2010/67

Measuring the variability in supply chains with the peakedness

Philippe Chevalier, Jean-Christophe Van den Schrieck and Ying Wei

CORE

DISCUSSION PAPER

Center for Operations Research
and Econometrics

Voie du Roman Pays, 34
B-1348 Louvain-la-Neuve
Belgium
http://www.uclouvain.be/core

CORE DISCUSSION PAPER
2010/67

Measuring the variability in supply chains with the peakedness

Philippe CHEVALIER [1], Jean-Christophe VAN DEN SCHRIECK[2] and Ying WEI [3]

October 2010

Abstract

This paper introduces a novel way to measure the variability of order flows in supply chains, the peakedness. The peakedness can be used to measure the variability assuming the order flow is a general point pro- cess. We show basic properties of the peakedness, and demonstrate its computation from real-time continuous demand processes, and cumulative demand collected at fixed time intervals as well. We also show that the peakedness can be used to characterize demand, forecast, and inventory variables, to effectively manage the variability. Our results hold for both single stage and multistage inventory systems, and can further be extended to a tree-structured supply chain with a single supplier and multiple retailers. Furthermore, the peakedness can be applied to study traditional inventory problems such as quantifying bullwhip effects and determining safety stock levels. Finally, a numerical study based on real life Belgian supermarket data verifies the effectiveness of the peakedness for measuring the order flow variability, as well as estimating the bullwhip effects.

Keywords: variability, peakedness, supply chain.

---

[1] Université catholique de Louvain, CORE and Louvain School of Management, B-1348 Louvain-la-Neuve, Belgium. E-mail: philippe.chevalier@uclouvain.be. This author is also member of ECORE, the association between CORE and ECARES.

[2] Université catholique de Louvain, CORE and Louvain School of Management, B-1348 Louvain-la-Neuve, Belgium. E-mail: jc.vandenschrieck@uclouvain.be

[3] Université catholique de Louvain, CORE, B-1348 Louvain-la-Neuve, Belgium.
E-mail: ying-wei_2008@hotmail.com

The authors would like to gratefully acknowledge the useful suggestions and comments of Jean-François Macq. This research was partially funded by the Walloon region with the grant FIRST – Post Doc "OPTIFLUX".

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 Introduction

In a supply chain for make-to-stock items (such as most consumer goods) the root of variability lies in the variability of customer demand. Difficulties arise in how to measure this variability and how to represent this variability in a model to help understand its propagation and improve the management of the supply chain. Taking the demand in a grocery store for example, it often exhibits different types of seasonalities according to the period of the day, the day of the week, the day of the month or the month of the year. Measuring the variability of demand flows with such patterns is not easy. Trying to predict how the variability of such flows might impact the supply chain is even harder.

Indeed, variability is a major issue for supply chain management. Order flow variability is responsible for an important portion of the inventory in the supply chain. It is also a complicated issue in the sense that on one hand, order flow variability is amplified up along the supply chain which is notoriously known as the bullwhip effect (Lee et al. (1997)); on the other hand, as explained by Cachon (1999), the variability of flows is reduced with aggregated flows; i.e., in a multiechelon system with N retailers and one wholesaler, balanced ordering will reduce the supplier's demand variance.

This paper seeks to address the issue of measuring the variability in supply chains, especially in fast moving consumer goods supply chains. We first take a look at how to model the underlying order flows. A traditional approach to model order flows in a supply chain is to use a renewal process assuming that the inter-arrival times of the order flow are i.i.d., and follow a certain distribution. As a special form of the renewal process, the Poisson process is perhaps the most widely used stochastic process in operations management because of its theoretical and computational simplicity. Noticing that in practice the arrival rate of a Poisson process is not necessarily constant, Cox (1955) proposes to represent the arrival rate function by a stationary stochastic process, the resulting arrival process is known as a doubly stochastic Poisson process.

Working with renewal processes, however, poses some problems. First it implies making rather detailed assumptions about the interarrival distributions, which most of the time are hard to characterize. An important difficulty in practice is the fact that the arrival process will most often exhibit different seasonality patterns at different time scales. – Looking at a supermarket for example, there are clear seasonality patterns in the day, the week, the month, the year... the combination being very hard to include in a distribution for interarrival times – Second, the renewal process assumption is not compatible with the observation that in practice the variability of flows does not diminish as much as one would expect with aggregated flows. Although a doubly stochastic process can explain the aggregated order variability, it imposes more assumptions on the process, that makes it less tractable for implementation.

Another approach to characterize the demand flows in a supply chain is using a time series model. For example, the Autoregressive Integrated Moving Average (ARIMA) process, which uses historical data and forecasts demand based on correlations between consecutive demand realizations (Lee et al. (1997), Graves (1999), and Gilbert (2005)). Or more generally, vector autoregressive time series (Aviv (2003)). However, the autoregressive process, even though popular in the literature, requires to estimate complex parameters. Here again the seasonalities make the estimation of the parameters very difficult.

This paper proposes a new measurement called the *peakedness*, which has been used for a long time in the telecommunication literature to measure the variability of flows in networks (see (Wilkinson, 1956) for a seminal paper). The peakedness is calculated based on the number of busy servers in a queue with infinite servers, namely the ratio of its variance over its mean value. As demonstrated in this paper, the peakedness is easy to compute and practical to apply in real life supply chains. Moreover, it is amenable to closed form formulas for order variability control and system optimization.

The peakedness functional provides a new approach for measuring order variability without relying on detailed assumptions of the underlying process. It does not require the interarrival times of the order flows to be i.i.d as well. Simply assuming that the order flow follows a point process with a known long-run average arrival rate, the peakedness can be easily computed from real life data based on a sufficiently large sample of observations. We also show in this paper how to estimate the peakedness from real life demand processes.

The peakedness can also be used for order variability control in a supply chain. Taking similar modeling hypotheses as in Graves (1999), we are able to characterize the demand, forecast and inventory variables based on the peakedness functional. Our results hold for single stage inventory systems and multistage inventory systems, and can further be extended to tree-structured supply chains with a single supplier and multiple retailers. Furthermore, the peakedness can be applied to study traditional supply chain questions such as quantifying bullwhip effects, and determining inventory safety stock levels.

The rest of the paper is organized as follows. We first give a review of the existing literature in section 2. We introduce the peakedness in section 3 and present some important properties for our context. We develop our single-stage inventory system model, as well as the extensions to the multi-stage inventory model and to the supply chain network in section 4 . In section 5, we compute the peakedness based on the real life data of a Belgium supermarket chain, and compare the estimated order flows variability with simulation results and that of Graves (1999). Finally, we conclude the paper with practical implications of this work.

2 Related work

There is a substantial body of literature concerning the variability in supply chains. In terms of modeling the order flows, several approaches have been used. See, for example, Browne and Zipkin (1991), Sivazlian (1974), Sahin (1979), Sahin (1982) for renewal processes, and Hadley and Whitin (1963) for Poisson process. For doubly stochastic Poisson processes, the reader is referred to Cox (1955) and Ozekici and Soyer (2006). See also Song and Zipkin (1993) for modeling the order process as a Markov chain.

Another popular approach for modeling order flows in supply chains is a time series based model, notably the Autoregressive Integrated Moving Average (ARIMA) process. Assuming a nonstationary ARIMA (0,1,1) demand process, Graves (1999) shows that the order process of the agent is an ARIMA(0,1,1) process as well, if employing an adjusted base-stock policy, which replenishes the inventory and adjusts the base stock level to the new forecasts. He further characterizes the inventory random variable and uses it to find the safety stock requirement. Gilbert (2005) generalizes the results for any ARIMA(p,d,q)-type demand series. He shows that the orders and inventories at all stages can also be modeled as ARIMA time series and provides their closed form expressions for the analysis of the bullwhip effect.

The bullwhip effect and its consequences are widely discussed in the literature, mostly using the time-series model. In the economics area, Kahn (1987) is the first to show the bullwhip effect for a firm that faces uncertain demand modeled by an AR(1) process. The author observes that the variability of the demand has a double effect: first the variability is reproduced in the production line through the replenishment of the stock and second the variability is amplified due to the adjustment of the inventory level itself in order to reflect changes in the forecasts for the future periods.

Extending Kahn's work on AR(1) model, Lee et al. (1997) identify causes for the bullwhip effect and analyze the implication on managerial decisions, such as demand signal, order batching, rationing game, and price variation. They also demonstrate that the supplier's demand variance is minimized when the retailers' orders are balanced, i.e., the same number of retailers order each period. Assuming balanced orders, Cachon (1999) studies the management of supply chain demand variability in a model with one supplier and N retailers by "scheduled ordering". He shows that the supplier's demand variance is further reduced when the retailer order intervals are lengthened or when the retailer's batch size is reduced. Chen et al. (2000) show the impact of forecasting and lead-times on the variability. They also demonstrate that variability amplification is reduced but not eliminated by information sharing.

Croson and Donohue (2006) find that the bullwhip effect still exists when all major drivers of variability amplification are removed. Based on experimental results of the beer game, they emphasize the impact of the behavior of the agents on the supply chain such as the misperception of change in demand. Studying the bullwhip effect in industry level, Cachon et al. (2007) find that wholesale industries exhibit a bullwhip effect, but retail industries generally do not, nor do most manufacturing industries. In addition, they point out that industries with highly seasonal demand tend to smooth the variability amplification. A recent paper of Chen and Lee (2009) indicates that advanced demand information such as projected future orders of the retailer is also an effective way to pass demand information and reduce the order variability along the supply chain.

The impact of the variability has been studied by a number of papers, focusing on how to set strategic safety stock levels in complex supply chain networks. Assuming demand is stationary, Graves and Willems (2000) propose an algorithm optimizing the strategic safety stock placement in supply chains. Graves and Willems (2008) further extend their previous work to nonstationary demand process. Both of these work are following the so-called "constant service time" policy.

In this paper, we follow the similar model framework and ordering policy as in Graves (1999), but we take a broader perspective. We contribute to the literature in the following senses: (1) Rather than assuming that the demand follows an ARIMA process with some known parameters, we model the order flows in a supply chain by a general point process, which makes the renewal process to be a specific form. (2) Based on the general order flow process, we propose an easy to calculate but effective variability measurement, the peakedness. (3) Last but not least, our model can handle the propagation of variability in supply chains of general structure.

3 The peakedness functional

In this section, we introduce a new measure of flow variability in supply chains, the peakedness. We first model order flows of supply chains by a point process. We then introduce the concept of peakedness. We pursue by presenting some properties that enable us to estimate the peakedness from different forms of arrival data in practice.

3.1 A general arrival process model

We model the order flows of supply chains by a general arrival process: *the point process*. A point process is used to describe a simple arrival process with single arrivals, and is mathematically defined by

$$X(t) = \sum_{i=0}^{+\infty} \delta(t - T_i),$$

where $\{T_i\}$, $i = 0, 1, ...$, constitutes the sequence of arrival times, and $\delta(\cdot)$ is the Dirac delta function (Macq (2005), Daley and Vere-Jones (2002)). $X(t)$ is thus equal to one when an arrival occurs and to zero otherwise. By convention, $T_0 = 0$. The corresponding counting process of $X(t)$ is

$$N(t) = \sum_{i=0}^{\infty} u(t - T_i),$$

where $u(t)$ is the unit step function:

$$u(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{otherwise.} \end{cases}$$

We assume that the point process $X(t)$ is stationary and give its definition from Daley and Vere-Jones (2002).

Definition 3.1. *A point process is stationary when for every $r = 1, 2, ...$ and all bounded Borel subsets $A_1, ..., A_r$ of the real line, the joint distribution of*

$$\{N(A_1 + t), ..., N(A_r + t)\}$$

does not depend on t, $(-\infty < t < \infty)$.

Here the sets A_i are Borel sets of $(0, \infty)$ and we require $t > 0$. Intuitively, by stationarity, it means that the distribution of number of arrivals does not depend on time t. Alternatively, the *arrival rate* of $X(t)$, if denoted by λ, $0 < \lambda < \infty$, is a constant and given by $\lambda = E[X(t)]$.

A point process is a general arrival process for which it is assumed that the inter-arrival times are random variables. Unlike renewal processes, it does not require the inter-arrival times to follow a certain distribution. As a result, a renewal process is a special point process.

3.2 A general variability measure: the Peakedness

Suppose that all arrivals of $X(t)$ go to a queue with an infinite number of servers. Their service time distribution is denoted by G. We observe $S(t)$, the number of busy servers, and define the *peakedness* of the arrival flow $X(t)$ as

$$z(X, \mathsf{G}) = \frac{\text{Var}[S(t)]}{E[S(t)]}. \tag{1}$$

Specifically, when G is an exponential distribution with service rate s, $0 < s < \infty$, we denote the peakedness by $z(X, \mathsf{M}(s))$; and $z(X, \mathsf{D}(s))$ when G is deterministic with service rate s. Figure 1 shows the fictitious infinite server pool that is used to evaluate the peakedness. We refer the interested reader to Jagerman et al. (1997) or Wolff (1989) for a more extensive introduction to the peakedness.

The peakedness measures the total variability of the order flow. To illustrate this, consider two deterministic arrival processes X_1 and X_2. For X_1, one order arrives every period, and for X_2, 10 orders arrive every ten periods, with no order arrivals in-between. As a result, the arrival rate for both deterministic arrival process are the same, namely 1. Suppose service time is constant at one period. For X_1, the number of busy servers $S(t) = \{1, 1, ..., 1\}$ and the peakedness is obviously 0; while for X_2, $S(t) = \{10, 0, ...; 10, 0, ...\}$, the peakedness is larger than zero. Therefore, the peakedness can differentiate these two deterministic arrival processes.

Property 1. *If $X(t)$ is a Poisson process, its peakedness is 1.*

This property is a direct consequence of the fact that the number of busy servers is geometrically distributed for an $M/G/\infty$ system whatever the service time distribution. This is handy, as this gives an easy benchmark for the variability of any process. Moreover this benchmark coincides with the more familiar coefficient of variation.

In terms of implementation, the estimation of the peakedness of an arrival flow poses some problems. The first problem is that the definition

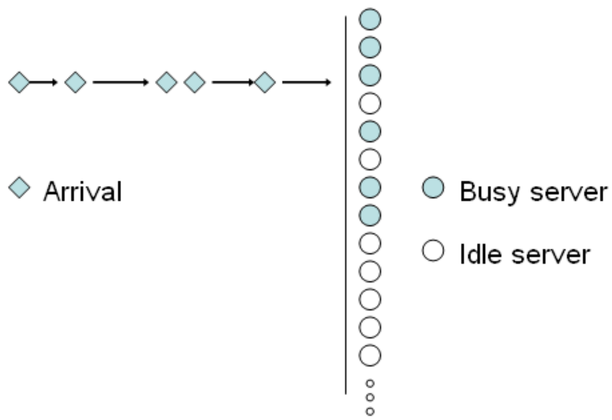

Figure 1: A schematic representation of the arrival stream and the number of busy servers

involves observing the behavior of an infinite server system with random service times, these service times will introduce some randomness in the estimation. The second problem is that in order to accurately simulate the infinite server system we should keep all individual arrival times, this information can quickly take a lot of storage space. In the next section we show how we can eliminate the first problem by using the so called fluid peakedness. In the subsequent subsections we show how it is possible to work with simplified data.

3.3 The fluid peakedness

It is problematic that two sources of variability are present in the calculation of $z(X, \mathsf{G})$: one from the demand arrival process $X(t)$, and another from the service time uncertainty G. In order to focus on the variability of the demand process, we propose a new measure: the *fluid peakedness*, based on a "fluidity" assumption of the service process. Each infinitesimal amount of demand is independently served by its own infinitesimal server according to the given service time distribution G. For a single arrival at time t, the number of servers still answering this arrival at time $t' > t$ would in this case be $F_{\mathsf{G}}(t' - t) = P[\mathsf{G} > t' - t]$. The fluid peakedness is then calculated

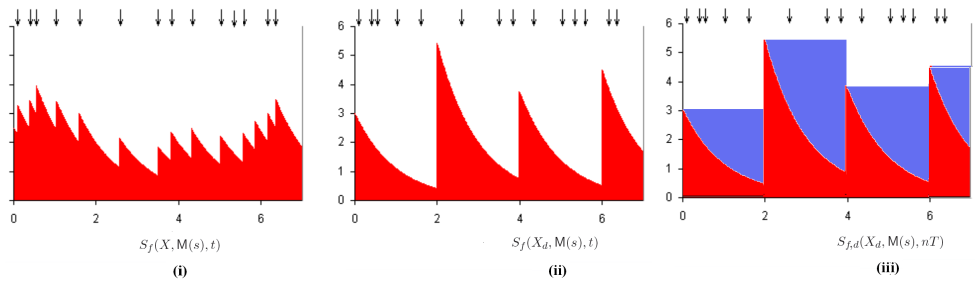

Figure 2: Illustration of the difference among $S_f(X, \mathsf{M}(s), t)$ (i), $S_f(X_d, \mathsf{M}(s), t)$ (ii), and $S_{f,d}(X_d, \mathsf{M}(s), nT)$ (iii) ($T = 2$).

by

$$z_f(X, \mathsf{G}) = \frac{\mathrm{Var}[S_f(X, \mathsf{G}, t)]}{E[S_f(X, \mathsf{G}, t)]}, \tag{2}$$

with $S_f(X, \mathsf{G}, t)$ the amount of demand still being served at any time t, given the service distribution G.

With this definition, all the randomness in $S_f(X, \mathsf{G}, t)$ comes from the arrival flow and not from the service times. When the service time distribution G is exponential with service rate s, $S_f(X, \mathsf{M}(s), t)$ can be seen as a process that is subject to an exponential decay and increases by one each time a new demand arrives, i.e., $S_f(X, \mathsf{M}(s), t) = \sum_{T_i \leq t} e^{-s(t-T_i)}$. Figure 2 (i) gives an illustration of $S_f(X, \mathsf{M}(s), t)$, with each arrow representing a demand arrival.

A more general discussion about the relationships of the peakedness and the fluid peakedness is provided in Appendix A. In particular, we have the following:

Property 2. *For any point process X, $z(X, \mathsf{M}(s)) = z_f(X, \mathsf{M}(s)) + \frac{1}{2}$.*

A sketch of the proof is presented in Appendix A. The interested reader may find a more detailed proof in Macq (2005).

Property 3. *For any point process X, $z(X, \mathsf{D}(s)) = z_f(X, \mathsf{D}(s))$.*

This property follows directly from the fact that a deterministic service distribution contains no randomness.

Hereafter in this paper, when we say peakedness, we mean fluid peakedness.

3.4 Estimating the peakedness from batch arrival data

Notice that in practice the demand arrival process $X(t)$ is often difficult to observe continuously for every $t \geq 0$. One might only have access to information about the cumulated demand over a time interval such as an hour or a day. We here show how to use cumulated data to compute a peakedness.

Denote by T the length of the time interval, and by D_n the demand observed during the n^{th} time period, i.e., $D_n = N(nT) - N((n-1)T)$, $n \in \mathbb{Z}^+$. $X_d = \{D_1, D_2, ...\}$ thus constitutes a discrete-time batch arrival process with batch size D_n, $n = 1, 2, ...$, and interval arrival time length T. In other words, X_d is obtained by aggregating the arrivals of $X(t)$ over periods of length T.

The number of busy servers of the discrete-time arrival process for an exponential service time distribution $\mathsf{M}(s)$, if denoted by $S_f(X_d, \mathsf{M}(s), t)$, is then computed recursively by

$$\begin{aligned} S_f(X_d, \mathsf{M}(s), t) &= S_f(X_d, \mathsf{M}(s), (n-1)T)e^{-s[t-(n-1)T]}, \qquad (3)\\ &\qquad (n-1)T < t < nT \quad \forall n \in \mathbb{N}_0 \\ S_f(X_d, \mathsf{M}(s), nT) &= S_f(X_d, \mathsf{M}(s), (n-1)T)e^{-sT} + D_n, \quad \forall n \in \mathbb{N}_0 \qquad (4)\\ S_f(X_d, \mathsf{M}(s), 0) &= 0. \qquad (5) \end{aligned}$$

Figure 2 (ii) shows $S_f(X_d, \mathsf{M}(s), t)$ when $T = 2$.

As arrivals only occur at epochs nT, $n \in \mathbb{N}_0$, one would likely use the values of $S_f(X_d, M(s), t)$ at those times as sample values to compute the peakedness. In other words one would only consider the set of measurements $\{S_f(X_d, M(s), T), S_f(X_d, M(s), 2T), ...\}$. With this set of measurements, we compute the peakedness of the process presented in Figure 2 (iii), noted $S_{f,d}(X_d, \mathsf{M}(s), t)$ and which differs from process $S_f(X_d, \mathsf{M}(s), t)$. Property 4 explains how to adjust this peakedness to obtain an unbiased peakedness for process X_d.

Property 4. *Let $z_{f,d}(X_d, \mathsf{M}(s))$ be the ratio of the variance over the mean of the elements in $\{S_f(X_d, M(s), T), S_f(X_d, M(s), 2T), ...\}$. Then,*

$$z_f(X_d, M(s)) = z_{f,d}(X_d, M(s)) \frac{1+\alpha}{2} + \frac{\lambda T}{2} \frac{1+\alpha}{1-\alpha} - \frac{\lambda T}{sT}, \qquad (6)$$

where $\alpha = e^{-sT}$.

Proof. For batch arrival data, the arrivals only occur at nT, $n \in Z^+$. Given an exponential service distribution $\mathsf{M(s)}$, $S_f(X_d, \mathsf{M}(s), t)$ can be seen as a process that is fed at every arrival time and that is subject to an exponential decay between arrival times. The value of $S_f(X_d, \mathsf{M}(s), t)$ can therefore been calculated from $S_f(X_d, \mathsf{M}(s), nT)$, for every $t \in [nT, (n+1)T)$. In other words for any value of $\tau \in (0; T]$,

$$S_f(X_d, \mathsf{M}(s), nT + \tau) = S_f(X_d, \mathsf{M}(s), nT)e^{-s\tau} \tag{7}$$

Therefore, we have $E[S_f(X_d, \mathsf{M}(s), nT+\tau)] = e^{-s\tau}E[S_f(X_d, \mathsf{M}(s), nT)]$ and $E[S_f(X_d, \mathsf{M}(s), nT+\tau)^2] = e^{-2s\tau}E[S_f(X_d, \mathsf{M}(s), nT)^2]$. Consequently, we can write:

$$\begin{aligned} E[S_f(X_d, \mathsf{M}(s), t)] &= \frac{1}{T}\int_0^T E[S_f(X_d, \mathsf{M}(s), nT)]e^{-s\tau}d\tau \\ &= \frac{1-\alpha}{sT}E[S_f(X_d, \mathsf{M}(s), nT)], \end{aligned} \tag{8}$$

$$\begin{aligned} E[S_f(X_d, \mathsf{M}(s), t)^2] &= \frac{1}{T}\int_0^T E[S_f(X_d, \mathsf{M}(s), nT)^2]e^{-2s\tau}d\tau \\ &= \frac{1-\alpha^2}{2sT}E[S_f(X_d, \mathsf{M}(s), nT)^2], \end{aligned} \tag{9}$$

where $\alpha = e^{-sT}$. Combining these two expressions together, we have:

$$\begin{aligned} \mathrm{Var}[S_f(X_d, \mathsf{M}(s), t)] &= \frac{1-\alpha^2}{2sT}E[S_f(X_d, \mathsf{M}(s), nT)^2] \\ &\quad -\frac{(1-\alpha)^2}{(sT)^2}E^2[S_f(X_d, \mathsf{M}(s), nT)] \\ &= \frac{1-\alpha^2}{2sT}\big(\mathrm{Var}[S_f(X_d, \mathsf{M}(s), nT)] \\ &\quad +E^2[S_f(X_d, \mathsf{M}(s), nT)]\big) \\ &\quad -\frac{(1-\alpha)^2}{(sT)^2}E^2[S_f(X_d, \mathsf{M}(s), nT)] \end{aligned} \tag{10}$$

We easily deduce from (4) and from the definition of D_n that

$$E[S_f(X_d, \mathsf{M}(s), nT)] = \frac{E[D_n]}{1-\alpha} \tag{11}$$

$$= \frac{\lambda T}{1-\alpha}. \tag{12}$$

Dividing (10) by (12) eventually yields (6). □

Furthermore it is possible to find an approximation of $z_f(X, M(s))$, i.e. the peakedness associated to the point process X based on $z_{f,d}(X_d, \mathsf{M}(s))$. The approximation relies on the following property.

Property 5. *Let $z_{f,d}(X, M(s))$ be the peakedness computed with the sample values in $\{S_f(X, M(s), T), S_f(X, M(s), 2T), ...\}$, i.e. the values of $S_f(X, M(s), t)$ at the time of the batch arrivals in X_d. Then $z_{f,d}(X, M(s))$ can be approximated as*

$$z_{f,d}(X, M(s)) \simeq \frac{1}{2} - \frac{1}{sT} \cdot \frac{1-\alpha}{1+\alpha} + z_{f,d}(X_d, M(s)) \left(\frac{1-\alpha}{sT} \right).$$

Proof. $S_f(X, M(s), nT)$ can be computed with the following equation

$$S_f(X, M(s), nT) = S_f(X, M(s), (n-1)T)\alpha + \sum_{i=1}^{D_n} e^{-s(nT-t_i)} \tag{13}$$

where t_i, $(n-1)T < t_i \leq nT$ is the time of arrival i. Similarly, $S_f(X_d, M(s), nT)$ is easily computed in the following way:

$$S_f(X_d, M(s), nT) = S_f(X_d, M(s), (n-1)T)\alpha + D_n. \tag{14}$$

The difference between $S_f(X, M(s), nT)$ and $S_f(X_d, M(s), nT)$ is that the service of arrival i in process X is delayed by a time $T - t_i$ in $S_f(X_d, M(s), nT)$. Therefore, upon the beginning of its service in $S_f(X_d, M(s), nT)$, the number of remaining associated busy servers in $S_f(X, M(s), nT)$ is equal to $e^{-s(nT-t_i)}$ whereas it is equal to 1 in $S_f(X_d, M(s), nT)$. The time difference $\tau_i = nT - t_i$ is a random variable which we will suppose to be uniformly distributed over the interval [0,T].

(12) gives $E[S_f(X_d, M(s), nT)]$. The related variance is equal to

$$\begin{aligned} \mathrm{Var}[S_f(X_d, M(s), nT)] &= \mathrm{Var}[S_f(X_d, M(s), nT)]\alpha^2 \\ &\quad + \mathrm{Var}[D_n] + 2\,\mathrm{Cov}[S_f(X_d, M(s), nT), D_n]\alpha \\ &= \frac{\mathrm{Var}[D_n] + 2\,\mathrm{Cov}[S_f(X_d, M(s), nT), D_n]\alpha}{1-\alpha^2}. \end{aligned}$$

We therefore find

$$z_{f,d}(X_d, M(s)) = \frac{\mathrm{Var}[D_n] + 2\,\mathrm{Cov}[S_f(X_d, M(s), nT), D_n]\alpha}{E[D_n](1+\alpha)}.$$

Similarly, for $S_f(X, M(s), nT)$, we deduce from (13)

$$\begin{aligned} E[S_f(X, M(s), nT)] &= \frac{E[\sum_{i=1}^{D_n} e^{-\mu\tau_i}]}{1-\alpha} \\ &= \frac{E[D_n]E[e^{-\mu\tau}]}{1-\alpha}. \end{aligned}$$

and

$$\begin{aligned} &\mathrm{Var}[S_f(X, M(s), nT)] \\ &\quad = \mathrm{Var}[S_f(X, M(s), nT)]\alpha^2 + \mathrm{Var}[\sum_{i=1}^{D_n} e^{-\mu\tau_i}] \\ &\quad\quad +2\,\mathrm{Cov}[S_f(X, M(s), nT), \sum_{i=1}^{D_n} e^{-\mu\tau_i}]\alpha \\ &\quad = \frac{\mathrm{Var}[\sum_{i=1}^{D_n} e^{-\mu\tau_i}] + 2\,\mathrm{Cov}[S_f(X, M(s), nT), \sum_{i=1}^{D_n} e^{-\mu\tau_i}]\alpha}{1-\alpha^2} \end{aligned}$$

Applying the law of total variance, we easily solve $\mathrm{Var}[\sum_{i=1}^{D_n} e^{-\mu\tau_i}]$:

$$\mathrm{Var}[\sum_{i=1}^{D_n} e^{-\mu\tau_i}] = E[\mathrm{Var}[\sum_{i=1}^{D_n} e^{-\mu\tau_i}|B_n]] + \mathrm{Var}[E[\sum_{i=1}^{D_n} e^{-\mu\tau_i}|D_n]].$$

If we assume that all τ_i's are independent, this expression becomes

$$\mathrm{Var}[\sum_{i=1}^{D_n} e^{-\mu\tau_i}] = \mathrm{Var}[e^{-\mu\tau}]E[D_n] + E^2[e^{-\mu\tau}]\mathrm{Var}[D_n].$$

As the service of any arrival is delayed by a time uniformly distributed over $[0, T]$ in $S_f(X_d, M(s), nT)$, the amount of fluid operators answering any arrival in $S_f(X, M(s), nT)$ is on average smaller by a factor $E[e^{-\mu\tau}]$ than the amount of operators in $S_f(X_d, M(s), nT)$. We therefore assume that $\mathrm{Cov}[S_f(X, M(s), nT), \sum_{i=1}^{D_n} e^{-\mu\tau_i}] = \mathrm{Cov}[S_f(X_d, M(s), nT), D_n]E^2[e^{-\mu\tau}]$. Therefore,

$$\begin{aligned} \mathrm{Var}[S_f(X, M(s), nT)] \;\simeq\; & \frac{\mathrm{Var}[e^{-\mu\tau}]E[D_n] + E^2[e^{-\mu\tau}]\mathrm{Var}[D_n]}{1-\alpha^2} \\ & + \frac{2\mathrm{Cov}[S_f(X_d, M(s), nT), D_n]E^2[e^{-\mu\tau}]\alpha}{1-\alpha^2}. \end{aligned}$$

We obtain our approximation for $z_{f,d}(X, M(s))$ by dividing this last expression by $E[S_f(X, M(s), nT)]$:

$$\begin{aligned} z_{f,d}(X, M(s)) &\simeq \frac{\mathrm{Var}[e^{-\mu\tau}]}{E[e^{-\mu\tau}](1+\alpha)} \\ &+ E[e^{-\mu\tau}]\left(\frac{\mathrm{Var}[D_n] + 2\mathrm{Cov}[S_f(X_d, M(s), nT), D_n]e^{-\mu T}}{E[D_n](1+e^{-\mu T})}\right) \\ &\simeq \frac{\mathrm{Var}[e^{-\mu\tau}]}{E[e^{-\mu\tau}](1+\alpha)} + E[e^{-\mu\tau}]\frac{\mathrm{Var}[S_f(X_d, M(s), nT)]}{E[S_f(X_d, M(s), nT)]}. \end{aligned}$$

As,

$$\begin{aligned} E[e^{-\mu\tau}] &= \frac{1}{T}\int_0^T e^{-\mu t} dt \\ &= \frac{1 - e^{-\mu T}}{\mu T} \\ \mathrm{Var}[e^{-\mu\tau}] &= \frac{1}{T}\int_0^T e^{-2\mu t} dt - \left(\frac{1 - e^{-\mu T}}{\mu T}\right)^2 \\ &= \frac{1 - e^{-2\mu T}}{2\mu T} - \left(\frac{1 - e^{-\mu T}}{\mu T}\right)^2, \end{aligned}$$

we can write

$$z_{f,d}(X, M(s)) \simeq \frac{1}{2} - \frac{1}{sT} \cdot \frac{1-\alpha}{1+\alpha} + z_{f,d}(X_d, M(s))\left(\frac{1-\alpha}{sT}\right).$$

□

If we make the additional assumtion that $z_{f,d}(X, M(s)) = z_f(X, M(s))$, i.e. that taking sample measurements of $S_f(X, M(s), t)$ at times nT does not introduce a bias, we obtain an estimate for $z_f(X, M(s))$. this last assumption essentially means that the arrival process is sufficiently ergodic that in the long run the number of busy virtual servers at the times nT has the same distribution as at any random time. Simulation experiments confirmed the quality of this approximation for different types of demand processes.

4 A peakedness based model

In this section we model the variability of flows in a supply chain based on the peakedness, and investigate how to manage the variability using the peakedness functional. We consider a similar model as in Graves (1999), and assume that all participants use exponential smoothing to make their forecasts and order according to an adjusted base stock type policy. Rather than assuming that the demand follows an ARIMA process with some known parameters, we simply suppose that the demand is a point process and that the long term demand arrival rate λ and its peakedness are known.

4.1 A single-stage inventory system

We start with an agent selling a single product and facing a single-stage inventory problem in a periodic review system, with T the time length for each review period. We assume that the order replenishment lead-time is fixed and known as HT. In other words H is the number of periods included in the lead-time. In each period n, $n \in \mathbb{Z}^+$, the agent receives the quantity ordered H periods ago, observes the demand for the current period D_n, fulfills demand from inventory as much as possible and backlogs unsatisfied demand; and then he makes a forecast F_n for the demand in the next periods and finally places an order for a quantity O_n. The demand process, the forecast model, and the inventory control variables in each period n are as follows.

- **Demand process**: We model demand as a point process with long-term arrival rate λ, and observed at discrete-time intervals $\{T, 2T, ...\}$. The aggregated demand process is thus a discrete batch arrival process X_d constituted by $\{D_1, D_2, ...\}$, and the average number of arrivals per time interval is equal to $E[D_n] = \lambda T$, $n \in Z^+$. As mentioned earlier, the peakedness can be calculated based on the aggregated demand arrival process.

- **Forecast model**: After observing the demand D_n in period n, the agent makes a forecast F_n for the demand in period $n+1$. The forecasts are made based on an exponential smoothing technique with parameter α and initial forecast λT:

$$\begin{aligned} F_0 &= \lambda T \\ F_n &= (1-\alpha)D_n + \alpha F_{n-1}, \end{aligned} \tag{15}$$

  with $\alpha = e^{-sT}$, and $n = 1, 2, ...$ It has been shown in the literature that for an ARIMA(0,1,1) demand process, the exponential-weighted moving average provides the best forecast (Graves (1999)). In this paper, given that the demand process is a point process, we still use the exponential smoothing technique considering its theoretical and practical popularity. We suppose that the smoothing factor s is chosen such as to minimize the mean square forecasting error. Intuitively, a small smoothing factor s means that more weight is given to past observations, or, equivalently, that a longer period is taken into account in the forecast.

- **Inventory control variables**: In each period, the agent receives the order issued H periods ago, and places a new order to its supplier. We write the inventory balance equation as follows:

$$I_n = I_{n-1} - D_n + O_{n-H}, \tag{16}$$

  where I_n is the ending inventory level of period n, and O_n is the order quantity placed at period n, $n = 1, 2, ...$

  We adopt an adjusted base-stock policy for the order placement, which is widely used in the literature. See, for example, Kahn (1987) and Graves (1999). The base-stock policy is adjusted as the demand forecast changes over time, and indicated by.

$$O_n = D_n + H(F_n - F_{n-1}) \tag{17}$$

  The variability of the orders placed by the agent is thus twofold. It comes first from the variability of the demand observed, and second from the variability of adjusting the forecasts.

Substituting (17) to (16) yields:

$$I_n = I_{n-1} - D_n + D_{n-H} + H(F_{n-H} - F_{n-H-1}) \tag{18}$$

By repeated backward substitution, we can then rewrite (16) as

$$I_n = I_0 - D_n - ... - D_{n-H+1} + HF_{n-H}, \tag{19}$$

with I_0 the initial inventory level set at the beginning of the planning horizon.

4.1.1 Single stage peakedness model

We first model the variability of the order flow using the peakedness. Note that the number of busy servers for the peakedness computed with a deterministic service time of T would correspond to the number of arrivals during a period of T. Consequently we have

$$\begin{aligned} D_n &= S_f(X_d, \mathsf{D}(1/T), nT) & (20)\\ \mathrm{Var}[D_n] &= \lambda T z_{f,d}(X_d, \mathsf{D}(1/T)), & (21) \end{aligned}$$

where $S_{f,d}((X_d, \mathsf{D}(1/T), nT)$ and $z_f(X_d, \mathsf{D}(1/T), nT)$, are the number of busy servers and peakedness defined previously. (21) is obtained based on the definition of fluid peakedness, and the expected value λT of the demand in each period.

Analogously, we can measure the forecast and its variance by the peakedness functional, as indicated in Lemma 4.1. See technical proofs in Appendix B.

Lemma 4.1. *The forecast F_n and its variance can be computed by*

$$\begin{aligned} F_n &= (1-\alpha) S_f(X_d, \mathsf{M}(s), nT) & (22)\\ Var[F_n] &= (1-\alpha)\lambda T z_{f,d}(X_d, \mathsf{M}(s)) & (23) \end{aligned}$$

with $\alpha = e^{-sT}$ and where $S_f(X_d, \mathsf{M}(s), nT)$ and $z_{f,d}(X_d, \mathsf{M}(s))$ are the number of busy servers and peakedness defined previously.

We continue to characterize the inventory control variables by the peakedness in Lemma 4.2.

Lemma 4.2. *If ignoring the autocovariance effect of demands across periods, i.e., $Cov[D_n, D_{n-i}] = 0$, for $i = 1, 2, ...$, the order quantity O_n and its variance can be computed by*

$$\begin{aligned} O_n &= S_f(X_d, \mathsf{D}(1/T), nT) + H(1-\alpha)[S_f(X_d, \mathsf{M}(s), nT) \\ &\quad - S_f(X_d, \mathsf{M}(s), (n-1)T)] & (24)\\ Var[O_n] &= \lambda T z_{f,d}(X_d, \mathsf{D}(1/T))(1 + 2H(1-\alpha)) \\ &\quad + 2(1-\alpha)^2 H^2 \lambda T z_{f,d}(X_d, \mathsf{M}(s)), & (25) \end{aligned}$$

if employing a base-stock policy.

The proof is given in Appendix B.2. Ignoring the autocovariance effect is for analytical convenience and popular in the literature (see e.g., Graves (1999) and Gilbert (2005)). Observing (25) we can see that the first item indicates the variability propagation of the downstream node demand, and the second item indicates the variability propagation due to the forecast.

By substituting (20) and (22) into equation (19), we find that the inventory level I_n and its variance are equal to

$$\begin{aligned} I_n &= I_0 - \sum_{i=0}^{H-1} S_f(X_d, \mathsf{D}(1/T), (n-i)T) \\ &\quad + H(1-\alpha) S_f(X_d, \mathsf{M}(s), (n-H)T) && (26) \\ \text{Var}[I_n] &= H\text{Var}[D_n] + H^2 \text{Var}[F_n] \\ &= H\lambda T z_{f,d}(X_d, \mathsf{D}(1/T)) + H^2(1-\alpha)\lambda T z_{f,d}(X_d, \mathsf{M}(s)) && (27) \end{aligned}$$

4.1.2 Extended discussion

Suppose the initial inventory level I_0 is the control variable determined at the beginning of the planning period, as the system safety stock to assure some service level. If following the well-known assumption that inventory fluctuations follows a Normal distribution, we then have the following proposition

Proposition 4.3. *If inventory fluctuations follow a Normal distribution, the initial inventory level I_0 is then set to be a multiple of $Std[I_n]$, and*

$$I_0 = \xi \sqrt{H\lambda T z_{f,d}(X_d, \mathsf{D}(1/T)) + H^2(1-\alpha)\lambda T z_{f,d}(X_d, \mathsf{M}(s))} \qquad (28)$$

with ξ the critical-fractile of service.

It should be noted that Proposition 4.3 is satisfied by a number of demand models, for example, normally distributed demand models, and demands of ARIMA models with the random noise ϵ_t belonging to $N(0, \sigma^2)$, and mean demand in a period equal to λT.

If we further include a unit holding cost h and a unit penalty cost b to the model, the objective is to decide I_0 to minimize the expected cost in each period such that the probability of not stocking out equals $b/(b+h)$, which is alternatively a newsvendor problem. As indicated in Graves (1999) and Veinott (1965), the adjusted base-stock inventory policy is thus myopically optimal if O_n is allowed to be negative.

Noticing $E[S_f(X_d, \mathsf{D}(1/T), nT)] = \lambda T$, and $\text{Var}[S_f(X_d, \mathsf{D}(1/T), nT)] = \text{Var}[D_n] = \sigma^2$, when $\alpha = 1$, i.e. demands across periods are i.i.d., we find

$$I_0 = \xi \sqrt{H\lambda T z_{f,d}(X_d, \mathsf{D}(1/T))} = \xi\sigma\sqrt{H},$$

which is the same result as in Graves (1999) and Lee et al. (1997). In addition, the marginal safety stock decreases with the increasing lead-time.

On the other hand, when $\alpha < 1$, demands across periods are correlated, the marginal safety stock increases as the lead-time grows.

The impact of the length of the review period is more complicated. On one hand, as the review length increases, the variability increases due to the larger batch size, which equals λT. On the other hand, the variability might decrease because the mean demand tends to be stationary with larger review periods.

4.2 Single-item multiple-stage system

In this section, we extend the previous analysis to a multi-stage inventory system. We start with a two-stage series supply chain with a downstream node and an upstream node. We are interested in estimating how the demand variability is propagated. Results in the previous section apply to the downstream node. We now investigate the upstream node.

The event sequence for each node in each period remains the same as described in the single-stage model. The demand process, forecast and inventory control variables for the upstream node at period n, $n = 1, 2, ...$, are analyzed as follows.

- **Demand process:** The demand process of the upstream node is thus the order flow issued by the downstream node, and denoted by $\{O\}$. $\{O\} = \{O_1, O_2, ..., O_n, ...\}$ with O_n the orders from downstream stage at period n. Noticing $E[F_n] = E[D_n] = \lambda T$, by (17), we then have $E[O_n] = \lambda T$ for $N = 1, 2, ...$ For $n \leq 0$, let $O_n = \lambda T$.

- **Forecast model:** At period n, after the demand at the upstream node O_n is observed, a new forecast $F_n^{(2)}$ for the demand in period $n + 1$ is made based on an exponential smoothing technique,

$$F_n^{(2)} = (1 - \beta)O_n + \beta F_{n-1}^{(2)} \tag{29}$$

  with $\beta = e^{-rT}$, and r is chosen such as to minimize the mean square forecasting error. Let $F_n^{(2)} = \lambda T$ for $n \leq 0$.

- **Inventory control variables:** The ending inventory level at the upstream node at period n, if denoted by $I_n^{(2)}$, is then:

$$I_n^{(2)} = I_{n-1}^{(2)} - O_n + O_{n-L}^{(2)} \tag{30}$$

  where L is the number of periods included in the lead-time for order replenishment at the upstream node, and $O_n^{(2)}$ is the order placed at period n.

Similarly, we assume that in the upstream node an order is placed based on the adjusted base-stock policy, and calculated by

$$O_n^{(2)} = O_n + L(F_n^{(2)} - F_{n-1}^{(2)}) \tag{31}$$

The first item is the demand observed at n-th period, and the second item is the adjusted level for the future L periods.

Let $O_n^{(2)} = \lambda T$ for $n \leq 0$. Substituting (31) to (30) and by backward repetition , we obtain

$$I_n^{(2)} = I_0^{(2)} - O_n - ... - O_{n-L+1} + L \cdot F_{n-L}^{(2)} \tag{32}$$

4.2.1 Multiple stage peakedness model

Noticing that the inventory order quantities issued by the downstream node are the order arrivals to the upstream node, we are now able to characterize the demand and forecast of the upstream node as

$$\begin{aligned} O_n &= S_f(O, \mathsf{D}(1/T), nT) && (33) \\ \mathrm{Var}[O_n] &= \lambda T z_{f,d}(O, \mathsf{D}(1/T)) && (34) \\ F_n^{(2)} &= (1-\beta) S_f(O, \mathsf{M}(r), nT) && (35) \\ \mathrm{Var}[F_n^{(2)}] &= (1-\beta)\lambda T z_{f,d}(O, \mathsf{M}(r)) && (36) \end{aligned}$$

The ending inventory level at the end of period n is then

$$I_n^{(2)} = I_0^{(2)} - \sum_{i=0}^{L-1} S_f(O, \mathsf{D}(1/T), (n-i)T) + L(1-\beta) S_f(O, \mathsf{M}(r), (n-L)T) \tag{37}$$

Lemma 4.4 establishes the variability propagation equations based on the peakedness functional, relating (37) to the peakedness of the downstream node.

Lemma 4.4. *If ignoring the autocovariance effect of demands across periods, i.e., $Cov[D_n, D_{n-i}] = 0$ for $i = 1, 2, ...,$ the peakedness of the orders issued by the downstream node can be calculated by*

$$\begin{aligned} z_{f,d}(O, \mathsf{D}(1/T)) &= (1 + 2H(1-\alpha)) z_{f,d}(X_d, \mathsf{D}(1/T)) \\ &\quad + 2(1-\alpha)^2 H^2 z_{f,d}(X_d, \mathsf{M}(s)) && (38) \end{aligned}$$

with $\alpha = e^{-sT}$*, where* $z_{f,d}(X_d, \mathsf{D}(1/T))$ *and* $z_{f,d}(X_d, \mathsf{M}(s))$ *are defined previously for the downstream demand. In addition,*

$$\begin{aligned} z_{f,d}(O, \mathsf{M}(r)) &= \frac{1}{(1+\beta)}\left[\left(1 + 2H\frac{(1-\alpha)(1-\beta)}{1-\beta\alpha}\right)z_{f,d}(X_d, \mathsf{D}(1/T))\right. \\ &\quad \left. +2H^2\frac{(1-\alpha)^2(1-\beta)}{1-\beta\alpha}z_{f,d}(X_d, \mathsf{M}(s))\right] \end{aligned} \tag{39}$$

with $\beta = e^{-rT}$.

The proof is given in Appendix B.3.

By (38), we see that $z_{f,d}(O, \mathsf{D}(1/T))$ increases with the lead-time H of the node, while decreases with α. Similar results apply to $z_{f,d}(O, \mathsf{M}(r))$, which increases with the lead-time of the downstream agent.

As a remark, equations (38) and (39) permit to evaluate the amplification of the variability at any consecutive nodes in a supply chain.

4.2.2 Extended discussion

Similarly, suppose the initial inventory level $I_0^{(2)}$ is the control variable determined at the beginning of the planning period. Assume $D_n \in N(\lambda T, \sigma^2)$, $I_0^{(2)}$ is then set to be a multiple of $Std[I_n^{(2)}]$, and

$$I_0^{(2)} = \xi^{(2)}\sqrt{L\lambda T z_{f,d}(O, \mathsf{D}(1/T)) + L^2(1-\beta)\lambda T z_{f,d}(O, \mathsf{M}(r))} \tag{40}$$

with $\xi^{(2)}$ the critical-fractile of service for the upstream node.

Substituting (38) and (39) into (40), we then continue to investigate the impact of lead-time, demand correlation, and length of review period on the safety stock level of the upstream stage.

When $\alpha = 1$, the downstream demands are independent and identically distributed. By (38) , $z_{f,d}(O_d, \mathsf{D}(1/T)) = z_{f,d}(X_d, \mathsf{D}(1/T))$ and by (39) $z_{f,d}(O, \mathsf{M}(r)) = z_{f,d}(X_d, \mathsf{D}(1/T))/(1+\beta)$, where β is the forecast parameter of the upstream agent. Noticing $z_{f,d}(X_d, \mathsf{D}(1/T)) = \sigma^2/(\lambda T)$,

$$I_0^{(2)} = \xi\sqrt{L\sigma^2 + L^2(1-\beta)\sigma^2/(1+\beta)}.$$

The safety stock is dependent on the order lead-time L.

When $\alpha < 1$, the safety stock of the upstream agent is then dependent not only on the upstream lead-time, but also the downstream lead-time, similar to Graves (1999).

Observing (40), we also find that as the review length gets smaller, the inventory requirement decreases.

4.2.3 Quantifying the bullwhip effect

Based on the variability propagation equation, we are now able to quantify the well-known bullwhip effect using the peakedness measurement. We focus on a two-echelon model while the results can be extended to the general case.

The following results hold based on the previous analysis.

Corollary 4.5. *The variance of the order flow is propagated as*

$$\begin{aligned} Var[O_n^{(2)}] &= Var[O_n](1+2L(1-\beta)) + 2(1-\beta)L^2\, Var[F^{(2)}] \\ &= \lambda T z_{f,d}(O, \mathsf{D}(1/T))(1+2L(1-\beta)) \\ &\quad +2(1-\beta)^2 L^2 \lambda T z_{f,d}(O, \mathsf{M}(r)) \end{aligned} \tag{41}$$

In addition, $Var[O_n^{(2)}] \geq Var[O_n] \geq Var[D_n]$.

As indicated in Corollary 4.5, the bullwhip effect of the upstream stage depends not only on its lead-time L, but also on the lead-time of the downstream stage H. We can also easily verify that the batch order size λT and the forecasting parameters α and β impact the bullwhip effect.

4.3 A tree-structured supply chain model

It is relatively easy to extend the model presented so far to predict the variability in a tree-structured supply chain. Consider a supply chain model with a single upstream agent and multiple downstream agents. We denote by λ_i the arrival rate of demand at downstream agent i, and λ_{in} at the upstream node. For technical simplicity, we need to assume that demands at different downstream stages are independent. The analysis for the single-stage model applies to each downstream agent.

With various order flows arriving to an upstream node, the arrival rate and peakedness are merged according to Proposition 4.6.

Proposition 4.6.

The arrival flows of downstream agents are merged to the upstream agent with the arrival rate $\lambda_{in} = \sum_i \lambda_i$. *In addition,* $z_{in} = \frac{\sum_i \lambda_i z_i}{\sum_i \lambda_i}$, *where* z_i *is the peakedness for the downstream agent* i, *and* z_{in} *is the peakedness for the upstream agent.*

The proof is easily obtained using basic computations on the variance of $S(t)$. See Tabordon (2002) for the details.

By Proposition 4.6, we can therefore identify the initial safety stock level for the upstream agent based on (40), assuming that all unsatisfied demands are backlogged.

We now apply Proposition 4.6 to illustrate the possible utilization of our model. We use a simple example with one supplier and N retailers, the total time needed for the supplier to produce an item and deliver it to a retailer is L. We will compare two types of policies, in the first case the policy is completely decentralized and the retailers manage their inventories locally and order to the supplier with a lead-time of L, consequently the supplier does not need to anticipate orders, she can start production as soon as an order arrives and deliver it on time. In the second case, which models a VMI type policy, the supplier observes store inventory in real time, she manages the global inventory and delivers to the retailer as needed to satisfy demand (we suppose she can do so sufficiently fast in order to neglect safety stocks at the retailer). These are in fact the two extreme cases, we could imagine any situation where the retailer orders with a lead-time of $L_r < L$ and the supplier maintains some inventory such as to anticipate demand over a time of $L - L_r$.

Suppose a tree-structured supply chain that consists in one supplier and N retailer stores. Retailers observe their demand, make their forecast, and place orders to the supplier in each period. Demands are assumed to be independent. Denote by λ_i the arrival rate and z_i the peakedness for the demand process at store i. For simplicity, we suppose that review length, lead-times and service levels for the different retailers are the same.

We first consider the case where each store holds inventory independently. By Proposition 4.3, the initial inventory decision for store i, $I_{i,0}$, is then

$$I_{i,0} = \xi \sqrt{H\lambda_i T z_{f,d,i}(X_d, \mathsf{D}(1/T)) + H^2(1-\alpha_i)\lambda_i T z_{f,d,i}(X_d, \mathsf{M}(s))}. \quad (42)$$

The total inventory stocks for the supply chain system is then $\sum_i I_{i,0}$.

If employing VMI, the retailer stores provide demand information λ_i and $z_{f,d,i}$ to the supplier and the supplier delivers the products to the retailer stores while maintaining an agreed service level. The initial inventory decision for the supplier, $I_{d,0}$, is

$$I_{d,0} = \xi \sqrt{H\lambda_{in} T z_{f,d,in}(X_d, \mathsf{D}(1/T)) + H^2(1-\alpha_{in})\lambda_{in} T z_{f,d,in}(X_d, \mathsf{M}(s))}, \quad (43)$$

while λ_{in} and $z_{f,d,in}$ are obtained with Proposition 4.6. Comparing the inventory stock requirements under both cases, i.e., $\sum_i I_{i,0} - I_{d,0}$, we can quantify the impact of VMI on inventories depending on the characteristics of the demand.

5 Numerical study

In this section we report a numerical study conducted (1) to demonstrate how to calculate the peakedness from real life order flows, (2) to test whether the variability is adequately estimated by the peakedness approach. We compare the bullwhip effects of a two-stage supply chain by our peakedness results with those of simulation approach and Graves (1999) as well. We further show that by the peakedness we can determine the safety stock level at the upstream stage of the supply chain, and estimate the total holding and shortage cost.

Our numerical study is based on real life data collected from a supermarket of Delhaize Group, which is a Belgian food retailer consisting of more than 2600 stores on three continents. We suppose that the retail store manager observes the sales every day and aggregates the demand of the entire week, i.e., $T = 1$ week. Every week he makes a new forecast and issues an order to the distribution center (DC). The DC delivers the orders within that week, and the order fulfillment lead-time is $H = 1$ week. Similarly, the DC issues orders to its upstream agent weekly, and the order fulfillment lead-time is $L = 2$ weeks.

We first simulate the two-stage supply chain process. For the purpose of this study, we consider the aggregated weekly sales for three different products over a year, and regard it as the weekly demand, i.e., $\{X_d\}$. We choose the smoothing factor value α that minimizes the forecast error, where $\alpha = e^{-sT}$ minimizes the mean square forecasting error. Weekly forecasts are computed using (15), and the warm up effects with the forecasts are avoided by duplicating the dataset to obtain a two-year horizon while using the data for the second year only. Order quantity for each week can thus be determined by (17) assuming that it employs an adaptive base-stock policy. We can then obtain the bullwhip effect at the retailer stage by computing the ratio of the variance of the order and the variance of the demand, i.e. $\mathrm{Var}[O_n]/\mathrm{Var}[D_n]$.

The sum of the order flows from the retailers is the demand arrival process to the wholesaler, we repeat the simulation process with $L = 2$ week. Forecasts and orders of the wholesaler are determined by (29) and (31), respectively. The smoothing factor $\beta = e^{-rT}$ is obtained by minimizing the mean square forecasting error, like α for the demand process. It should

Table 1: Simulation results

	Product	1	2	3
	Mean Demand	120.83	120.98	1.73
Retailer	Variance of Demand	750.03	416.06	2.08
	Smoothing factor (α)	0.65	0.80	0.95
	Variance of Orders	1243.30	572.31	2.24
Wholesaler	Variance of Demand	1243.30	572.31	2.24
	Smoothing factor (β)	0.95	0.85	1.00
	Variance of Orders	1506.17	934.96	2.24

be noted that β in Graves (1999) is obtained by $1-\beta = (1-\alpha)/(1+H(1-\alpha))$. We summarize the simulation results including the demand and order variances in Table 1. Similarly, the bullwhip effect (BWE) at the wholesaler stage is measured by computing the ratio of the variance of its order and the variance of its incoming demand, i.e. $\text{Var}[O_n^{(2)}]/\text{Var}[O_n]$.

The data received from the supermarket chain are daily sales for each of the three products, the data about the exact checkout time of products is aggregated every night and is not conserved by the chain in order to have a more compact database. We used the results of subsection 3.4 to compute the peakedness for each product.

Now we are ready to estimate the order flow variability and the bullwhip effect by the peakedness. The peakedness of the order quantities $z_{f,d}(O, D(1/T))$ can be computed by (38). We can estimate $z_{f,d}(O, \mathsf{M}(r))$ using (39), and $z_{f,d}(O^{(2)}, D(1/T))$ by dividing (41) by $E(O_n) = \lambda T$. All the peakedness values are listed in Table 2. The bullwhip effect of the retailer can then be calculated by $z_{f,d}(O, D(1/T))/z_{f,d}(X_d, D(1/T))$, noticing that the order flow rates stay the same (we assume no orders are lost).

As indicated in Graves (1999), if the demand process of the downstream stage is an ARIMA(0,1,1) process, the order process, namely, the demand process of the upstream stage is also an ARIMA (0,1,1) process. In addition, a simplified measure of order variance amplification can be computed by $(1+H*(1-\alpha))^2$ (note that α in this paper corresponds to '$1-\alpha$' in Graves (1999)), as compared to the variance of the demand. We summarize the estimated bullwhip effects by different approaches in Table 3.

Table 3 reveals that overall the bullwhip effects estimated by the peakedness approach is close to the simulation results and those estimated by Graves (1999). We also see that bullwhip effect estimated by Graves (1999)

Table 2: The peakedness result

	Product	1	2	3
Retailer	$z_{f,d}((X_d, D(1/T))$	6.21	3.44	1.20
	$z_{f,d}((X_d, \mathsf{M}(s))$	6.90	4.28	0.51
	$z_{f,d}((O, D(1/T))$	12.24	5.16	1.33
Wholesaler	$z_{f,d}(O, D(1/T))$	12.24	5.16	1.33
	$z_{f,d}((O, M(r))$	3.59	2.29	0.60
	$z_{f,d}((O^{(2)}, D(1/T))$	14.76	8.66	1.33

Table 3: Comparison of bullwhip effects

	Product	1	2	3
Retailer	Simulation	1.66	1.34	1.07
	Peakedness	1.97	1.50	1.10
	Graves (1999)	1.82	1.44	1.10
Wholesaler	Simulation	1.21	1.63	1.00
	Peakedness	1.21	1.68	1.00
	Graves (1999)	2.31	1.78	1.20

is a little over-estimated at the upstream stage level.

In order to compare the safety stocks for these three different approaches, we calculate the safety stock levels at the wholesaler's stage based on a simplified formula of (28) (assuming $\alpha = 1$):

$$ss = \xi \cdot \sqrt{Var[O] \cdot L},$$

where ξ corresponds to the service level (SL), and L is the lead time for replenishment at the wholesaler stage, $Var[O]$ is the variance of the demand at the wholesaler stage. In addition, $Var[O] = BWE \cdot Var[D]$. Here BWE is computed and listed in Table 3 and $Var[D]$ is listed in Table 1 for different approaches. We summarize the results of safety stock levels in Tables 4 to 6 for different service levels . As observed, the safety stock levels of the peakedness approach is slightly higher as compared to the other two. This is because the estimated bullwhip effects at the retailer's stage by the peakedness approach is slightly larger.

Suppose a simple setting with unit holding cost h and shortage cost s. We calculate the weekly holding and shortage cost based on the real data, with different safety stock levels, that is, $TC = h \cdot max[I_n^{(2)}, 0] + s \cdot max[-I_n^{(2)}, 0]$. The inventory level $I_n^{(2)}$ is calculated based on (32). When $I_n^{(2)} \geq 0$, holding cost occurs, otherwise, shortage occurs. We consider different service levels with $SL = 1 - h/s$ for different s/h ratio. We present the relative cost values in Table 4 (costs were normalized so that the cost of the policy based on the simulation is 100).

We are also interested in the effective service level with these different safety stock levels. We calculate the effective service level by the number of periods when a stockout occurs as compared to the total number periods in the planning horizon. The section "Effective service level" in Table 4 gives the difference between the effective service levels reached and the target service levels (with $'-'$ meaning below the target and vice versa).

As observed from Tables 4 to 6, if we calculate the average of the normalized cost for these 27 cases, we find that the peakedness approach achieves the smallest average cost, that is, $TC_{peakedness} = 99.31$, while Graves (1999) approach is the largest cost with $TC_{graves} = 102.96$, given the simulation cost is 100. On the contrary, the distance of the effective service level to the target service level by the peakedness approach is the highest, that is, $ESL_{peakedness} = -0.7\%$, whilst Graves (1999) approach obtains the smallest effective service level with $ESL_{graves} = -3.5\%$. Again, the simulation approach is in the middle at the value of -1.1%. As a summary, we notice that the better precision of the peakedness model makes it possible to determine safety inventory levels that generate savings for the supply chain.

6 Conclusion and further research

In this paper, we first propose to use the peakedness as a way of measuring the variability of flows in a supply chain. The main advantage of this approach is that it requires less assumptions on the underlying order flows. Assuming that demand is a general point process, we generalize the most popular demand models in the literature, such as time-series models, Poisson processes, and renewal processes. Though the peakedness might seem a less intuitive measure of variability, it is actually very easy to compute from real life data based on a sufficiently large sample of observations.

We further show that the peakedness can be used to characterize basic inventory models in a supply chain system. We start with a single-stage inventory model. Using the peakedness of the demand, we are able to characterize the forecast (assuming forecasts are made using exponential smoothing), the order decisions, and the inventory levels. We further make an extended discussion of demand correlation and lead-time, and their impact on the initial inventory stock level. By establishing the variability propagation equation of the peakedness, we can further extend the results to multiple stage inventory system, and quantify the bullwhip effects of the supply chain. Finally, we extend the peakedness analysis to a tree-structured supply chain network by proposing the merging equations of the peakedness.

Employing real life data from a Belgian supermarket, we show numerically how the peakedness can be calculated easily. The results also verify that the peakedness can measure the order flow variability effectively. As compared to the ARIMA (0,1,1) assumption in Graves (1999), the peakedness approach can be applied to general point processes, and the estimated bullwhip effects are close to the simulation results.

In addition, our model can also includes cost considerations such as inventory holding and shortage costs to find the optimal initial inventory levels in order to achieve a specified service level. Such an analysis is based on the assumption that inventory fluctuations follow a normal distribution.

This paper is the first attempt to use the concept of peakedness to analyze the variability in supply chain. We can see a lot of future applications based on the peakedness model considering its simplicity for implementation and amenability for optimizing purposes. It is likely that more decision variables could be introduced in the model, and there are many interesting research questions. For example, how to use the peakedness for supply chain planning, and manufacturing strategy decision. In addition, as the resulting model is easy to compute, it could be integrated in economics based models such as the principal-agent based models that are very common in the supply chain management literature. As Cachon (1999) observed the cost of variability might be very different at different stages of the supply chain, it would be interesting to link a cost model to the model presented here. It would then be possible to build an optimization procedure to find the most cost effective supply chain structure.

A Link between the peakedness and the fluid peakedness

Given a point process $X(t)$ with rate λ, the peakedness for a general service process $\mathsf{G}(s)$ is computed by

$$z(X,\mathsf{G}(s)) = 1+\frac{s}{\lambda_X}\int_{-\infty}^{\infty}(k_X(t)-\lambda_X\delta(t))\rho_{F^c}(t)dt, \tag{44}$$

where $k_X(t)$ is the covariance density function of the point process $X(t)$, $\delta(t)$ is the Dirac delta function and $\rho_{F^c}(t)$ is the autocorrelation function of the service time distribution function. The fluid peakedness is computed by

$$z_f(X,\mathsf{G}(s)) = \frac{s}{\lambda_X}\int_{-\infty}^{\infty}k_X(t)\rho_{F^c}(t)dt. \tag{45}$$

See Eckberg (1983) for more details about different ways of measuring the peakedness. By comparing (44) and (45), we notice that

$$z_f(X,\mathsf{G}(s)) = z(X,\mathsf{G}(s))-1+s\rho_{F^c}(0). \tag{46}$$

Equation (46) implies that, for a given service time distribution G, both definitions of peakedness only differ by a constant, i.e. a value independent of the point process $X(t)$ considered, and therefore provide us with the same information about the point process.

Proof of Property 2. When the service times are exponentially distributed, $\rho_{F^c}(t)=\frac{1}{2s}e^{-s|t|}$. Substituting $\rho_{F^c}(0)=\frac{1}{2s}$ into (46) proves Property 2. □

B Other technical proofs

B.1 Proof of Lemma 4.1

Based on equation (15), we can write

$$F_n = (1-\alpha)\sum_{i=0}^{n-1}\alpha^i D_{n-i}. \tag{47}$$

In the mean time, by (4) we have

$$S_f(X_d,\mathsf{M}(s),nT) = \sum_{i=0}^{n-1}\alpha^i D_{n-i}. \tag{48}$$

Combining (47) and (48) yields Equation (22), the first statement of Lemma 4.1. Furthermore, we already proved (see (12)) that

$$E[S_f(X_d, \mathsf{M}(s), nT)] = \frac{\lambda T}{1-\alpha}.$$

We can therefore express the variability of the forecasts as a function of the peakedness,

$$\begin{aligned} \mathrm{Var}[F_n] &= (1-\alpha)^2 \mathrm{Var}[S_f(X_d, \mathsf{M}(s), nT)] \\ &= (1-\alpha)\lambda T z_{f,d}(X_d, \mathsf{M}(s), nT). \end{aligned}$$

This completes the proof of Lemma 4.1. □

B.2 Proof of Lemma 4.2

Equation (24) is obtained by substituting (20) and (22) into (17).

From equation (17), we find that

$$\mathrm{Var}[O_n] = \mathrm{Var}[D_n] + H^2 \mathrm{Var}[\Delta F_n] + 2H \mathrm{Cov}[D_n, \Delta F_n], \qquad (49)$$

with ΔF_n defined as the difference between two successive forecasts, that is $\Delta F_n = F_n - F_{n-1}$. We prove (25) by expressing (49) as a function of $\mathrm{Var}[D_n]$ and $\mathrm{Var}[F_n]$. We then substitute them with (21) and (23). For the second term we have:

$$\begin{aligned} \mathrm{Var}[\Delta F_n] &= \mathrm{Var}[F_n] + \mathrm{Var}[F_{n-1}] - 2\,\mathrm{Cov}[F_n, F_{n-1}] \\ &= 2\mathrm{Var}[F_n] - 2\,\mathrm{Cov}[(1-\alpha)D_n + \alpha F_{n-1}, F_{n-1}] \\ &= 2(1-\alpha)\mathrm{Var}[F_n] - 2(1-\alpha)\,\mathrm{Cov}[D_n, F_{n-1}]). \end{aligned}$$

Noticing that $F_{n-1} = (1-\alpha)D_{n-1} + \alpha F_{n-2} = (1-\alpha)\sum_{i=0}^{\infty} \alpha^i D_{n-1-i}$, we can write,

$$\begin{aligned} \mathrm{Cov}[D_n, F_{n-1}] &= \mathrm{Cov}[D_n, (1-\alpha)\sum_{i=0}^{\infty} \alpha^i D_{n-1-i}] \\ &= (1-\alpha)\sum_{i=0}^{\infty} \alpha^i \mathrm{Cov}[D_n, D_{n-1-i}] \\ &= (1-\alpha)\sum_{i=1}^{\infty} \alpha^{i-1} \mathrm{Cov}[D_n, D_{n-i}]. \end{aligned}$$

Thus,

$$\text{Var}[\Delta F_n] = 2(1-\alpha)\text{Var}[F_n] - 2(1-\alpha)^2 \sum_{i=1}^{\infty} \alpha^{i-1}\text{Cov}[D_n, D_{n-i}].$$

Focusing now on the third term of the right-hand side of (49), we can write:

$$\begin{aligned}
\text{Cov}[D_n, (F_n - F_{n-1})] &= \text{Cov}[D_n, F_n] - \text{Cov}[D_n, F_{n-1}] \\
&= \text{Cov}[D_n, (1-\alpha)\sum_{i=0}^{\infty} \alpha^i D_{n-i}] \\
&\quad -\text{Cov}[D_n, (1-\alpha)\sum_{i=0}^{\infty} \alpha^i D_{n-1-i}] \\
&= (1-\alpha)\text{Var}[D_n] + (1-\alpha)\sum_{i=1}^{\infty} \alpha^i \text{Cov}[D_n, D_{n-i}] \\
&\quad -(1-\alpha)\sum_{i=0}^{\infty} \alpha^i \text{Cov}[D_n, D_{n-i-1}] \\
&= (1-\alpha)\text{Var}[D_n] \\
&\quad -(1-\alpha)^2 \sum_{i=1}^{\infty} \alpha^{i-1}\text{Cov}[D_n, D_{n-i}].
\end{aligned}$$

As a result, by including these two equalities into equation (49), we obtain:

$$\begin{aligned}
\text{Var}[O_n] &= (1 + 2H(1-\alpha))\text{Var}[D_n] + 2(1-\alpha)H^2\text{Var}[F_n] \\
&\quad -2H(1-\alpha)^2[H+1]\sum_{i=1}^{\infty} \alpha^{i-1}\text{Cov}[D_n, D_{n-i}].
\end{aligned}$$

If we ignore the last item, which is the autocovariance effect of demands across periods, we then obtain a simplified expression of $\text{Var}[O_n]$, and thus Lemma 4.2 is proven. □

B.3 Proof of Lemma 4.4

Equation (38) is obtained by dividing (25) with the expectation, i.e., λT.

To prove equation (39), let us first find $S_f(O, \mathsf{M}(r), nT)$:

$$S_f(O, \mathsf{M}(r), nT) = \sum_{i=0}^{\infty} \beta^i O_{n-i}$$

where $\beta = e^{-rT}$. From this, we have:

$$\begin{aligned}
E[S_f(O, \mathsf{M}(r), nT)] &= \frac{E[O]}{1-\beta} \\
&= \frac{\lambda T}{1-\beta} \\
\mathrm{Var}[S_f(O, \mathsf{M}(r), nT)] &= \sum_{i=0}^{\infty} \beta^{2i} \mathrm{Var}[O_n] \\
&\quad +2 \sum_{i=0}^{\infty} \mathrm{Cov}[\beta^i O_{n-i}, \sum_{j=1}^{\infty} \beta^{i+j} O_{n-j-i}] \\
&= \frac{1}{1-\beta^2} \mathrm{Var}[O_n] \\
&\quad - \frac{2\beta(1-\alpha)^2}{(1-\beta^2)(1-\alpha\beta)} (\mathrm{Var}[D_n]H + H^2 \mathrm{Var}[F_n]).
\end{aligned}$$

To obtain this last expression, we used the fact that when we assume $\mathrm{Cov}[D_n, D_{n-1}] = 0$, we have:

$$\begin{aligned}
\mathrm{Cov}[O_n, O_{n-1}] &= -(1-\alpha)^2 (\mathrm{Var}[D_n]H + H^2 \mathrm{Var}[F_n]) \\
\mathrm{Cov}[O_n, O_{n-2}] &= \alpha \mathrm{Cov}[O_n, O_{n-1}]
\end{aligned}$$

With this result, it is relatively straigth forward to obtain a closed form expression for the peakedness of the order placed by an agent: we divide $\mathrm{Var}[S_f(O, \mathsf{M}(r), nT)]$ by $E[S_f(O, \mathsf{M}(r), nT)]$.

$$\begin{aligned}
&z_{f,d}(O, \mathsf{M}(r)) \\
&= \left(\frac{\mathrm{Var}[O_n]}{1-\beta^2} - \frac{2\beta(1-\alpha)^2}{(1-\beta^2)(1-\alpha\beta)} (\mathrm{Var}[D_n]H + H^2 \mathrm{Var}[F_n])] \right) / \frac{\lambda T}{1-\beta} \\
&= \frac{1}{(1+\beta)\lambda T} \Big(\mathrm{Var}[D_n] + H^2 \mathrm{Var}[\Delta F_n] + 2F(1-\alpha) \mathrm{Var}[D_n] \\
&\qquad - \frac{2\beta(1-\alpha)^2}{1-\alpha\beta} (\mathrm{Var}[D_n]H + H^2 \mathrm{Var}[F_n]) \Big) \\
&= \frac{1}{(1+\beta)\lambda T} \Big(\mathrm{Var}[D_n](1 + 2H \frac{(1-\alpha)(1-\beta)}{1-\beta\alpha}) \\
&\qquad + 2H^2 \mathrm{Var}[F_n] \frac{(1-\alpha)(1-\beta)}{1-\beta\alpha} \Big).
\end{aligned}$$

The last step is to replace $\mathrm{Var}[D_n]$ and $\mathrm{Var}[F_n]$ by the expressions provided by equations (21) and (23). □

References

Aviv, Y. 2003. A time-series framework for supply chain inventory management *Oper. Res.* , **51**(2) 210–227.

Browne, S. and Zipkin, P. 1991. Inventory Models with Continuous, Stochastic Demands. *Ann. Appl. Probab.*, **1**(3) 419–435.

Balakrishnan, A., Geunes J., and Pangburn M. S. 2004. Coordinating Supply Chains by Controlling Upstream Variability Propagation. *Manufacturing Service Oper. Management*, **6**(2) 163–183.

Cachon, G. P. 1999. Managing Supply Chain Demand Variability with Scheduled Ordering Policies. *Management Sci.*, **45**(6) 457–479.

Cachon, G. P., Randall, T., and Schmidt, G.M. 2007. In Search of the Bullwhip Effect. *Manufacturing Service Oper. Management*, **9**(4) 457–479.

Chen, F. 2003. Information sharing and supply chain coordination. A. G. de Kok, S. C. Graves, eds. Handbooks in Operations Research and Management Science, Vol. 11. Supply Chain Management: Design, Coordination, and Operation: 341 - 421, Elsevier, Armsterdam.

Chen, L., Lee, H. L. 2009. Information Sharing and Order Variability Control Under a Generalized Demand Model. *Management Sci.*, **55**(5) 781–797.

Chen, F., Drezner, Z., Ryan, J. K., and Simchi-levi, D. 2000. Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times, and Information. *Management Sci.*, **46**(3) 436–443 .

Cox, D.R. 1955. Some Statistical Methods Connected with Series of Events. *J. Roy. Stat. Soc. B Met.* **17**(2) 129–164.

Croson, R. and Donohue, K. 2006. Behavorial Causes of the Bullwhip Effect and the Observed Value of the Inventory Information. *Management Sci.* **52**(3) 323–336.

Daley D.J. and Vere-Jones D. 2002. An introduction to the theory of point process *Volume I: Elementary theory and methods, Chapter 3* 2nd edition. Springer-Verlag, New York.

Dejonckheere, J., Disney, S., Lambrecht, M., and Towill, D. 2004. The impact of information enrichment on the Bullwhip effect in supply chains: A control engineering perspective. *Eur. J. Oper. Res.* **153**(3) 727–750.

A. E. Eckberg, J. 1983. Generalized peakedness of teletraffic processes. In *10th Int. Teletraffic Congress.*

Ettl, M., G. E. Feigin, G. Y. Lin, D. D. Yao. 2000. A supply network model with base-stock control and service requirements. *Oper. Res.* **49**(2) 216–232.

Gilbert, K. 2005. An ARIMA Supply Chain Model. *Management Sci.*, **51**(2) 305–310.

Graves, S. C. 1999. A Single-Item Inventory Model for a Non Stationary Demand Process. *Manufacturing Service Oper. Management*, **1**(1) 50–61.

Graves, S. C., Willems, S. P 2000. Optimizing Strategic Safety Stock Placement in Supply Chains. *Manufacturing Service Oper. Management*, **2**(1) 68–83.

Graves, S. C., Willems, S. P 2008. Strategic Inventory Placement in Supply Chains: Nonstationary Demand. *Manufacturing Service Oper. Management*, **10**(2) 278–287.

Graves, S. C., Willems, S. P 2003. Supply chain design: Safety stock placement and supply chain configuration. A. G. de Kok, S. C. Graves, eds. Handbooks in Operations Research and Management Science, Vol. 11, Chapter 3. Supply Chain Management: Design, Coordination and Operation. North-Holland, Amsterdam, The Netherlands, 95–132.

Hadley, G., and Whitin, T. M. 1963. Analysis of Inventory Systems. Prentice-Hall Englewood Cliffs, N.J.

Jagerman, D. L., Melamed, B., and Willinger, W. 1997. Stochastic Modeling of Traffic Processes. Technical Report 7-96, Rutgers Center for Operations Research (RUTCOR).

Kahn, J. A. 1987. Inventories and the Volatility of Production. *Amer. Econ. Rev.*, **77** (4) 667–679.

Lee, H. L., Padmanabhan, V., and Whang, S. 1997. Information Distortion in a Supply Chain: The Bullwhip Effect. *Management Sci.*, **43**(4) 546–558.

Lee, H, L., So, K. C., Tang, C. S. 2000. The value of information sharing in a two-level supply chain. *Management Sci.*, **46**(5) 626–643.

Macq, J.-F. 2005. *Optimization of Multimedia Flows over Data Networks: the Core Location Problem and the Peakedness Characterization*. PhD thesis, Faculté des Sciences Appliquées.

Ozekici, S., and Soyer, R. 2006. Semi-Markov Modulated Poisson Process: Probabilistic and Statistical Analysis. *Math. Method Oper. Res.*, 64(1): 125–144.

Raghunathan, S. 2001. Information sharing in a supply chain: A note on its value when demand is nonstationary. *Management Sci.*, **47**(4) 605–610.

Sahin, I. 1979. On the Stationary Analysis of Continuous Review (s, S) Inventory Systems with Constant Lead Times. *Oper. Res.*, **27**(4) 717–729.

Sahin, I. 1982. On the Objective Function Behavior in (s, S) Inventory Models. *Oper. Res.*, **30**(4) 709–724.

Schoenmeyr, T., Graves, S. C. 2009. Strategic Safety Stocks in Supply Chains with Evolving Forecasts. *Manufacturing Service Oper. Management*, **11**(4) 657–673.

Simchi-Levi, D., Zhao, Y. 2005. Safety Stock Positioning in Supply Chains with Stochastic Lead Times. *Manufacturing Service Oper. Management*, **7**(4) 295–318.

Sivazlian, B.D. 1974. A Continous-Review (s, S) Inventory System with Arbitrary Interarrival Distribution between Unit Demand. *Oper. Res.*, **22**(1) 65–71.

Song, J.S. and Zipkin, P. 1993. Inventory Control in a fluctuating demand environmen. *Oper. Res.*, **41**(2) 351–370.

Tabordon, N. 2002. *Modeling and Optimizing the Management of Operator Training in a Call Center*. PhD thesis, Institut D'Administration et de Gestion.

Veinott, A. 1965. Optimal policy for a multi-product, dynamic, non-stationary inventory problem. *Management Sci.*, **12** 206–222.

Wilkinson, R. 1956. Theories for toll traffic engineering in the usa. *AT&T Tech. J.* , **35**(2) 421–514.

Wolff, R. W. 1989. *Stochastic Modeling and the Theory of Queues*. Prentice-Hall, Englewood Cliffs, NJ.

Table 4: Safety stock, normalized cost and effective service level. Part 1.

Product	1	2	3	1	2	3	1	2	3
Service Levels	80%	($\frac{s}{h} = 5$)		90%	($\frac{s}{h} = 10$)		93.3%	($\frac{s}{h} = 15$)	
Safety stock level									
Simulation	42	29	2	64	44	3	75	51	4
Peakedness	46	30	2	70	46	3	82	54	4
Graves (1999)	45	30	2	68	45	3	79	52	4
Normalized cost									
Simulation	100	100	100	100	100	100	100	100	100
Peakedness	101.38	100.69	100	99.68	100.12	100	101.24	100.25	100
Graves (1999)	110.69	101.20	101.14	113.44	100.45	104.04	116.53	101.23	104.00
Effective service level									
Simulation	8.5%	8.5%	-8.8%	0.4%	0.4%	-9.2%	0.9%	-1.0%	-2.9%
Peakedness	8.5%	10.4%	-8.8%	4%	2.3%	-9.2%	4.8%	0.9%	-2.9%
Graves (1999)	2.7%	8.5%	-6.9%	-3.5%	4%	-9.2%	-4.8%	-1.0%	-4.8%

Table 5: Safety stock, normalized cost and effective service level. Part 2.

Product	1	2	3	1	2	3	1	2	3
Service Levels	95%	($\frac{s}{h} = 20$)		96%	($\frac{s}{h} = 25$)		97.5%	($\frac{s}{h} = 40$)	
Safety stock level									
Simulation	83	56	4	88	60	4	98	67	5
Peakedness	90	59	4	96	62	4	107	70	5
Graves (1999)	87	57	4	92	61	4	103	68	5
Normalized cost									
Simulation	100	100	100	100	100	100	100	100	100
Peakedness	103.3	99.39	199	102.97	99.11	100.00	101.24	96.86	100.00
Graves (1999)	113.51	100.64	103.90	110.02	100.17	103.81	99.29	99.04	103.69
Effective service level									
Simulation	3.1%	-0.8%	-4.6%	2.1%	-1.8%	-5.6%	0.6%	-3.3%	0.6%
Peakedness	3.1%	-0.8%	-4.6%	2.1%	-1.8%	-5.6%	0.6%	-3.3%	0.6%
Graves (1999)	-6.5%	-2.7%	-6.5%	-5.6%	-3.7%	-7.5%	-5.2%	-5.2%	-1.3%

Table 6: Safety stock, normalized cost and effective service level. Part 3.

Product	1	2	3	1	2	3	1	2	3
Service Levels	98%	($\frac{s}{h} = 50$)		98.5%	($\frac{s}{h} = 70$)		99%	($\frac{s}{h} = 100$)	
Safety stock level									
Simulation	103	70	5	110	75	5	117	79	5
Peakedness	112	73	5	120	78	5	127	83	5
Graves (1999)	108	72	5	115	76	5	122	81	5
Normalized cost									
Simulation	100	100	100	100	100	100	100	100	100
Peakedness	100.11	95.80	100.00	97.93	93.72	100.00	95.03	92.64	100.00
Graves (1999)	96.89	97.00	103.39	98.00	97.20	103.39	99.36	94.92	103.39
Effective service level									
Simulation	0.1%	-3.8%	0.1%	-0.4%	-4.3%	-0.4%	-0.9%	-4.8%	-0.9%
Peakedness	0.1%	-3.8%	0.1%	-0.4%	-4.3%	-0.4%	-0.9%	-4.8%	-0.9%
Graves (1999)	-1.8%	-5.7%	-1.8%	-2.3%	-6.2%	-2.3%	-2.8%	-6.7%	-2.8%

Recent titles

CORE Discussion Papers

2010/27. Olivier DURAND-LASSERVE, Axel PIERRU and Yves SMEERS. Uncertain long-run emissions targets, CO_2 price and global energy transition: a general equilibrium approach.
2010/28. Andreas EHRENMANN and Yves SMEERS. Stochastic equilibrium models for generation capacity expansion.
2010/29. Olivier DEVOLDER, François GLINEUR and Yu. NESTEROV. Solving infinite-dimensional optimization problems by polynomial approximation.
2010/30. Helmuth CREMER and Pierre PESTIEAU. The economics of wealth transfer tax.
2010/31. Thierry BRECHET and Sylvette LY. Technological greening, eco-efficiency, and no-regret strategy.
2010/32. Axel GAUTIER and Dimitri PAOLINI. Universal service financing in competitive postal markets: one size does not fit all.
2010/33. Daria ONORI. Competition and growth: reinterpreting their relationship.
2010/34. Olivier DEVOLDER, François GLINEUR and Yu. NESTEROV. Double smoothing technique for infinite-dimensional optimization problems with applications to optimal control.
2010/35. Jean-Jacques DETHIER, Pierre PESTIEAU and Rabia ALI. The impact of a minimum pension on old age poverty and its budgetary cost. Evidence from Latin America.
2010/36. Stéphane ZUBER. Justifying social discounting: the rank-discounting utilitarian approach.
2010/37. Marc FLEURBAEY, Thibault GAJDOS and Stéphane ZUBER. Social rationality, separability, and equity under uncertainty.
2010/38. Helmuth CREMER and Pierre PESTIEAU. Myopia, redistribution and pensions.
2010/39. Giacomo SBRANA and Andrea SILVESTRINI. Aggregation of exponential smoothing processes with an application to portfolio risk evaluation.
2010/40. Jean-François CARPANTIER. Commodities inventory effect.
2010/41. Pierre PESTIEAU and Maria RACIONERO. Tagging with leisure needs.
2010/42. Knud J. MUNK. The optimal commodity tax system as a compromise between two objectives.
2010/43. Marie-Louise LEROUX and Gregory PONTHIERE. Utilitarianism and unequal longevities: A remedy?
2010/44. Michel DENUIT, Louis EECKHOUDT, Ilia TSETLIN and Robert L. WINKLER. Multivariate concave and convex stochastic dominance.
2010/45. Rüdiger STEPHAN. An extension of disjunctive programming and its impact for compact tree formulations.
2010/46. Jorge MANZI, Ernesto SAN MARTIN and Sébastien VAN BELLEGEM. School system evaluation by value-added analysis under endogeneity.
2010/47. Nicolas GILLIS and François GLINEUR. A multilevel approach for nonnegative matrix factorization.
2010/48. Marie-Louise LEROUX and Pierre PESTIEAU. The political economy of derived pension rights.
2010/49. Jeroen V.K. ROMBOUTS and Lars STENTOFT. Option pricing with asymmetric heteroskedastic normal mixture models.
2010/50. Maik SCHWARZ, Sébastien VAN BELLEGEM and Jean-Pierre FLORENS. Nonparametric frontier estimation from noisy data.
2010/51. Nicolas GILLIS and François GLINEUR. On the geometric interpretation of the nonnegative rank.
2010/52. Yves SMEERS, Giorgia OGGIONI, Elisabetta ALLEVI and Siegfried SCHAIBLE. Generalized Nash Equilibrium and market coupling in the European power system.
2010/53. Giorgia OGGIONI and Yves SMEERS. Market coupling and the organization of counter-trading: separating energy and transmission again?
2010/54. Helmuth CREMER, Firouz GAHVARI and Pierre PESTIEAU. Fertility, human capital accumulation, and the pension system.
2010/55. Jan JOHANNES, Sébastien VAN BELLEGEM and Anne VANHEMS. Iterative regularization in nonparametric instrumental regression.

Recent titles

CORE Discussion Papers - continued

2010/56. Thierry BRECHET, Pierre-André JOUVET and Gilles ROTILLON. Tradable pollution permits in dynamic general equilibrium: can optimality and acceptability be reconciled?
2010/57. Thomas BAUDIN. The optimal trade-off between quality and quantity with uncertain child survival.
2010/58. Thomas BAUDIN. Family policies: what does the standard endogenous fertility model tell us?
2010/59. Nicolas GILLIS and François GLINEUR. Nonnegative factorization and the maximum edge biclique problem.
2010/60. Paul BELLEFLAMME and Martin PEITZ. Digital piracy: theory.
2010/61. Axel GAUTIER and Xavier WAUTHY. Competitively neutral universal service obligations.
2010/62. Thierry BRECHET, Julien THENIE, Thibaut ZEIMES and Stéphane ZUBER. The benefits of cooperation under uncertainty: the case of climate change.
2010/63. Marco DI SUMMA and Laurence A. WOLSEY. Mixing sets linked by bidirected paths.
2010/64. Kaz MIYAGIWA, Huasheng SONG and Hylke VANDENBUSSCHE. Innovation, antidumping and retaliation.
2010/65. Thierry BRECHET, Natali HRITONENKO and Yuri YATSENKO. Adaptation and mitigation in long-term climate policies.
2010/66. Marc FLEURBAEY, Marie-Louise LEROUX and Gregory PONTHIERE. Compensating the dead? Yes we can!
2010/67. Philippe CHEVALIER, Jean-Christophe VAN DEN SCHRIECK and Ying WEI. Measuring the variability in supply chains with the peakedness.

Books

J. GABSZEWICZ (ed.) (2006), La différenciation des produits. Paris, La découverte.
L. BAUWENS, W. POHLMEIER and D. VEREDAS (eds.) (2008), *High frequency financial econometrics: recent developments*. Heidelberg, Physica-Verlag.
P. VAN HENTENRYCKE and L. WOLSEY (eds.) (2007), *Integration of AI and OR techniques in constraint programming for combinatorial optimization problems*. Berlin, Springer.
P-P. COMBES, Th. MAYER and J-F. THISSE (eds.) (2008), *Economic geography: the integration of regions and nations*. Princeton, Princeton University Press.
J. HINDRIKS (ed.) (2008), *Au-delà de Copernic: de la confusion au consensus* ? Brussels, Academic and Scientific Publishers.
J-M. HURIOT and J-F. THISSE (eds) (2009), *Economics of cities.* Cambridge, Cambridge University Press.
P. BELLEFLAMME and M. PEITZ (eds) (2010), *Industrial organization: markets and strategies*. Cambridge University Press.
M. JUNGER, Th. LIEBLING, D. NADDEF, G. NEMHAUSER, W. PULLEYBLANK, G. REINELT, G. RINALDI and L. WOLSEY (eds) (2010), *50 years of integer programming, 1958-2008: from the early years to the state-of-the-art*. Berlin Springer.

CORE Lecture Series

C. GOURIÉROUX and A. MONFORT (1995), Simulation Based Econometric Methods.
A. RUBINSTEIN (1996), Lectures on Modeling Bounded Rationality.
J. RENEGAR (1999), A Mathematical View of Interior-Point Methods in Convex Optimization.
B.D. BERNHEIM and M.D. WHINSTON (1999), Anticompetitive Exclusion and Foreclosure Through Vertical Agreements.
D. BIENSTOCK (2001), Potential function methods for approximately solving linear programming problems: theory and practice.
R. AMIR (2002), Supermodularity and complementarity in economics.
R. WEISMANTEL (2006), Lectures on mixed nonlinear programming.