make a difference in our undiscounted model). Finally, each period, inventory or backlog incur a holding cost h per unit on hand or a backlog cost b per unit short. The average cost over horizon T becomes:

$$C_{I_{-1}}^{T} = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=l,g}\left[k^{i}K^{i} + o^{i}\left(q_{t}^{i} - K^{i}\right)^{+} + h\left(I_{t}\right)^{+} + b\left(I_{t}\right)^{-} + c^{i}q_{t}^{i}\right].$$

When the inventory and order process is ergodic (as it will be), the (random) average cost converges w.p. 1 to

$$\lim_{T\to\infty} C_{I_{-1}}^{T} = C = \sum_{i=l,g}\left[k^{i}K^{i} + o^{i}\mathbb{E}\left(q^{i} - K^{i}\right)^{+} + h\mathbb{E}\left(I\right)^{+} + b\mathbb{E}\left(I\right)^{-} + c^{i}\mathbb{E}\left(q^{i}\right)\right]. \tag{3}$$

Ideally, one would like to characterize the initial inventory I_{-1}, a capacity vector K and an admissible dual sourcing policy (which defines the replenishments q_t^l and q_{t+L}^g as a function of I_{t-k} and D_{t-k} for $t, k = 0, 1, ...$) that minimizes C. The optimal policy to this problem is to our knowledge unknown. It has been shown that if you disregard capacity costs, the dual-index base-stock policy is optimal when the leadtime difference is one period, but the optimal policy becomes intractable if $L > 1$. In this paper we will focus on optimizing a specific family of sourcing policies that are sufficiently simple to analyze and adopt in practice. Our specific attention goes to smoothing policies because they can be analyzed exactly and generate strategic insight in the average sourcing allocation. We introduce our class of dual sourcing smoothing policies in section 2.

1.2 Related literature

Our work directly relates to two streams of research: dual sourcing inventory models and order smoothing policies. The dual-sourcing literature refers to inventory models where replenishment occurs through a regular channel and/or a more expensive, but faster expedited channel. The objective is to minimize the expected sum of procurement, holding and shortage costs over multiple periods. The dual-sourcing literature is very rich; we focus primarily on discrete review models. Fukuda (1964) shows that when the leadtime difference is one period, dual-base-stock policies are optimal. In a dual-base-stock policy, an expedite order is placed to bring the inventory position up to a first *(expedite)* base-stock level, after which a regular order is placed to bring it up to a second and higher *(regular)* base-stock level. Fukuda uses first-order conditions to derive expressions for the base-stock levels. Whittemore and Saunders (1977) extend Fukuda's (1964) model and show that when leadtimes differ by more than one period, the optimal policy is no longer a dual base-stock, but it depends on the entire ordering history and requires multidimensional dynamic programming.

Globally optimal policies are in general highly complex. Therefore, various heuristic policies are proposed in the literature. Rosenshine and Obee (1976) consider a standing order policy, which Allon and Van Mieghem (2010) call a tailored base-surge (TBS) policy. TBS orders at a constant rate from the regular source ("base"demand) and uses a base-stock policy for the emergency replenishment (the remaining "surge"demand). Tagaras and Vlachos (2001) extend this policy to allow emergency replenishment within the regular review period. Even this simple TBS policy is not amenable to exact analysis – instead Allon and Van Mieghem (2010) develop a Brownian analytical model that is asymptotically optimal for high sourcing volumes. It is noteworthy that, by definition, a TBS policy is independent of the slow source's leadtime.

Veeraraghavan and Scheller-Wolf (2008) introduce a dual-index dual-base-stock policy that tracks inventory positions over both regular and expedited leadtimes. Order-up-to levels for both inventory positions are computed using a simulation-based optimization procedure. The authors show that such a dual-index policy is nearly optimal when compared to state-dependent policies. Sheopuri et al. (2010) generalize the class of dual-index policies. They show that the "lost sales inventory problem"is a special case of the dual sourcing problem and use this property to suggest two new classes of policies with an order-up-to structure that perform equal to, or even slightly better than, the dual-index policy with the same computational requirements. Scheller-Wolf et al. (2006) consider single-index dual-base-stock policies, which are identical to dual-index policies except that only one inventory position is tracked instead of two. The authors computationally show that its performance is comparable to the more complex dual-index heuristic.

Recently, the dual-sourcing literature is adopted in the context of global sourcing strategies, thus combining the advantages of global low-cost sourcing and local quick response manufacturing. Allon and Van Mieghem (2010) introduce the standing order policy in this perspective and provide guidelines for determining the "strategic allocations,"i.e. how the average total sourcing volume should be allocated to the global and local sources. Wu and Zhang (2011) develop a game-theoretic model where multiple firms in a competitive setting may choose between efficient sourcing and responsive sourcing; a key feature of the game is that depending on the sourcing strategy, a firm may observe different signals about the uncertain market demand. Liu and Nagurney (2011) address the impact of demand and cost uncertainty in a supply chain network with offshoring and quick-response production. Using variational inequality theory, the authors formulate the governing equilibrium conditions of the competing manufacturers and a simulation study investigates the quantitative impact of demand and cost uncertainty. Recent empirical work is done by Jain et al. (2011) who study the impact of global sourcing and supplier diversification on inventory investment based on US firm level performance.

In this paper we allocate the order volume to the global and local sources by introducing a class of order smoothing policies. Smoothing is a well known method to reduce variability. The benefit of order smoothing derives from the fact that the order pattern is less variable than the demand. Therefore, the total installed safety capacity is reduced compared to demand-replacing chase policies such as traditional base-stock policies. The introduction of order smoothing in a global dual sourcing context is not new. The TBS policy can actually be interpreted as an order smoothing policy: the presumption is that the low-cost source cannot rapidly change volumes because of frictions such as long leadtimes or an inflexible level production process which is essential to achieve this cost advantage (Allon and Van Mieghem 2010). Indeed, under TBS the global source needs no safety capacity. Moreover, an increase in the standing order reduces the variability of the responsive order stream ("peak-shaving behavior") and thus the required safety capacity of the responsive sources also reduces. Veeraraghavan and Scheller-Wolf (2008) specify a capacitated scenario in their dual-index policy to constrain the order volatility and thus the capacities at each source.

Smoothing is justified when production and holding costs are convex or when there is a cost of changing the level of production (Sobel 1969, 1971). Simon (1952) and Vassian (1955) did pioneering work on the development of smoothing rules using servomechanism (or control) theory and Laplace transform methods. Forrester (1961) and Magee (1958) suggest that production smoothing can be achieved by distributing the transient part of the required production over a number of successive periods. Bertrand (1986) extends this approach to a multi-product multi-phase production system.

Graves has contributed to the smoothing literature over the course of the last 25 years. In 1988, he reviewed the existing literature on safety stocks for manufacturing systems and criticized that the literature did not properly consider the role of safety stocks in the presence of inflexibility in manufacturing systems. He characterized the need for additional safety stocks as a result of the smoothing or decoupling function within a manufacturing operation. Of particular interest to our work is the linear production control rule described in his paper which smoothes the aggregate production and permits an explicit examination of the tradeoff between safety stocks and production flexibility (Graves 1988). A similar rule is used by Balakrishnan et al. (2004) who set the order quantity equal to a convex combination of the previously observed consumer demands. They make use of these order smoothing rules downstream in the chain to coordinate the entire supply chain. They also characterize the optimal smoothing parameter values and assess the potential cost savings that these order-smoothing strategies can yield compared to the uncoordinated case when individual firms separately minimize their costs. In our paper we use the same linear control rules to allocate orders to the global and local source, thereby smoothing production over both sources.

Recent empirical work on production smoothing is done by Cachon et al. (2007) who found, based on industry-level US data, that order smoothing exists in the retail industry and in some manufacturing industries, but not in the wholesale industry. Chen and Lee (2011) show how the prevalence of capacity constraints in these industries (e.g., limited shelf/warehouse space and manufacturing capacity) drives order smoothing. Cantor and Katok (2011) use a series of laboratory experiments to demonstrate Cachon's (2007) findings: when the cost of varying orders is higher than the cost of holding inventory, production and order smoothing is indeed a rational and cost-minimizing behavior. Bray and Mendelson (2011) study firm-level US data and show that firms generally amplify last-minute shocks, yet smooth seasonal variations. There is also a large economics literature preceding the work in operations management, which empirically investigates production smoothing – we refer to Cachon et al. (2007) for an overview and discussion.

Our dual sourcing model also relates to the choice of mixed-mode transportation systems where a shipper can use two transport modes together for a single commodity flow. Recently, Combes (2011) studied this problem by minimizing the total landed cost using approximations and simulations. Our exact analysis and formula for strategic allocation may be directly applicable to this setting.

1.3 Our positioning in the literature

The dual-sourcing literature has taken two approaches: One stream focuses on numerically optimizing sophisticated dynamic policies that approach optimal performance (Veeraraghavan and Scheller-Wolf 2008, Sheopuri et al. 2010). In a quest to generate strategic insight in the average sourcing allocation, the other stream focuses on near-optimal policies that are simpler and therefore easier to adopt in practice and to analyze, e.g. the "standing order policy"(Rosenshine and Obee 1976, Allon and Van Mieghem 2010). Our policy is positioned within this second stream of allocation policies.

The underlying assumption of the standing order policy is that implementing feedback control of the global source is ineffective or not desirable due to various frictions such as long transportation leadtimes or inflexible production. But nothing necessitates the order to the global source to be constant over time. Indeed, the order to the global source *can* fluctuate over time, as long as its order quantity is known at least L periods in advance of the order to the local source (for delivery in the same period). We therefore propose a different allocation policy which does take this condition into account, but allows some flexibility in the offshore order stream. The local source is still utilized to dynamically react to recent changes in demand.

Our dual sourcing policy is the first dual sourcing policy that is easy to implement *and* analytically tractable. It captures most of the relevant parameters, including the leadtime, by exact analysis and it provides a lower bound to the strategic allocation volumes. In addition it is capable to give guidance on practice. Finally, we are able to provide simple formulae and parameter sensitivity that includes the shift from dual to single sourcing as the leadtime increases – we knew that intuitively but are not aware of any analytic or structural results in a dual sourcing model.

2 Dual Sourcing – Order Smoothing Policies

We start with the simpler case where the leadtime difference $L = 1$ and later extend to general L.

2.1 Order and Inventory Dynamics

The proposed "dual sourcing smoothing" policy allocates the sourcing volume q_t to each source as follows: For $t = 0, 1, ...T$:

$$q_t = q_t^g + q_t^l \text{ where } q_t^g = \alpha q_{t-1} \text{ and } q_t^l = (1-\alpha)D_t, \tag{4}$$

with $0 \leq \alpha \leq 1$ and q_{-1} as initial condition. If $\alpha = 1$, then $q_t = q_t^g = q_{t-1} = q_{t-1}^g = q_{-1}$ and we obtain a level strategy that sole sources a constant quantity each period from the offshore source. In contrast, if $\alpha = 0$, then $q_t = q_t^l = D_t$ is a demand replacement policy or a chase strategy that sole sources from the responsive source. Any value in between is a compromise between a chase and level strategy.

This order policy finds its origin in linear control theory (Forrester 1961, Magee 1958). Graves (1988) used this control rule to smooth the aggregate production. The policy also relates to the exponential smoothing order policy, which is used by Balakrishnan et al. (2004) to reduce the variability in orders in a single source setting. Boute et al. (2007) show that this order smoothing policy is identical to a generalized base-stock policy, where the inventory deficit is not recovered in one period, but instead spread out over time: $q_t = (1-\alpha)(s - I_t)$, where $1/(1-\alpha)$ refers to the *adjustment time* and s the base-stock level.

This sourcing policy has multiple benefits: Given that the quantity q_t^g is based on information in period $t-1$, it can be ordered to the global supplier one period in advance for timely delivery in period t. The local source supplies the remaining portion of the delivery in period t, which is dependent on the most recent demand observed in period t. The linear control makes the policy analytically tractable and allows even closed form solutions. This analysis will also show that, by leveraging advance information, the order streams have smaller variability than the demand, which is referred to a smoothing.

Analytic tractability stems from the linear control structure that allows iterative solutions:

Proposition 1 *The order and net inventory processes are linear combinations of the demand process: for $t = 0, 1, ...T$:*

$$q_t = \alpha^{t+1} q_{-1} + \sum_{i=0}^{t} (1-\alpha)\alpha^i D_{t-i} \tag{5}$$

$$I_t = \begin{cases} I_{-1} + \frac{1-\alpha^t}{1-\alpha}\alpha q_{-1} - \sum_{i=0}^{t} \alpha^i D_{t-i} & \text{if } \alpha < 1, \\ I_{-1} + t q_{-1} - \sum_{i=0}^{t} D_i & \text{if } \alpha = 1. \end{cases} \tag{6}$$

Proof:

Iterating the recursion (4) directly yields (5). Given this order process, we can also analytically track the net inventory dynamics for $t = 1, ..., T$ and $\alpha \in [0, 1)$:

$$\begin{aligned} I_t &= I_{t-1} + q_{t-1} - D_t \\ &= I_{-1} + \sum_{i=0}^{t-1} q_i - \sum_{i=0}^{t} D_i \\ &= I_{-1} + \sum_{i=0}^{t-1} \alpha^{i+1} q_{-1} + \sum_{i=0}^{t-1}\sum_{j=0}^{i} (1-\alpha)\alpha^j D_{i-j} - \sum_{i=0}^{t} D_i \\ &= I_{-1} + \frac{1-\alpha^t}{1-\alpha}\alpha q_{-1} + \sum_{i=0}^{t-1}\sum_{j=0}^{i} (1-\alpha)\alpha^j D_{i-j} - \sum_{i=0}^{t} D_i \end{aligned}$$

and using $k = i - j$, the double sum becomes

$$\sum_{i=0}^{t-1}\sum_{j=0}^{i}(1-\alpha)\alpha^j D_{i-j} = \sum_{k=0}^{t-1}\sum_{i=k}^{t-1}(1-\alpha)\alpha^{i-k} D_k = \sum_{k=0}^{t-1}(1-\alpha)\frac{1-\alpha^{t-k}}{1-\alpha} D_k = \sum_{k=0}^{t-1}\left(1-\alpha^{t-k}\right) D_k$$

so that

$$I_t = I_{-1} + \frac{1-\alpha^t}{1-\alpha}\alpha q_{-1} - \sum_{i=0}^{t}\alpha^{t-i} D_i = I_{-1} + \frac{1-\alpha^t}{1-\alpha}\alpha q_{-1} - \sum_{j=0}^{t}\alpha^j D_{t-j}.$$

If $\alpha = 1$, we get $q_t = q_{-1}$ and

$$I_t = I_{-1} + \sum_{i=0}^{t-1} q_i - \sum_{i=0}^{t} D_i = I_{-1} + t q_{-1} - \sum_{i=0}^{t} D_i.$$

■

Taking expectations in (5) and (6) yields for $t = 0, 1, ...T$:

$$\begin{aligned} \mathbb{E}q_t &= \mu + \alpha^{t+1}\left(q_{-1} - \mu\right) \\ \mathbb{E}I_t &= \begin{cases} I_s + \frac{1-\alpha^t}{1-\alpha}\alpha\left(q_{-1} - \mu\right) & \text{if } \alpha < 1, \\ I_s + t\left(q_{-1} - \mu\right) & \text{if } \alpha = 1. \end{cases} \end{aligned}$$

For the average inventory to converge for $\alpha = 1$, we must have $q_{-1} = \mu$ so that the average total order replaces average demand, as is necessary for a stable system. When $\alpha < 1$, however, this is accomplished irrespective the value of q_{-1} (whose effect on the total order washes out because of the exponential smoothing as $t \to \infty$). It nevertheless has a lasting impact on the average inventory. By definition, $\mathbb{E}I_t$, average inventory after demand is fulfilled but before replenishments, is the safety stock. Given that I_s is a decision variable, nothing is lost by henceforth setting $q_{-1} = \mu$ also for $\alpha < 1$.

Proposition 2 *If $q_{-1} = \mu$, the expectation and variance of the order and inventory processes are:*

$$\begin{array}{ll} \mathbb{E}q_t = \mu & Var(q_t) = \left[1 - \alpha^{2(t+1)}\right]\frac{1-\alpha}{1+\alpha}\sigma^2 \le \sigma^2 \\ \mathbb{E}q_t^g = \alpha\mu & Var(q_t^g) = \left[1 - \alpha^{2(t+1)}\right]\frac{1-\alpha}{1+\alpha}\alpha^2\sigma^2 \le \sigma^2 \\ \mathbb{E}q_t^l = \left(1-\alpha\right)\mu & Var(q_t^l) = \left(1-\alpha\right)^2\sigma^2 \le \sigma^2 \\ \mathbb{E}I_t = I_s & Var(I_t) = \begin{cases} \frac{1-\alpha^{2(t+1)}}{1-\alpha^2}\sigma^2 \ge \sigma^2 & \textit{if } \alpha < 1 \\ t^2\sigma^2 & \textit{if } \alpha = 1. \end{cases} \end{array}$$

Proof: Given that D_{t-i} are identically distributed, taking expectations and variances in (5) directly yields $\mathbb{E}q_t$ and $Var(q_t)$. The expectations and variances of q_t^g and q_t^l then follow directly from (4). The expectation of (6) yields $\mathbb{E}I_t = I_{-1} - \mu = I_s$.■

Given that $\mathbb{E}q_t^g = \alpha\mu$, the smoothing parameter α also equals the strategic base or offshoring allocation. (Later we will see that this no longer holds when $L > 1$.) The variance of each order stream is less than the demand variance (hence the name "order smoothing"). The total order variance is convex decreasing in α (i.e., a higher reliance on the global source). The order variances are increasing in t and a positive reliance on the local source ($\alpha < 1$) is necessary for the variances to be converging. The variance of the inventory process grows without bound under sole global sourcing with smoothing ($\alpha = 1$). (The global source then supplies a constant quantity μ while demand remains random with same mean μ. The resulting net inventory process behaves as a random walk with traffic intensitity 1 and is unstable.)

2.2 Cost Components

With the order and inventory dynamics known, we now turn our attention to the three components of the cost structure: sourcing, capacity and inventory costs. Proposition 2 directly yields that the:

$$\text{average sourcing cost rate } C_s = \left[c^l - (c^l - c^g)\alpha\right]\mu. \tag{7}$$

Given that $c^l > c^g$, the sourcing cost C_s decreases in the smoothing parameter α. Consequently, this sourcing cost decrease is the first marginal benefit of offshoring/smoothing

$$MB_s = (c^l - c^g)\mu. \tag{8}$$

Each source needs a capacity level that supplies its order stream using an optimal combination of installed capacity and overtime. This optimal trade-off is found using a standard newsvendor solution. For general distributions and/or finite horizons, this solution involves a critical fractile on a convolution of the demand distribution. When demands are normally distributed, all orders are normally distributed. In the long run, then, the optimal capacities are:

Proposition 3 *With normal demands, the long-run optimal capacities are*

$$\begin{array}{ll} K^{l*} = (1-\alpha)\left(\mu + z_K^l\sigma\right) & \textit{where} \quad \Phi_N(z_K^l) = \frac{o^l - k^l}{o^l}, \\ K^{g*} = \alpha\left(\mu + z_K^g\sqrt{\frac{1-\alpha}{1+\alpha}}\sigma\right) & \textit{where} \quad \Phi_N(z_K^g) = \frac{o^g - k^g}{o^g}. \end{array} \tag{9}$$

The associated capacity cost rate is

$$\begin{aligned} C_K &= k^l(1-\alpha)\mu + \left[k^l z_K^l + o^l I_N(z_K^l)\right](1-\alpha)\sigma \\ &+ k^g\alpha\mu + \left[k^g z_K^g + o^g I_N(z_K^g)\right]\alpha\sqrt{\frac{1-\alpha}{1+\alpha}}\sigma. \end{aligned} \tag{10}$$

Given that the overall order variability is lower than the demand variability, the total installed capacity size $K^l + K^g$ under "smoothed dual sourcing" is less than under single sourcing. To compare the capacity cost as a function of the smoothing degree α, let us denote

$$\kappa^l = k^l z_K^l + o^l I_N(z_K^l), \quad (11)$$
$$\kappa^g = k^g z_K^g + o^g I_N(z_K^g), \quad (12)$$
$$\Delta = (c^l - c^g) + (k^l - k^g). \quad (13)$$

Δ denotes the typical unit-cost advantage of the offshore global source, both in direct sourcing costs as well as in capacity and labor costs. κ^i represents the expected safety capacity and overtime cost per period at location i for a unit standard deviation in demand. The sourcing and capacity cost then becomes:

$$C_s(\alpha) + C_K(\alpha) = C_0 - \left(\Delta\mu + \kappa^l\sigma\right)\alpha + \kappa^g\sigma\alpha\sqrt{\frac{1-\alpha}{1+\alpha}}, \quad (14)$$

where

$$C_0 = (c^l + k^l)\mu + \kappa^l\sigma. \quad (15)$$

Proposition 4 *The capacity cost $C_K(\alpha)$ is concave and decreasing over $\alpha \in [0,1]$ if $k^g \leq k^l$ and $\kappa^g \leq \kappa^l$. Then both sourcing and capacity costs are decreasing, representing the marginal benefit*

$$MB(\alpha) = -C_s' - C_K' = \Delta\mu + \kappa^l\sigma - \kappa^g\sigma\frac{1-\alpha-\alpha^2}{(1-\alpha)^{1/2}(1+\alpha)^{3/2}}, \quad (16)$$

which is convex increasing from $MB(0) = \Delta\mu + (\kappa^l - \kappa^g)\sigma > 0$ to $MB(1) = +\infty$.

Proof: The cost $C_K(\alpha)$ is the positive sum of a linear function and $f(\alpha) = \alpha\sqrt{\frac{1-\alpha}{1+\alpha}}$, where

$$f'(\alpha) = \frac{1-\alpha-\alpha^2}{(1-\alpha)^{1/2}(1+\alpha)^{3/2}},$$
$$f''(\alpha) = \frac{-2(1-\alpha)(1+\alpha)(1+2\alpha) - (1+5\alpha)}{2(1-\alpha)^{3/2}(1+\alpha)^{5/2}}.$$

f'' is negative in $\alpha \in [0,1]$ so that f' is decreasing with $f'(0) = 1$ and minimal value $f'(-1) = -\infty$. Thus, f is concave and so is C_K. Given that $C_K' \leq C_K'(0) = \left(k^g - k^l\right)\mu + \left(\kappa^g - \kappa^l\right)\sigma < 0$, C_K is decreasing everywhere if the right-hand-side is negative.■

The marginal benefit from more offshoring stems from (i) the sourcing cost decrease (8) and (ii) reduced capacity costs due to lower average investment capacity costs $\left(k^l - k^g\right)\mu$ and safety capacity costs. The fact that the latter disappears completely but only in the limit case of level sourcing ($\alpha = 1$) yields (theoretically) the infinite marginal benefit $MB(1)$.

This marginal benefit, however, must be weighted against the marginal cost of increased inventory: With normal demand, the net inventory process I_t is also normally distributed. If $\alpha < 1$, then also the limiting net inventory process I_∞ is normally distributed with mean I_s and variance $\frac{1}{1-\alpha^2}\sigma^2$. The optimal safety stock level then follows from a newsvendor solution:

Proposition 5 *With normal demand and $\alpha < 1$, the long-run optimal safety stock is*

$$I_s^* = z_I \frac{\sigma}{\sqrt{1-\alpha^2}} \text{ where } \Phi_N(z_I) = \zeta_s = \frac{b}{b+h}. \tag{17}$$

The associated inventory holding and backlogging cost rate is

$$C_I = [hz_I + (h+b)I_N(z_I)] \frac{\sigma}{\sqrt{1-\alpha^2}} = \kappa_I \frac{\sigma}{\sqrt{1-\alpha^2}}, \tag{18}$$

which is convex increasing in α, representing the marginal cost

$$MC(\alpha) = C_I'(\alpha) = \kappa_I \sigma \frac{\alpha}{(1-\alpha^2)^{3/2}}. \tag{19}$$

Proof: (*This is not a standard newsvendor problem because the decision variable I_s is the mean of the distribution but we shall show that it reduces to a newsvendor model*). The limiting net inventory process I_∞ is normally distributed with mean $\mathbb{E}I_t = I_s$ and standard deviation $\sigma_I = \frac{\sigma}{\sqrt{1-\alpha^2}}$. Let $f(x|I_s, \sigma_I)$ denote its density. The associated expected inventory cost rate is

$$C_I(I_s) = \int_0^\infty hxf(x|I_s,\sigma_I)dx - \int_{-\infty}^0 bxf(x|I_s,\sigma_I)dx.$$

Substitute to standardized units $z = (x - I_s)/\sigma_I$ and define $z_I = I_s/\sigma_I$. With $x = I_s + z\sigma_I = (z_I + z)\sigma_I$ and $f(x|I_s, \sigma_I) = \phi(z)/\sigma_I$, we get

$$\begin{aligned} C_I(I_s) &= \int_{-z_I}^\infty h(z_I+z)\sigma_I\phi(z)dz - \int_{-\infty}^{-z_I} b(z_I+z)\sigma_I\phi(z)dz \\ &= \sigma_I\left[hz_I(1-\Phi(-z_I)) + h\phi(-z_I) - bz_I\Phi(-z_I) + b\phi(-z_I)\right], \\ &= \sigma_I\left[hz_I + (h+b)\left(\phi(-z_I) - z_I\Phi(-z_I)\right)\right] \end{aligned}$$

where we used the identity $\phi'(x) = -x\phi(x) \Rightarrow -\int_{-\infty}^x z\phi(z)dz = \int_x^\infty z\phi(z)dz = \phi(x)$. Invoking $\phi(-x) = \phi(x)$ and $\Phi(-x) = 1 - \Phi(x)$, it directly follows that

$$\begin{aligned} C_I(I_s) &= \sigma_I\left[hz_I + (h+b)\left(\phi(z_I) - z_I(1-\Phi(z_I))\right)\right] = \sigma_I\left[hz_I + (h+b)I_N(z_I)\right] = \frac{\kappa_I\sigma}{\sqrt{1-\alpha^2}} \\ &= \sigma_I\left[-bz_I + (h+b)\left(\phi(z_I) + z_I\Phi(z_I)\right)\right] \end{aligned}$$

The inventory cost is clearly convex increasing in α :

$$\frac{d}{d\alpha}C_I(I_s) = \kappa_I\sigma\frac{\alpha}{(1-\alpha^2)^{3/2}} > 0, \text{ and } \frac{d^2}{d\alpha^2}C_I(I_s) = \kappa_I\sigma\frac{5\alpha}{(1-\alpha^2)^{5/2}} > 0.$$

Now taking the first derivative

$$\frac{d}{dI_s}C_I(I_s) = \frac{d}{\sigma_I dz_I}C_I(I_s) = -b + (h+b)\left(\phi'(z_I) + \Phi(z_I) + z_I\phi(z_I)\right) = -b + (h+b)\Phi(z_I).$$

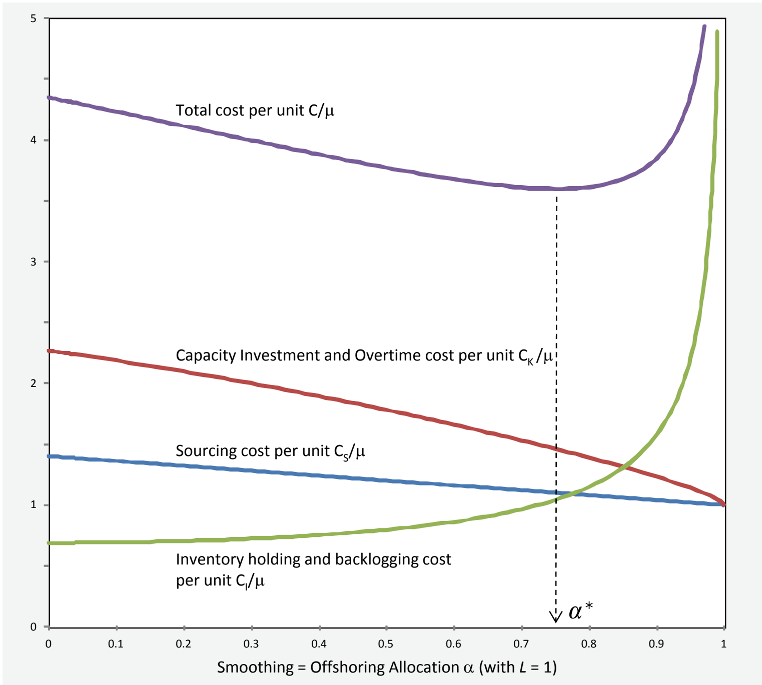

Figure 1: A larger reliance on the global source decreases sourcing and capacity costs but increases inventory cost. There is a unique optimal smoothing or offshoring allocation α^*. (Parameters: $\mu = 50$, $\sigma = 20$, $c^g = 1$, $c^l = 1.4$, $k^l = 1.5$, $k^g = 1$, $o^i = 4k^i$, $h = 0.8$, $b = 19h$.).

Clearly $\frac{d^2}{dI_s^2} C_I(I_s) = \sigma_I^{-1}(h+b)\phi(z_I) > 0$ so that C_I is convex with unique minimum given by the familiar critical fractile condition $\Phi(z_I) = b/(h+b)$.■

The trade-offs are now clear: A larger reliance on the global source has the benefit of decreasing sourcing and capacity cost (in concave fashion) but increases inventory costs (in convex fashion), as shown in Fig. 1. This directly raises the question whether there is an optimal trade-off and, if so, how to characterize it.

If we define the inventory financial parameter

$$\kappa_I = hz_I + (h+b)I_N(z_I) \tag{20}$$

as the expected inventory related cost rate per period for a unit standard deviation in demand, and add it to the financial parameters κ^l, κ^g and Δ defined in (11)-(13), we arrive at the total cost as

$$C(\alpha) = C_0 - \left(\Delta\mu + \kappa^l\sigma\right)\alpha + \kappa^g\sigma\alpha\sqrt{\frac{1-\alpha}{1+\alpha}} + \kappa_I\frac{\sigma}{\sqrt{1-\alpha^2}}. \tag{21}$$

It is useful to introduce four dimensionless parameters: the coefficient of variation measures the relative amount of demand variability and is denoted by

$$COV = \frac{\sigma}{\mu}, \tag{22}$$

and three model-specific parameters θ:

$$\theta_1 = \frac{\Delta}{\kappa_I COV}, \theta_2 = \frac{\kappa^g}{\kappa_I}, \theta_3 = \frac{\kappa^l}{\kappa_I}. \tag{23}$$

Proposition 6 *The total scaled cost rate has two independent parameters $\theta_{13} = \theta_1 + \theta_3$ and θ_2:*

$$\widehat{C}(\alpha; \theta_1, \theta_2) = \frac{C(\alpha) - C_0}{\kappa_I \sigma} = -\theta_{13}\alpha + \theta_2 \alpha \sqrt{\frac{1-\alpha}{1+\alpha}} + \frac{1}{\sqrt{1-\alpha^2}} \tag{24}$$

and is concave-convex in $\alpha \in [0, 1]$ with a unique interior minimum α^* satisfying:

$$\widehat{MB}(\alpha^*) = \theta_{13} - \theta_2 \frac{1 - \alpha^* - \alpha^{*2}}{(1-\alpha^*)^{1/2}(1+\alpha^*)^{3/2}} = \widehat{MC}(\alpha^*) = \frac{\alpha^*}{(1-\alpha^{*2})^{3/2}}. \tag{25}$$

Proof: Both marginal cost and marginal benefits are convex increasing, with MB initially dominating MC and then reversing. Thus, they always have a unique interior intersection, represented by point B in Fig. 2. ■

Proposition 7 *The optimal offshoring allocation α^* increases when θ_{13} increases and, if $\alpha^* > \frac{\sqrt{5}-1}{2}$, when θ_2 increases.*

Proof: The implicit function theorem yields that $\frac{d\alpha^*}{d\theta} = -\frac{\frac{\partial^2 C}{\partial\theta\partial\alpha}}{\frac{\partial^2 C}{\partial\alpha^2}}$, evaluated at α^*. Given that then $\frac{\partial^2 C}{\partial \alpha^2} > 0$, the sign of $\frac{d\alpha^*}{d\theta}$ equals the sign of the numerator which equals $\frac{\partial \widehat{MB}}{\partial \theta}$. Clearly, $\frac{\partial \widehat{MB}}{\partial \theta_{13}} = 1$ so that $\frac{d\alpha^*}{d\theta_{13}} > 0$ while $sign\frac{\partial \widehat{MB}}{\partial \theta_2} = -sign(1 - \alpha^* - \alpha^{*2})$, which is positive when $\alpha > \frac{\sqrt{5}-1}{2}$ and negative otherwise.■

Interestingly, $\frac{\sqrt{5}-1}{2}$ is the well-known golden ratio which represents here an inflection point in capacity costs. The influence of θ_{13} on α^* is clear as it impacts the intercept of $\widehat{MB}$ in Fig. 2. In contrast, θ_2 has opposite impact on the intercept and slope of $\widehat{MB}$. Which effect dominates, and thus the impact of θ_2 on α^*, depends on the location of α^* with respect to the golden ratio.

By investigating how a model parameter impacts θ, we find its impact on the optimal offshoring allocation α^*:

1. If Δ increases, then θ_{13} and thus α^* increase.

2. If $COV = \sigma/\mu$ increases while $\Delta > 0$ (which is typical), then θ_{13} and thus α^* decrease.

3. If h increases, then κ_I increases[1] and if $\alpha^* > \frac{\sqrt{5}-1}{2}$, then both θ_{13} and θ_2 and thus α^* decrease.

[1] $\frac{d\kappa_I}{dh} = \frac{d}{dh}[hz_I + (h+b)I_N(z_I)] = \phi_N(z_I) + \Phi_N(z_I)z_I > 0$ for all z_I as optimal functions of h.

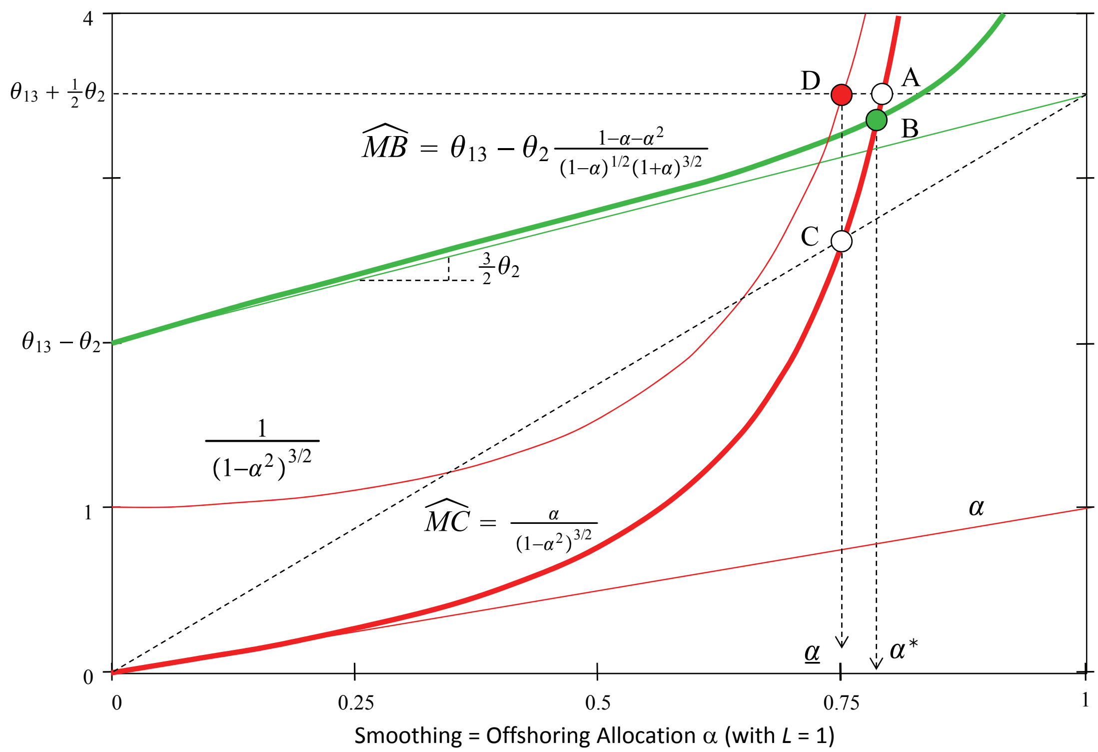

Figure 2: The marginal cost of inventory (solid red) intersects the marginal benefit of sourcing and capacity (solid green) at point B, defining the optimal smoothing level α^*. Bounds on MB and MC intersect at point D, defining the lower bound $\underline{\alpha}$ on α^*.

4. If k^l increases while $z^l_K > 0$ (which is typical), then κ^l and Δ increase[2] and thus θ_{13} and α^* increase.

5. If k^g increases while $z^g_K > 0$ (which is typical), then κ^g and θ_2 increase while Δ and thus θ_{13} decrease. The overall effect depends on their relative magnitudes but typically the latter effect dominates and α^* decreases.

These comparative statics give guidance on how to tailor and adapt the sourcing strategy to the financial, customer service requirements, and demand characteristics. One should source more from the local source (less from the global source) if

- the customer service requirements (and thus inventory costs) are higher,
- the volatility in demand is higher,
- the cost advantage of the global source compared to the holding costs is lower. (Later we shall show that a longer leadtime diminishes the comparative cost advantage of the global source.)

[2] $\frac{d\kappa^i}{dk^i} = \frac{d}{dk^i}[k^i z^i_K + o^i I_N(z^i_K)] = z^i_K$ for all z^i_K as optimal functions of k^i.

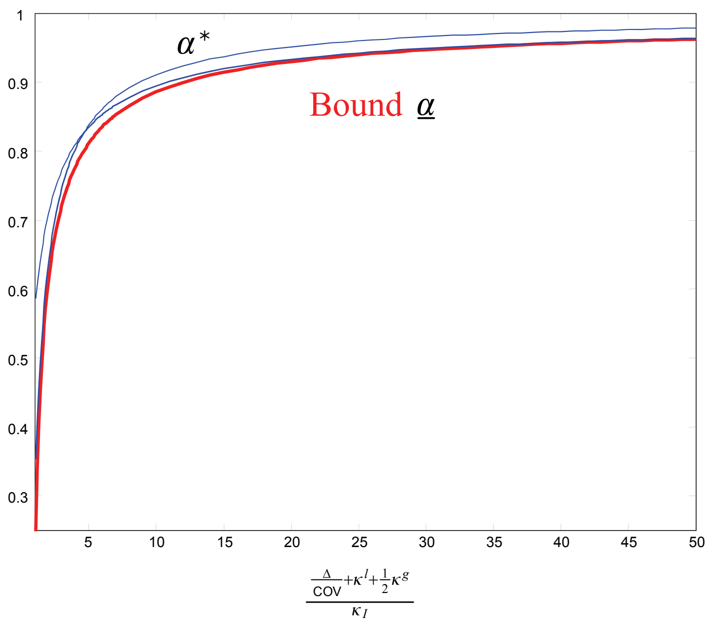

Figure 3: The lower bound $\underline{\alpha}$ provides a good approximation of the optimal strategic allocation α^*, which falls between the two blue curves depending on parameters θ_{13} and θ_2.

2.3 Bounds and Cost Sensitivity

Proposition 8 *If $\theta_{13} + \frac{1}{2}\theta_2 > 1$, then the optimal offshoring allocation has a lower bound*

$$\alpha^* \geq \underline{\alpha} = \sqrt{1 - \left(\theta_{13} + \frac{1}{2}\theta_2\right)^{-2/3}} = \sqrt{1 - \left(\frac{\kappa_I}{\frac{\Delta}{COV} + \kappa^l + \frac{1}{2}\kappa^g}\right)^{2/3}} \tag{26}$$

that is asymptotically tight as $\theta_{13} + \frac{1}{2}\theta_2 \to 0$. Otherwise, we have that $\theta_{13} - \theta_2 < 1 - \frac{3}{2}\theta_2$ and, for small optimal allocations ("light offshoring") a better approximation is

$$\alpha^* \approx \frac{\theta_{13} - \theta_2}{1 - \frac{3}{2}\theta_2} = \frac{\frac{\Delta}{COV} + \kappa^l - \kappa^g}{\kappa_I - \frac{3}{2}\kappa^g}. \tag{27}$$

Proof: The value $\underline{\alpha}$ is the solution of $f_1(\alpha) = \frac{1}{(1-\alpha^2)^{3/2}} = \theta_{13} + \frac{1}{2}\theta_2 = g_1(\alpha)$, where $f_1(\alpha) > \widehat{MC}(\alpha)$ and the right-hand-side is the maximal value of the first-order Taylor expansion $g_1(\alpha) = \theta_{13} - \theta_2 + \frac{3}{2}\theta_2\alpha$ of $\widehat{MB}(\alpha)$. (Represented by point D in Fig. 2.) To establish that $\underline{\alpha}$ is a lower bound, first note that, keeping $\theta_{13} + \frac{1}{2}\theta_2 = x_0$ constant, the minimal value of θ_2 is 0 for which $\widehat{MB}(\alpha) \to x_0$ as $\theta_2 \to 0$ and the associated optimal α^* is defined by point A in Fig. 2. As θ_2 increases to its maximal value where $\theta_{13} = \theta_2 = \frac{2}{3}x_0$ (recall that $\theta_{13} - \theta_2 \geq 0$), the associated optimal α^* "travels down" the $\widehat{MC}(\alpha)$ curve to point B, near point C and may increase again. But it never goes below point C and it is easily verified that C and D are on the same vertical, so that $\alpha^* \geq \underline{\alpha}$. Clearly, as $\theta_{13} + \frac{1}{2}\theta_2 \to 0$, then $\underline{\alpha} \to 1$ and thus also $\alpha^* \to 1$.

From Fig. 2 it can be verified that Eq.(26) only holds when the intercept of $\widehat{MB}(\alpha)$ is larger

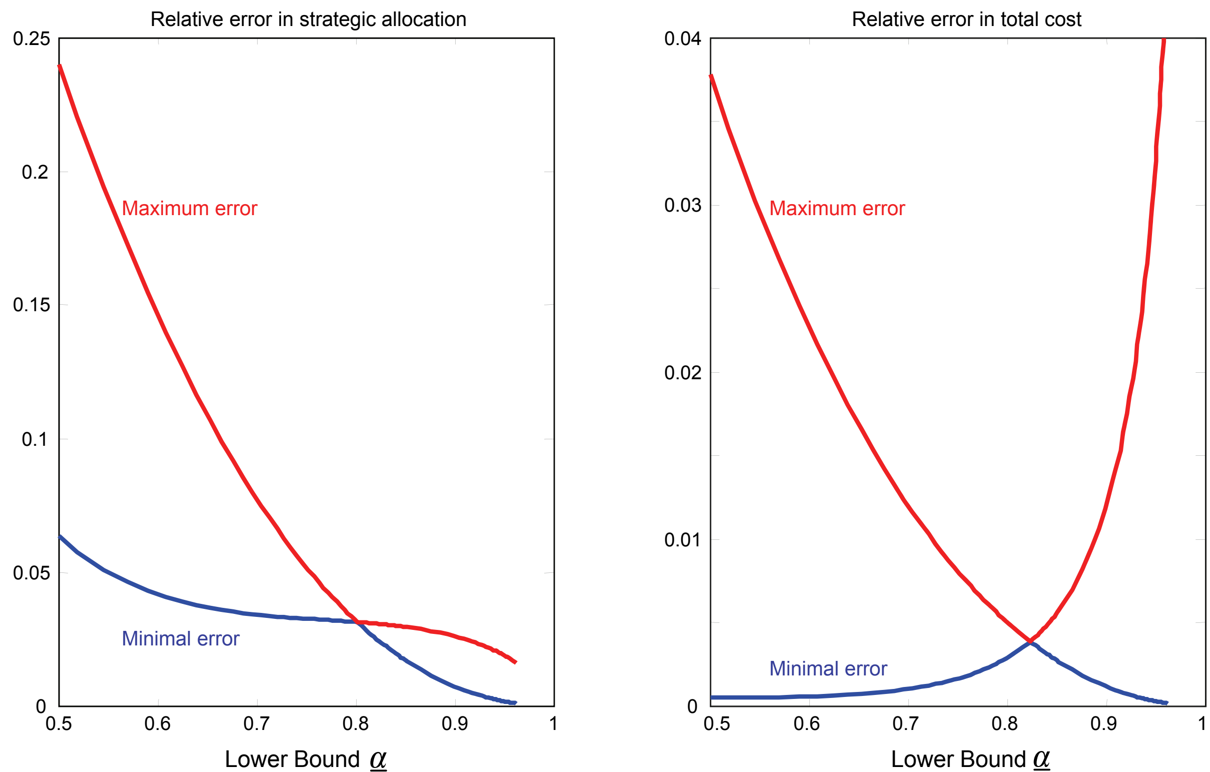

Figure 4: Upper and lower bounds on the error of using the lower bound $\underline{\alpha}$ as an approximation for the optimal strategic allocation α^* (left panel) and associated optimal cost (right panel).

than one. When $\theta_{13} - \frac{1}{2}\theta_2 < 1$, the strategic allocation is approximated by setting $\widehat{MB}(\alpha^*) \approx \theta_{13} - \theta_2 + \frac{3}{2}\theta_2\alpha^* = \widehat{MC}(\alpha^*) \approx \alpha^*$. ■

Fig. 3 shows that the lower bound provides an excellent approximation of the optimal offshoring allocation α^*. We performed a numerical study to quantify the minimal and maximal relative error (scanning all admissible values of θ) in allocation $\frac{\alpha^*-\underline{\alpha}}{\alpha^*}$ and corresponding error in cost. As shown in Fig. 4, the relative error is contained and the bound explains more than 96% of the cost variation.

Not only does the lower bound give a useful simple formula, it also highlights the dominant parameter that affects the strategic allocation is

$$\frac{\Delta}{\kappa_I \mathrm{COV}} = \frac{\frac{(c^l-c^g)+(k^l-k^g)}{h}}{[z_I + (1+b/h)I_N(z_I)]\,\mathrm{COV}} = \frac{\text{financial relative cost advantage}}{\text{service level } \times \text{ relative variability}}. \tag{28}$$

(The other driver is $\frac{\kappa^l+\frac{1}{2}\kappa^g}{\kappa_I}$, the ratio of safety capacity and overtime cost to inventory cost.) It directly shows how the strategic allocation increases in the global source's financial cost advantage, but decreases in the required customer service level and the relative variability in the demand. The bound therefore agrees with the earlier discussed comparative statics on α^*. (Note that the last fraction is key and applies both to the lower bound (for medium and high offshoring) as well as in the other approximation for light offshoring.)

3 General leadtime difference L

We can extend the dual sourcing smoothing policy to the setting where the slow source has a leadtime $L > 1$ by splitting the total received quantity (5) and allocating a fraction to each source. Recall that the quantity q_t^g received from the slow source in period t must be ordered at time $t - L$ and can only depend on D_{t-L-i} for $i = 0, 1, ...$ This directly suggests

$$q_t^l = \sum_{i=0}^{L-1} (1-\alpha)\alpha^i D_{t-i}, \tag{29}$$

$$q_t^g = \alpha^{t+L} q_{-L} + \sum_{i=L}^{t} (1-\alpha)\alpha^i D_{t-i}. \tag{30}$$

Given that the total quantity q_t remains unchanged, the inventory process and cost as described in Propositions 2 and 5 remain unchanged by the larger leadtime L. But obviously the orders to both sources change: taking expectations and variances in (29)-(30) yields: if we set $q_{-L} = \alpha^{1-L}\mu$

$$\begin{aligned}
&\mathbb{E}q_t = \mu, && Var(q_t) = \left[1 - \alpha^{2(t+1)}\right] \tfrac{1-\alpha}{1+\alpha}\sigma^2 \leq \sigma^2, \\
&\mathbb{E}q_t^g = \alpha^L \mu, && Var(q_t^g) = \left[1 - \alpha^{2(t-L+1)}\right] \tfrac{1-\alpha}{1+\alpha}\alpha^{2L}\sigma^2 \leq \sigma^2, \\
&\mathbb{E}q_t^l = \left(1 - \alpha^L\right)\mu, && Var(q_t^l) = \tfrac{1-\alpha}{1+\alpha}\left(1 - \alpha^{2L}\right)\sigma^2 \leq \sigma^2.
\end{aligned}$$

The important implication from $\mathbb{E}q_t^g = \alpha^L\mu$ is that with general leadtime $L > 1$ the strategic allocation (the fraction of average total orders allocated to the global source) is α^L. This is *different* and smaller than the smoothing level α.

We can now characterize the sourcing and capacity costs (recall that C_I remains as before):

$$\begin{aligned}
C_s &= \mu[c^l - \alpha^L(c^l - c^g)], \\
C_K^l &= k^l\mu(1 - \alpha^L) + \left[k^l z_K^l + o^l I_N(z_K^l)\right]\sigma\sqrt{\frac{1-\alpha}{1+\alpha}(1 - \alpha^{2L})}, \\
C_K^g &= k^g\mu\alpha^L + \left[k^g z_K^g + o^g I_N(z_K^g)\right]\sigma\sqrt{\frac{1-\alpha}{1+\alpha}\alpha^{2L}}.
\end{aligned}$$

We now also include the expected pipeline inventory cost $hL\mu\alpha^L$ because it impacts the optimal smoothing and allocation. If we slightly adjust notation by redefining

$$C_0 = \left(c^l + k^l\right)\mu \qquad \text{and} \qquad \Delta = (c^l - c^g) + (k^l - k^g) - hL,$$

then the total cost rate is

$$C(\alpha) = C_0 - \Delta\mu\alpha^L + \kappa^l\sigma\sqrt{\frac{1-\alpha}{1+\alpha}(1 - \alpha^{2L})} + \kappa^g\sigma\alpha^L\sqrt{\frac{1-\alpha}{1+\alpha}} + \kappa_I\frac{\sigma}{\sqrt{1-\alpha^2}}. \tag{31}$$

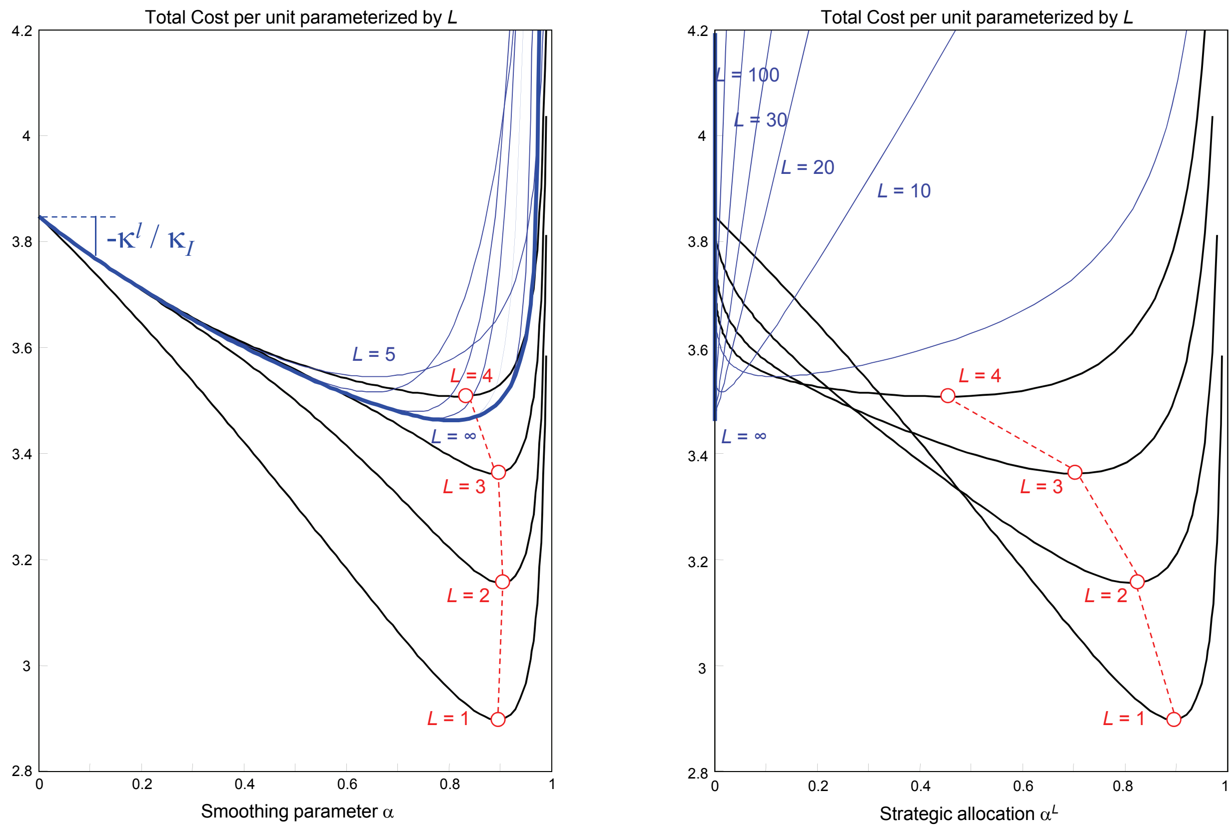

Figure 5: Total cost and corresponding optimal smoothing level α^* as a function of L (left panel). The optimal strategic allocation α^L decreases as L increases (right panel). As $L \to \infty$, the total cost curve converges uniformly with unique minimum α^*_∞. (Parameters: $\mu = 50$, $\sigma = 20$, $c^g = 1$, $c^l = 1.4$, $k^l = 1.5$, $k^g = 1$, $o^i = 4k^i$, $h = 0.224$, $b = 19h$.)

Figure 5 illustrates the impact of a change in leadtime L (using the same parameters as Fig. 1 but with lower unit holding cost $h = 0.224$ for Δ to remain positive up to $L = 4$). The left panel shows that, as L increases, the minimal cost typically increases. The corresponding unique optimal smoothing level initially increases as L increases to 2, but then decreases as L increases to 5, but then eventually increases again. As L increases further, the total cost curve converges to the curve in bold and the optimal smoothing level converges to α^*_∞ (= .8 in the figure).

Indeed, as $L \to \infty$, the cost function C converges uniformly over $[0, 1)$ to C_∞, where

$$C_\infty(\alpha) = C_0 + \kappa^l \sigma \sqrt{\frac{1-\alpha}{1+\alpha}} + \kappa_I \frac{\sigma}{\sqrt{1-\alpha^2}}, \tag{32}$$

which has a unique minimum

$$\alpha^*_\infty \quad = \quad \kappa^l / (\kappa^l + \kappa_I). \tag{33}$$

This limiting smoothing level thus decreases as inventory costs increase relative to capacity costs. The limiting cost C_∞ represents the cost of a single local source with exponential smoothing.

(Proposition 2 shows its order variance $Var(q_t)$.)

This does not mean, however, that one should rely more on the global source as L rises. Indeed, recall that the strategic allocation is α^L and looking at the cost as function of α^L (right panel) we see that the optimal allocation decreases as L increases. For larger leadtimes ($L > 10$ in the figure), the strategic allocation converges to a small but positive level $(\alpha^*_\infty)^{-L}$ (about $.8^{-L}$ in the figure). This convergence is quicker when inventory costs increase relative to capacity costs (then α^*_∞ decreases). Theoretically, this means that a dual sourcing strategy with very small reliance on the global source remains optimal for high leadtimes. In practice, this suggests single local sourcing with production smoothing (recall that our model does not include this strategy).[3]

Proposition 9 *If $L > 1$, the total scaled cost rate has three independent parameters θ_1, θ_2 and θ_3:*

$$\widehat{C}(\alpha;\theta_1,\theta_2,\theta_3) = \frac{C(\alpha) - C_0}{\kappa_I \sigma} = -\theta_1 \alpha^L + \theta_2 \alpha^L \sqrt{\frac{1-\alpha}{1+\alpha}} + \theta_3 \sqrt{\frac{1-\alpha}{1+\alpha}\left(1-\alpha^{2L}\right)} + \frac{1}{\sqrt{1-\alpha^2}}$$

and the optimal smoothing level α^* satisfies $0 < \alpha^* < 1$ and $\widehat{MB}(\alpha^*) = \widehat{MC}(\alpha^*)$ where:

$$\begin{aligned}\widehat{MB}(\alpha) &= \theta_1 L \alpha^{L-1} - \theta_2 \frac{L\alpha^{L-1}(1-\alpha)(1+\alpha) - \alpha^L}{(1-\alpha)^{1/2}(1+\alpha)^{3/2}} + \theta_3 \frac{L\alpha^{2L-1}(1-\alpha)(1+\alpha) + \left(1-\alpha^{2L}\right)}{\left(1-\alpha^{2L}\right)^{1/2}(1-\alpha)^{1/2}(1+\alpha)^{3/2}} \\ \widehat{MC}(\alpha) &= \frac{\alpha}{\left(1-\alpha^2\right)^{3/2}}.\end{aligned}$$

As $L \to \infty$, the optimal smoothing level and cost converge

$$\begin{array}{l}\alpha^* \to \alpha^*_\infty = \frac{\theta_3}{1+\theta_3} \\ \min \widehat{C} \to \underline{\widehat{C}}_\infty = \theta_3 \sqrt{\frac{1-\theta_3}{1+\theta_3}} + \frac{1+\theta_3}{\sqrt{1+2\theta_3}}\end{array}, \tag{34}$$

so that the optimal strategic allocation $\alpha^{*L} \simeq \alpha^{*L}_\infty \to 0$ as $L \to \infty$.

Proof: The necessary optimality conditions follow directly from differentiation. Given that $\widehat{MB}(0) = \theta_3 > \widehat{MC}(0) = 0$, $\alpha^* > 0$. Given that $C(\alpha) \to \infty$ as $\alpha \to 1$, $\alpha^* < 1$. As $L \to \infty$, the function $\widehat{MB}$ converges uniformly to $\widehat{MB}_\infty$ over $[0,1)$, where

$$\widehat{MB}_\infty(\alpha) = \theta_3 (1-\alpha)^{-1/2} (1+\alpha)^{-3/2}. \tag{35}$$

Consequently, the optimal smoothing level $\alpha^* \to \alpha^*_\infty$ as $L \to \infty$ where

$$\widehat{MB}_\infty(\alpha^*_\infty) = \frac{\theta_3}{(1-\alpha^*_\infty)^{1/2}(1+\alpha^*_\infty)^{3/2}} = \widehat{MC}(\alpha^*_\infty) = \frac{\alpha^*_\infty}{(1-\alpha^{*2}_\infty)^{3/2}} \Leftrightarrow \alpha^*_\infty = \frac{\theta_3}{1+\theta_3}.$$

■

[3] As long as the global source has a pure cost advantage $\Delta > 0$, which is the "natural" regime for dual sourcing practice, a higher leadtime results in a higher cost. Otherwise, pipeline costs overwhelm the sourcing cost advantage of the global source. Theoretically, it is still advantageous in our model to dual source because the very small reliance on the global source effectively allows the local source to smooth and reduce capacity requirements.

The proposition directly specifies an asymptotic estimate α_∞^L for the strategic allocation for large L. While that result is only theoretically appealing (as it implies infinitesimally small offshoring), it does also specify an estimate on the minimal cost under dual sourcing using smoothing. In addition, we propose a simple approximation for small to moderate L found by a similar procedure as that of $L = 1$: approximating the marginal benefit terms by their maximal first component $\theta_1 L$ and the values $\frac{1}{2}\theta_2 + \theta_3$ at $\alpha = 0$ to approximate other components (and make the formula consistent with the lower bound for $L = 1$), and the marginal cost by its upper bound $\left(1-\alpha^2\right)^{-3/2}$ yields:

$$\begin{aligned}\widehat{MB}(\alpha_0) &\simeq \theta_1 L + \theta_2 + \theta_3 \simeq \frac{1}{\left(1-\alpha_0^2\right)^{3/2}} \\ &\Rightarrow \alpha_0 = \sqrt{1-\left(\theta_1 L + \frac{1}{2}\theta_2 + \theta_3\right)^{-2/3}} = \sqrt{1-\left(\frac{\kappa_I}{\frac{L\Delta}{\text{COV}} + \kappa^l + \frac{1}{2}\kappa^g}\right)^{2/3}}\end{aligned}$$

Our analysis then culminates in our analytic estimate for the strategic allocation:

$$\alpha^{*L} \simeq \alpha_0^L = \left(1-\left(\theta_1 L + \frac{1}{2}\theta_2 + \theta_3\right)^{-2/3}\right)^{L/2} = \left(1-\left(\frac{\kappa_I}{\frac{L\Delta}{\text{COV}} + \kappa^l + \frac{1}{2}\kappa^g}\right)^{2/3}\right)^{L/2}. \quad (36)$$

The offshoring allocation estimate α_0^L is applicable when the denominator is positive, which is guaranteed in the natural regime when $\Delta > 0$. Unfortunately, when $L > 1$, the marginal cost function is too complex and the terms in θ_2 and θ_3 lack monotonicity properties to establish a bound on α^* but α_0 provides a useful approximation: Return to the example with base case parameters but $h = 0.224$ shown in Figure 5: the left panel of Figure 6 compares the exact optimal allocation α^{*L} with the approximation α_0^L, which results in a very small relative cost penalty $C(\alpha_0)/C(\alpha^*)$. The important observation is that the simple formulas capture the trends well and thus provide guidance in decision making.

4 Reference Policies

In the remainder of this paper we investigate the quality and robustness of our proposed dual sourcing - order smoothing policy. First we describe some important reference policies against which we will compare our policy's performance.

4.1 Single Sourcing Base-Stock Policies

Single sourcing from either the local or global source goes back to traditional inventory models. The standard periodic review base-stock policy is optimal in minimizing inventory related costs in absence of a fixed order cost (Zipkin 2000). Placing every period an order after demand is observed, leads to a chase policy, where the order to the single source equals observed demand.

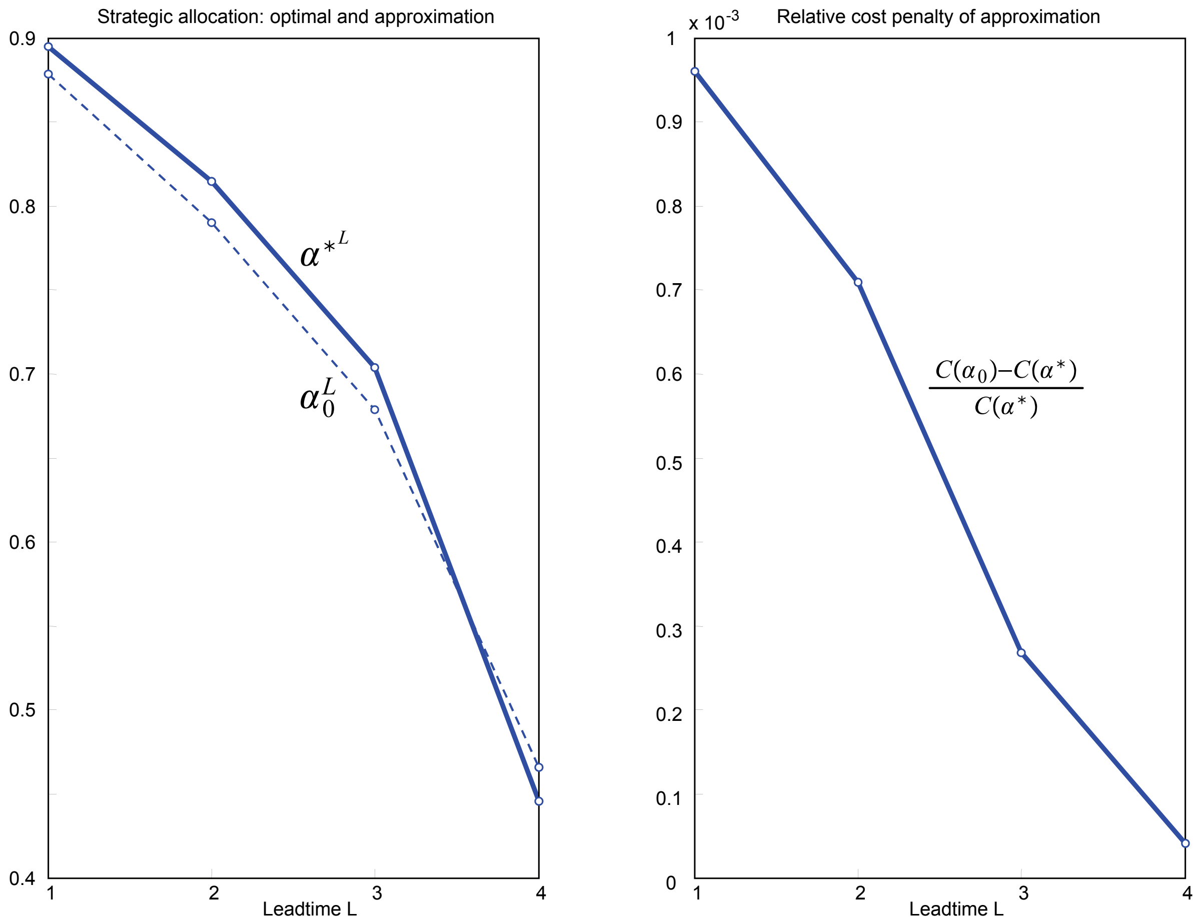

Figure 6: The approximation α_0^L provides a good estimate of the optimal strategic allocation α^{*L} (left panel) that results in very small relative cost penalty (right panel). (Parameters: $\mu = 50$, $\sigma = 20$, $c^g = 1$, $c^l = 1.4$, $k^l = 1.5$, $k^g = 1$, $o^i = 4k^i$, $h = 0.224$, $b = 19h$.)

Single sourcing from the local source under a standard base-stock policy yields $q_t^l = D_t$ and $q_t^g = 0$. Hence, $\mathbb{E}q_t^l = \mu$ and $Var(q_t^l) = \sigma^2$. The associated inventory process is

$$I_t = I_{t-1} + q_{t-1}^l - D_t = I_{-1} + q_{-1}^l - D_t = I_s + \mu - D_t,$$

so that $\mathbb{E}I_t = I_s$ and $Var(I_t) = \sigma^2$. The total average cost is $C = c^l\mu + k^l\mu + \kappa^l\sigma + \kappa_I\sigma$.

Under global single sourcing we have to take the leadtime into account, so that $q_t^g = D_{t-L}$ and $q_t^l = 0$. Then, $\mathbb{E}q_t^g = \mu$ and $Var(q_t^g) = \sigma^2$, and the inventory process is

$$I_t = I_{t-1} + q_{t-1}^g - D_t = I_{-1} + \sum_{i=1}^{L+1} q_{-i}^g - \sum_{i=t-L}^{t} D_i = I_s + (L+1)\mu - \sum_{i=t-L}^{t} D_i,$$

so that $\mathbb{E}I_t = I_s$ and $Var(I_t) = (L+1)\sigma^2$. The average cost is then $C = c^g\mu + k^g\mu + \kappa^g\sigma + \sqrt{(L+1)}\kappa_I\sigma$. Compared to single local sourcing, this policy benefits from lower sourcing and capacity unit costs while the total installed capacity remains the same. In contrast, inventory costs increase due to the longer leadtime.

4.2 Standing Order or Tailored Base-Surge (TBS) Policy

One dual sourcing benchmark is the standing order, or tailored base-surge (TBS), policy. The standing order policy is a base-stock type policy: the global source supplies every period a constant rate $q^g < \mu$, and an additional order to the local source is placed only when the inventory falls below the target level s, or

$$\begin{aligned} q_t^g &= q^g, \\ q_t^l &= (s - I_t - q^g)^+, \qquad \text{and} \qquad E(q_t^l) = \mu - q^g. \end{aligned}$$

The main difference with the standard (single source) base-stock policy is that this policy is not a "demand replacement"policy: when the standing order exceeds the observed demand, it results in excess inventory excursions above s. This is equivalent to a base-stock policy with negative demand studied by DeCroix et al. (2005): Let Z_t denote the "excess inventory process," which is the amount of inventory above the base-stock in period t. Then, under a TBS policy, the excess inventory dynamics follow $Z_t = [Z_{t-1} + (q^g - D_{t-1})]^+$. Its stationary random variable satisfies $Z = [Z + q^g - D]^+$, and the corresponding equilibrium distribution of the net inventory can be expressed as $I = s - D + Z$.

Essentially, Z is a regulated random walk capturing the occasional excess inventory excursions. This is equivalent to the Lindley equation describing the waiting time in a $GI/G/1$ queue. Unfortunately, Z cannot be solved analytically in general (Baron 2008). However, because Z is nonnegative, I is stochastically greater than $s - D$; i.e., $P(I \geq x) \geq P(s - D \geq x)$ for all x. Therefore, the optimal safety stock under a dual sourcing TBS policy will be at least as high as the safety stock under a standard single sourcing base-stock policy. Moreover it can be seen that the safety stock increases as the standing order to the global source goes up, since the exponential tail of the regulated random walk will be longer as $(\mu - q^g)$ decreases (Harrison 1985).

The constant standing order obviates the need for safety capacity in the global source and its average supply rate equals its installed capacity. In contrast, due to the variability in the order stream to the local source, its safety capacity will be positive. We do know that as the standing order increases, one will order less frequent from the local source, or equivalently, the orders to the local source will be more often *"cut off"* to zero. More specifically,

$$q_t^l = \begin{cases} 0 & \text{if } D_t < Z_t + q^g, \\ D_t - q^g & \text{if } D_t \geq Z_t + q^g. \end{cases} \tag{37}$$

As a result, increasing the standing order to the global source reduces the variability in orders to the local source, and so will its safety capacity.

In other words, more offshoring to the global source (i) reduces sourcing costs compared to sole sourcing from the local source; (ii) the total installed (safety) capacity goes down due to the smoothing, peak-shaving, nature of the order stream; (iii) this smoothing effect comes at the cost of an increased inventory costs. Hence, dual sourcing under a TBS policy boils down to the same sourcing-capacity-inventory trade-offs that we face in the dual sourcing smoothing policy. However, since the inventory process behaves like a regulated random walk, the TBS policy is analytically intractable: we are unable to analytically determine the variances of the local orders and the inventory process, and hence derive its total cost function.

If we use the standing order policy to sole-source from the global source, we essentially obtain a pure level policy: $q_t^g = \mu$ and $q_t^l = 0$. Due to the absence of variability in orders, there is no need to have safety capacity, so that $K = \mu$. However, as t grows, the inventory process has growing variance. Indeed, the safety stock under normal demand grows unbounded like $z_I\sqrt{t}\sigma$ and similarly the long-run total cost. While single global sourcing under a standing order policy minimizes capacity and sourcing costs, it is not a viable sourcing option due to the instability of the inventory process and the corresponding infinite inventory cost.

4.3 Dual-Index Base-Stock Policy

When only sourcing and inventory costs are considered, the dual-index base-stock policy is shown to be optimal for $L = 1$ (Fukuda 1964) and near-optimal for longer leadtimes (Veeraraghavan and Scheller-Wolf 2008). The dual-index policy tracks inventory positions over both the local and global leadtime: if the local inventory position is below the local base-stock level, it is brought back to this level by placing a local (expedited) order; after the local order is made, global (regular) orders are placed, restoring the global inventory position to its global target level. In this model, there is no cap to the local and global orders, but in case they exceed the installed capacity, the remaining units are produced in overtime. To constrain the order volatility and thus the capacities at each source, Veeraraghavan and Scheller-Wolf (2008, Section 5) suggest to implement caps on the order volumes. A capacitated model which truncates the order quantity is in essence a smoothing policy.

Similar to TBS, the dual-index base-stock is not a demand replacement policy: the global order may push the local inventory position above its target level, causing an overshoot (and in which case no local order is placed). Since there is no general closed-form distribution for the overshoot, and therefore neither for the orders and net inventory, this policy is not amenable to exact analysis and one has to rely on a numerical study.

5 Numerical study

We conduct a numerical simulation study to examine the performance of our dual sourcing policy. We use the long run average unit cost performance measure and we compare the performance of the dual sourcing smoothing (DSS) policy with the tailored base-surge (TBS), the dual-index (DI) policy and the local resp. global single sourcing base-stock policy (LSS resp. GSS) under a variety of parameter values. We do not compute the optimal cost or the optimal policy – if we

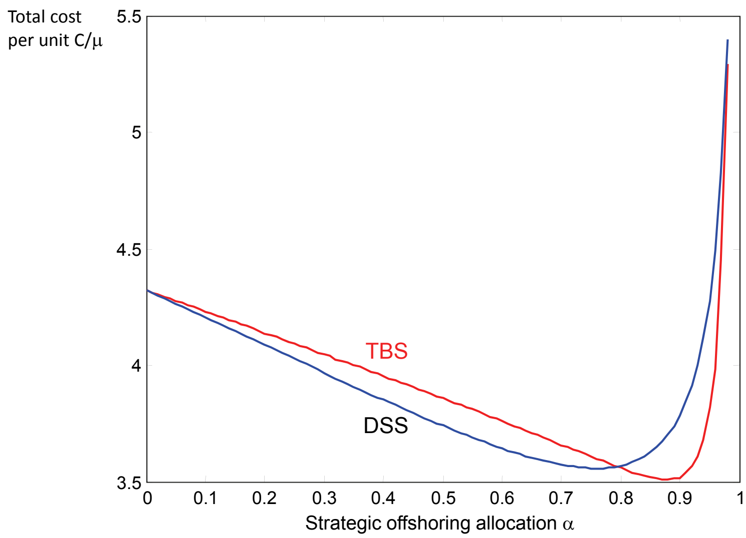

Figure 7: Total unit cost for DSS and TBS in function of the offshoring allocation (Parameters: COV= 0.4, $c^g = 1$, $c^l = 1.4$, $k^g = 1$, $k^l = 1.5$ $h = 0.8$, $L = 1$)

were to compute the cost of the optimal policy to benchmark our heuristics, we would have to restrict our computational study due to the computational complexity of the dynamic program (see Veeraraghavan and Scheller-Wolf (2008) and Sheopuri et al. (2010)). For the TBS and dual-index policies, we numerically optimized the target order-up-to levels and the installed capacities at each source to minimize the sum of average sourcing, inventory and capacity costs per unit.

Base case scenario

The base scenario assumes $c^l = \$1.4, c^g = \$1, k^l = \$1.5, k^g = \$1$ and $h = \$0.8$. A backlog cost $b = 19h$ and overtime costs $o^i = 4k^i$ ensures a service level of 95% in inventory and 75% in capacity. Demand is non-negative with coefficient of variation COV=0.4, and we start with leadtime $L = 1$. We focus on the average total cost per unit sourced, which is in steady state equivalent to C_t/μ.

We first compare the two "smoothing"policies, DSS and TBS. Figure 7 illustrates that under both DSS and TBS, the long run average total unit cost decreases when we offshore compared to single local sourcing ($\alpha = 0$), but for high levels of offshoring (α close to 1) it increases rapidly. The former is due to the decrease in sourcing costs and capacity costs when we offshore (and smooth orders), whereas the latter is due to the sharp increase in inventories as a result of the strong order dampening when we heavily rely on the global source. In the DSS policy, this sharp increase occurs for values $\alpha \geq 0.8$; a similar effect is observed in the TBS policy, albeit somewhat later ($\alpha \geq 0.9$), since the smoothing in orders only takes place in that region (this observation was also confirmed by monitoring the safety capacity levels under both policies). We observed this phenomenon for a wide variety of parameter values. Therefore, the optimal offshoring level in a TBS policy tends to

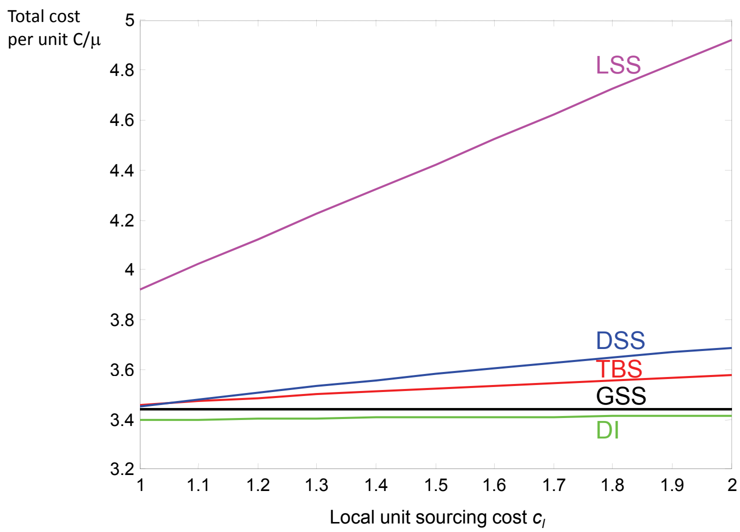

Figure 8: Average unit cost performance in function of the local sourcing cost (Parameters: COV= 0.4, $c^g = 1$, $k^g = 1$, $k^l = 1.5$, $h = 0.8$, $L = 1$)

be higher than the optimal allocation under DSS. Observe that DSS performs better for low levels of offshoring compared to TBS.

Under its optimal allocation, DSS performs 21.5% better than local single sourcing in this base case scenario. Dual sourcing lowers sourcing costs, but also capacity costs, while inventory costs increase due to the smoothing behavior. Under its optimal allocation in this base case, DSS performs 1.34% worse than TBS, and compared to the more complex pure dual-index (DI) policy, DSS has 4.2% higher costs. Notably, the order-up-to policy when sole sourcing from the global source performs 3.2% better than the DSS policy. Compared to global single sourcing, we find that dual sourcing under the DSS policy increases sourcing costs, but inventory costs tend to increase as well, i.e. the additional "smoothing"safety stock under dual sourcing is typically higher than the additional "leadtime"safety stock under sole global sourcing. So in this case the only benefit from dual sourcing is the capacity cost decrease due to smoothing.

Effect of the sourcing cost difference

Figure 8 illustrates the impact of changing the local sourcing cost on the average total unit cost. The dual sourcing smoothing (DSS) policy performs within a range of 3% of the dual-index (DI) and TBS policy. Note that the DI policy seems to be rather insensitive to the sourcing cost, which is due to the fact that under this range of parameter values, the allocation to the low-cost source is close to 100%. This is in line with the numerical study reported by Scheller-Wolf et al. (2006) where only a negligible portion of expedited orders is placed for comparable parameter values.

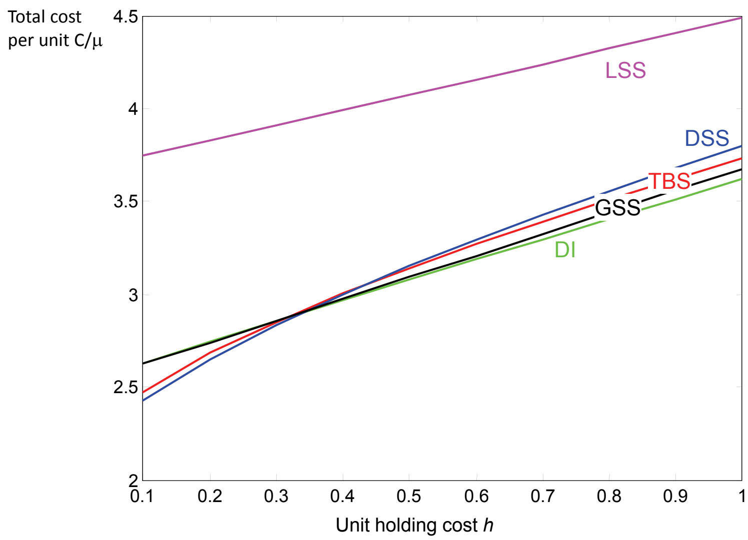

Figure 9: Average unit cost performance in function of holding cost (Parameters: COV$=0.4$, $c^g=1$, $c^l=1.4$, $k^g=1$, $k^l=1.5$, $L=1$)

Effect of the inventory cost relative to the capacity cost

Figure 9 shows the cost performance for different values of h. Given that both DSS and TBS are smoothing the order stream, and this smoothing comes at the cost of increased inventory, both DSS and TBS perform relatively better compared to the other policies when the unit inventory cost is reduced relative to its capacity cost. For h ranging from 0.1 to 1, the difference between DSS and TBS is less than 2%, with DSS performing better when capacity costs is of higher concern than inventory costs, because DSS dampens the orders to a larger extent than TBS. Besides, since the sharp increase in inventories for high levels of offshoring makes DSS perform worse than the other policies, this effect is dampened for lower inventory costs.

Interestingly, DSS performs better than the dual-index policy DI for low values of inventory cost. The pure dual-index policy performs close to optimal under the assumption of minimizing sourcing and inventory related costs, but our model additionally includes capacity costs at both sources. Since the offshoring levels under both DSS and TBS are increasing for lower inventory costs (e.g., for $h=0.1$, the optimal offshoring level under DSS and TBS is resp. 95% and 97%), it implies that the orders are considerably dampened in variability, with lower safety capacities (and hence lower capacity costs) as a result, whereas the DI policy does not dampen the order volumes.

Effect of local vs. global capacity cost

Figure 10 measures the impact of changing the local capacity costs. The relative cost difference between DSS and TBS is less than 2% for a range of k^l values, and DSS outperforms TBS if k^l increases beyond 2. In other words, as the local source becomes more expensive compared to the

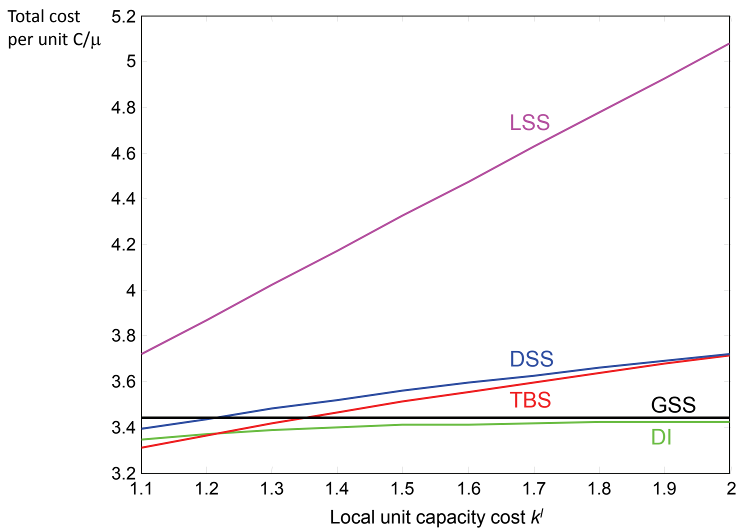

Figure 10: Average unit cost performance in function of local capacity cost (Parameters: COV= 0.4, $c^g = 1$, $c^l = 1.4$, $k^g = 1$, $h = 0.8$, $L = 1$)

global source, DSS performs better (closer) vis-a-vis TBS, because DSS smoothes local capacity to a larger extent compared to TBS: the DSS policy splits the total safety capacity over both the local and the global source; the TBS policy, on the contrary, has all its safety capacity on the local source. Moreover, as local capacity gets more expensive, both policies rely more on the global source, and the difference between TBS and DSS gets smaller.

Effect of demand variability

In this experiment we measure the impact of the demand variability. Given that we have normal demand, we simulate up to a coefficient of variation of 0.4 to minimize the probability of negative demand. Figure 11 illustrates that as demand is more variable, all policies increase in cost, but DSS has a slightly smaller slope than TBS and DI.

Effect of leadtime

Figure 12 illustrates the impact of the leadtime on the cost performance of the various policies. This experiment differs from the previous experiments in that we now include the pipeline inventory costs. To ensure that Δ remains positive (the "natural" regime), we set $c^l = 2.5$ and $h = 0.4$, and simulate up to a leadtime $L = 4$. We observe that the DSS policy provides a lower total cost performance compared to TBS and DI for large values of L. For large values of L, we also find that the TBS policy performs better than the dual index. They all perform better than the local single sourcing policy, indicating that even for large values of L, dual sourcing is beneficial.

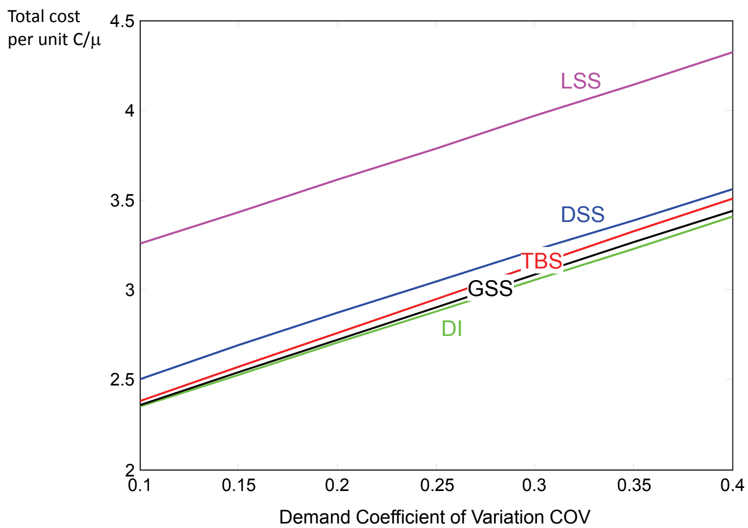

Figure 11: Average unit cost performance in function of demand variability (Parameters: $c^g = 1$, $c^l = 1.5$, $k^g = 1$, $k^l = 1.5$, $h = 0.8$, $L = 1$.)

Summary of computational study

Our computational experiments show that our dual sourcing smoothing policy performs very close to the tailored base-surge and dual-index policies. Especially when capacity is more expensive relative to inventory, when the sourcing cost advantage is moderate, when demand is highly variable, and when leadtimes are long, the DSS policy performs exceptionally well, and can even outperform the other policies.

A noteworthy observation is that single global sourcing may perform very close to dual sourcing in the context of our model. Yet it is fair to say that dual sourcing has various other risk mitigation benefits (e.g. supply risk, currency risk, dynamic changes in parameters) that we have not considered.

6 Summary and Discussion

The main research objective of this paper is strategic in nature: to determine a near-optimal average sourcing allocation to two sources. We used a linear control rule that allocates the sourcing volume in every period to a global cost-efficient source and a local responsive source. The linear control makes the policy analytically tractable and allows even closed form solutions. The analysis is much simpler and more accessible that any other analytical model discussed in the literature. Moreover, it captures many key parameters that managers consider relevant in their sourcing decisions. The model is also directly applicable to guide how a shipper can use two transport modes together for a single commodity flow.

The model provides the following insights and quantification on the strategic base or offshoring

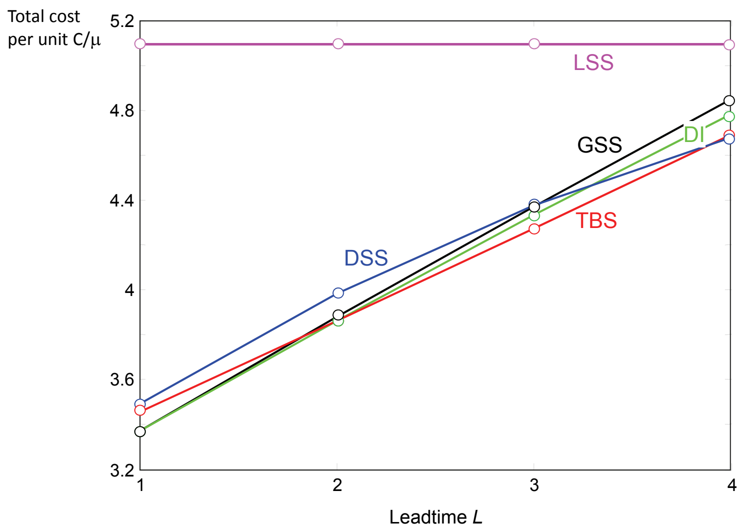

Figure 12: Average unit cost performance in function of Leadtime (Parameters: COV= 0.4, $c^g = 1$, $c^l = 2.5$, $k^g = 1$, $k^l = 1.5$, $h = 0.4$)

allocation. First, we provide a lower bound to the strategic allocation volumes when the leadtime difference is one review period. Not only does the lower bound give a useful simple formula, it also provides an excellent approximation and it highlights the dominant parameters that affect the strategic allocation: the strategic allocation increases in the global source's financial cost advantage, but decreases in the required customer service level and the relative variability in the demand. Second, we propose a simple approximation for small to moderate leadtime differences, which quantifies to what extent the optimal offshore allocation decreases as L increases. For large leadtimes we show that the dual sourcing strategy shifts to single local sourcing with production smoothing – we knew that intuitively but are not aware of any analytic or structural results in a dual sourcing model. Third, we determine the optimal safety stock at destination and capacity levels at each source. Our dual sourcing model is in essence a production smoothing model: increased offshoring to the global source allows to reduce the total capacity investments, but the smoothing comes at the cost of an increased safety stock investment. Finally, our model is capable to give guidance on practice. A comparative statics analysis confirms our intuition: higher sourcing cost difference and higher capacity costs at the local source increase the order allocation to the global cost efficient source. Higher inventory costs, more demand volatility and longer leadtimes increase the volume sourced from the local responsive source. The impact of the capacity costs at the global source is concave in the local order portion, and is less decisive in the optimal order allocation.

A numerical study confirms that our dual sourcing policy performs close to other existing dual sourcing policies in terms of total cost minimization. Especially when capacity is more expensive relative to inventory and when the leadtime difference is long, dual sourcing smoothing performs exceptionally well, and can even outperform the more sophisticated dual-index policy.

As with any analytic study, ours begs for extensions. Direct extensions could study demand distributions with "fat tails," which we expect to favor more local sourcing, or with serial demand correlation, where negative correlation will favor smoothing and global sourcing. Another extension could analyze non-stationary demand patterns to model for example how the sourcing allocation should adapt to the product life cycle.[4] A third extension could insert unit cost or quantity uncertainty to model for example currency exchange rate risk or supply quality risk. Last, but not least, would be a game-theoretic multi-decision maker model where one firm (the buyer) sources from one or two independent firms. This brings up the study of contracting and equilibrium investment and sourcing for which there is a literature on two-stage models (see, e.g., Van Mieghem (1999) for a first bargaining contract approach in a newsvendor network model and recently Wu et al. (2011) for volume commitment or capacity reservation contracts). These non-obvious model extensions are full research projects by themselves.

References

Allon, G., J. A. Van Mieghem. 2010. Global dual sourcing. *Management Science* **56**(1) 110–124.

Balakrishnan, A., J. Geunes, M. S. Pangburn. 2004. Coordinating supply chains by controlling upstream variability propagation. *Manufacturing & Service Operations Management* **6**(2) 163–183.

Baron, O. 2008. Regulated random walks and the lcfs backlog probability: analysis and application. *Operations Research* **56**(2) 471–486.

Bertrand, J.W.M. 1986. Balancing production level variations and inventory variations in complex production systems. *International Journal of Production Research* **24**(5) 1059–1074.

Boute, R. N., S. M. Disney, M. R. Lambrecht, B. Van Houdt. 2007. An integrated production and inventory model to dampen upstream demand variability in the supply chain. *European Journal of Operational Research* **178**(1) 121–142.

Bray, R., H. Mendelson. 2011. Information transmission and the bullwhip effect: an empirical investigation. *Proceedings of the 2011 MSOM Annual Conference*. Ann Arbor, Michigan.

Cachon, G.P., T. Randall, G.M. Schmidt. 2007. In search of the bullhwip effect. *Manufacturing & Service Operations Management* **9**(4) 457–479.

Cantor, D.E., E. Katok. 2011. The bullwhip effect and order smoothing in a serial supply chain. Working Paper, Smeal College of Business, Penn State University.

Chen, L., H. Lee. 2011. Bullwhip effect measurement and its implications. Working Paper, Graduate School of Business, Stanford University.

Combes, François. 2011. Inventory theory and mode choice in freight transport: The case of the simultaneous use of two transport modes on one shipper-receiver relationship. Working Paper, Université Paris Est and Association for European Transport and Contributors.

DeCroix, G., J.-S. Song, P. Zipkin. 2005. A series system with returns: stationary analysis. *Operations Research* **53**(2) 350–362.

Forrester, J. 1961. *Industrial Dynamics*. MIT Press, Cambridge MA.

[4]One could apply our formula for each phase separately to predict that the sourcing allocation will move from emphasizing local during product introduction (low mean but high variability in demand) to offshore during maturity and back to local during decline. However, this remains an unverified extrapolation of a stationary result.

Fukuda, Y. 1964. Optimal policies for the inventory problem with negotiable lead time. *Management Science* **10**(4) 690–708.

Graves, S. C. 1988. Safety stocks in manufacturing systems. *Journal of Manufacturing and Operations Management* **1**(1) 67–101.

Harrison, J.M. 1985. *Brownian motion and stochastic flow systems*. Krieger Publishing Company, Malabar, Florida.

Jain, N., K. Girotra, Netessine S. 2011. Managing global sourcing: an empirical study. *Proceedings of the 2011 MSOM Annual Conference*. Ann Arbor, Michigan.

Liu, Z., A. Nagurney. 2011. Supply chain networks with global outsourcing and quick-response production under demand and cost uncertainty. Working Paper, University of Massachusetts, Massachusetts.

Magee, J. F. 1958. *Production Planning and Inventory Control*. McGraw-Hill, New York.

Maher, K., B. Tita. 2010. Caterpillar joins 'onshoring' trend. *Wall Street Journal, March 11* .

Rosenshine, M., D. Obee. 1976. Analysis of a standing order inventory system with emergency orders. *Operations Research* **24**(6) 1143–1155.

Scheller-Wolf, A., S. Veeraraghavan, G.J. van Houtum. 2006. Effective dual sourcing with a single index policy. Working Paper, Stern School of Business, New York University, New York.

Sheopuri, A., G. Janakiraman, S. Seshadri. 2010. New policies for the stochastic inventory control problem with two supply sources. *Operations Research* **58**(3) 734–745.

Simon, H. 1952. On the application of servomechanism theory in the study of production control. *Econometrica* **20** 247–268.

Sobel, M. 1969. Production smoothing with stochastic demand I: finite horizon case. *Management Science* **16**(3) 195–207.

Sobel, M. 1971. Production smoothing with stochastic demand II: inffinite horizon case. *Management Science* **17**(11) 724–735.

Tagaras, G., D. Vlachos. 2001. A periodic review inventory system with emergency replenishments. *Management Science* **47**(3) 415–429.

Van Mieghem, Jan A. 1999. Coordinating investment, production and subcontracting **45**(7) 954–971.

Vassian, H. J. 1955. Application of discrete variable servo theory to inventory control. *Operations Research* **3** 272–282.

Veeraraghavan, S. K., A. Scheller-Wolf. 2008. Now or later: a simple policy for effective dual sourcing in capacitated systems. *Operations Research* **56**(4) 850–864.

Whittemore, A.S., S. C. Saunders. 1977. Optimal inventory under stochastic demand with two supply options. *European Journal of Operational Research* **178**(1) 121–142.

Wu, X., F. Zhang. 2011. Efficient supplier or responsive supplier? an analysis of sourcing strategies under competition. Working Paper, Olin Business School, Washington University, St. Louis.

Wu, Xiaole, Panos Kouvelis, Hirofumi Matsuo. 2011. Horizontal capacity coordination for risk management and flexibility: Pay ex-ante or commit a fraction of ex-post demand? Working Paper, Olin Business School, Washington University at St. Louis.

Zipkin, P. H. 2000. *Foundations of Inventory Management*. McGraw-Hill, New York.