Stockouts and Restocking:
Monitoring the Retailer From the Supplier's Perspective

Abstract:

Suppliers and retailers typically do not have identical incentives to avoid stockouts. Thus, the supplier needs to monitor the retailer's restocking efforts with the available data. We introduce a general model for this purpose and illustrate it using a specific application provided by a supplier to a national grocery chain. The model distinguishes between store stockouts (zero inventory in the store) vs. shelf stockouts (an empty shelf, but some inventory in other parts of the store), thereby identifying the cause of the stockout to be either a supply chain or a restocking issue. We find that the average stockout rates vary widely between stores, identifying two stores with stockout rates twice as high as for most other stores. Moreover, almost all stockouts are shelf stockouts. Thus, the model identifies stores that may have restocking issues. Stockouts lead to major losses in expected sales, where on average more than 60% of demand is lost when a stockout occurs. Finally, we also find major difference in shrinkage between stores, providing useful information to regional store managers.

Keywords: Out-of-Stock, Inventory, Bayesian Estimation, Particle Filter

INTRODUCTION

It is well known that stockouts have a major impact on profits (e.g. Gruen and Corsten 2008, Anderson et al. 2006). Given differences in the retailer's and the supplier's profit functions, the retailer's and the supplier's incentives in trying to avoid stockouts are generally not perfectly aligned. Imagine a situation in which a retailer carries two brands of a certain product, and consumers exhibit only weak brand preferences. If one of the brands is out of stock, the consumer will most likely simply buy the other brand rather than go to a different store or wait for his next shopping trip. Campo et al. (2003) conclude that retailers' losses from stockouts can remain very limited, while the suppliers' losses will be substantial. Thus, the retailer may expend less effort than desired by the supplier.

Given this discrepancy in interests, it is essential for the supplier to estimate stockouts in order to monitor the retailer's stocking efforts. Yet, physically checking the shelves in the store is typically expensive. Thus, the supplier needs to deduce the occurrence of stockouts from the available data. More importantly, the supplier also needs to know the root cause of the stockouts. If stockouts happen because there are no products in the store (a "store stockout"), the issue lies within the supply chain and the supplier itself may work on fixing the problem. However, if there is a stockout despite positive inventory in the store (a "shelf stockout"), the problem lies with the retailer who does not restock the shelves quickly enough, requiring a different solution.

In this paper, we present a model that allows the supplier to (a) determine the amount of stockouts and (b) distinguish the two types of stockouts, using only shipment and sales data that is readily available to the supplier. The model consists of two interrelated models, a sales model and an inventory model. The sales model yields daily estimates of stockout probabilities for each product at each store. The inventory model probabilistically tracks daily inventory levels, allowing for unobserved shrinkage. While the amount of stockouts could be estimated using a macro-level framework, the micro-level (i.e., daily or even shorter timeframe) analysis is necessary to link stockouts and inventory levels in order to distinguish the two types of stockouts.

Marketing studies concerning stockouts typically use data with information about product availability on the shelf based on physical checks and focus on estimating the effects of stockouts on demand and the resulting substitution patterns (e.g., Anupindi et al. 1998; Kalyanam et al. 2007; Musalem et al. 2010). In contrast, our model structure is crucial for a context where the supplier monitors the retailer. Instead of relying on information about the occurrence of stockouts as an input to the model, it estimates, among other things, the probability that a stockout occurred on a given day.

The results of our model not only indicate that the stockout rates vary widely across stores but our model also shows sales losses to be quite large when products are out of stock, given that stockouts arise due to unexpected demand shocks. We find that the loss in expected sales when a stockout occurs is on average 60 to 80%, i.e., the stockouts lead to major reductions of the supplier's profits. As for variations in stockouts, while seven out of 10 stores have stockout rates of less than 6%, two stores have stockouts rates of more than 10%. Thus, the model identifies stores that seem to have management issues leading to the higher stockout rates. Over 95% of stockouts are shelf stockouts, a situation that is clearly unacceptable to a supplier.

GENERAL MODEL

In this section, we present the general model applicable to the problem described in the introduction. This framework can be used for any situation in which a supplier needs to monitor the retailer's restocking efforts. Since the purpose of the model is application-driven, the distributional assumptions on the functions of the general model should match the structure of the data generating process of a given application. In contexts as here where there is relatively little observed information, information on the structural relationships between variables is essential to estimate the model. This is akin to the developments in the marketing literature on choice models in which the hypothesized structure of the latent choice process is directly modeled. This allows for insights into the unobserved choice process using only limited amounts of observed data (i.e., choice outcomes). Rather than solely relying on theory, the structure in our model can also come from information supplied by store managers or from inspection of the data, as we show below.

Data Generating Process

Figure (1) gives an overview of the underlying data generating process, presented as the path of an individual product through the system. The process starts with (observed) **shipments** to the individual stores. Total **inventory** can be separated into **backroom inventory** and **shelf inventory**. Backroom inventory refers to the products stored in the back of the store, while shelf inventory refers to the products actually on the shelf at a given time. Obviously, shipped products first become part of the backroom inventory, before turning into shelf inventory through restocking. Positive demand during non-zero shelf inventory leads to observed **sales**. However, if there is positive demand and zero shelf inventory, a **stockout** occurs. The positive demand during a stockout implies that at least one unit of sales was lost due to the stockout. So while there will be no sales during a stockout, stockouts have an impact on observed data by reducing realized sales (as indicated by the dashed line in Figure (1)). Note, however, that the type of stockout makes no difference for the stockout's impact on sales; thus, when talking about the relationship between stockouts and sales, we do not distinguish the two types. Finally, products may also leave inventory without being registered as sold through **shrinkage**. Examples for shrinkage are theft or spoilage.

<< Insert Figure (1) about here >>

Zero Inventory ≠ Stockout & (Shelf) Stockout ≠ Zero Inventory. Intuitively, there seems to be a deterministic relationship between inventory and stockouts. Whenever there is zero inventory, we must have a stockout; and whenever we have a stockout, we must have zero inventory. However, in this paper we distinguish between zero inventory and a stockout.

First, notice that the definition of a stockout above invokes the notion of positive demand, i.e., it is possible to have zero (shelf or total) inventory without observing a stockout. This definition has the advantage of estimating the "extent of out of stocks that actually matter to the retailer and the upstream supply chain members" (Gruen et al. 2002, p. 11) because it focuses on the impact on sales. If there are no units on the shelf, but nobody is trying to purchase the product at that time, neither the retailer's nor the supplier's profits are affected.

Second, a stockout implies zero shelf inventory, but not zero total inventory as there may well be products in the backroom at the time of a stockout. This subtle difference is at the heart of the distinction between store stockouts and shelf stockouts (i.e., the equation Stockout == Zero Inventory is only true when talking about a store stockout).

Stockout ≠ Zero Sales & Stockout ≠ (Sales==Inventory). Similarly, one may intuit a very simple relationship between stockouts and sales, namely that a stockout implies zero sales. While it is true that sales are zero for the duration of a stockout, *daily* sales may not be zero. In particular, there may be a stockout exactly *because* positive demand depleted the available shelf inventory. However, this should not lead one to the erroneous conclusion that a stockout implies that observed sales are equal to previous total inventory. Once again, one needs to realize that there may still be backroom inventory despite a stockout.

Model

For the remainder of the paper, we use the subscripts s, v, and t to refer to the different stores, products(vegetables in our application), and days (time), respectively. However, for the presentation of the general model, we suppress both s and v.

Depending on whether there was a stockout or not, the observed sales are either equal to the unobserved demand or less than the unobserved demand. We thus have

$$SALES_t = \begin{cases} D_t(\mathbf{X_t}, \gamma_t) & \text{with prob. } 1 - \rho_t \\ \\ SO_t(\mathbf{X_t}, \gamma_t) < D_t(\mathbf{X_t}, \gamma_t) & \text{with prob. } \rho_t \end{cases} \tag{1}$$

where $D_t(\cdot)$ is the unobserved demand, while $SO_t(\cdot)$ is the unobserved distribution of sales if there were a stockout (whether or not there actually is a stockout is of course to be estimated). $\mathbf{X_t}$ is a vector of variables affecting demand on a given day. This vector may include marketing variables like price and promotions as well as other factors. γ_t represents any unobserved demand shocks. Obviously, these same factors may also affect sales on a day with a stockout.

The probability of a stockout on day t, ρ_t, is a function of shelf inventory being zero or not (however, as explained in the following paragraph, we model only total inventory) as well as the unobserved demand shocks, i.e.

$$\rho_t = f(\mathcal{I}_{\{INV_{svt}=0\}}, \gamma_t) \tag{2}$$

While zero inventory does not imply a stockout (see above), it certainly makes it more likely. Similarly, large demand shocks may make stockouts more likely, as unexpectedly large demand makes timely restocking less likely.

In order to be able to distinguish between store and shelf stockouts, we need to estimate total inventory as time progresses. Ideally, one would like to distinguish shelf and backroom inventory. However, that is not possible with only sales and shipment data. Thus, we restrict ourselves to estimating total inventory, with the understanding that a stockout implies zero shelf inventory (but not vice versa). In other words, a stockout at a time for which the estimated total inventory is zero is a shelf stockout, whereas a stockout at a time for which the estimated total inventory is positive is a store stockout.

The progression of inventory is governed by the following equation:

$$INV_t = INV_{t-1} + SHIP_t - SALES_t - SHRINK_t \tag{3}$$

In words, inventory at the end of today is equal to yesterday's closing inventory plus today's shipments minus today's sales and shrinkage. We refer to shipments minus sales as the "observed part of inventory" (normalized such that the smallest value is equal to zero). Since total inventory must be non-negative at all times, the joint distribution g of starting inventory INV_0 and the vector of $SHRINK$ needs to ensure non-negative inventory for all days.

Additional insight into the timing of shrinkage may be gained from directly modeling the process of reordering. Note that shipment data also reveal order timing, provided that the lag between order and delivery is known. If the timing of orders is influenced by the current inventory levels (as in a pull strategy), observing orders provides information on inventory levels, which in turn provides information on shrinkage. Thus, we let

$$ORDER_t \sim h(INV_t, \mathbf{Z_t}) \tag{4}$$

where $\mathbf{Z_t}$ is a vector of other variables that may affect the ordering process. These relationships between the variables provide information to inform both model structure and estimation, yielding enough information in our application to obtain estimates that are precise enough for inferences and managerial decisions.

APPLICATION

We now apply this general model to a specific application. The application is motivated by a supplier to a national grocery chain. The supplier suspects high stockout rates in the stores due to the differing incentives to avoid stockouts discussed in the introduction. We first briefly describe the data, before specifying distributional assumptions on the general model based on interviews with store managers and detailed inspection of the data.

Data

The data come from a supplier to a major US grocery store chain. It consists of daily sales and shipment data for 181 days (from May 27, 2007 to November 23, 2007) from 10 grocery stores located in different US states for 17 different products in the frozen vegetable category (for one of the stores we have data for only 16 of the 17 products; thus, in total we have 169 time-series of 181 days each). These products include, for instance, frozen sweet corn, frozen sweet peas, and frozen mixed vegetables. However, the grocery stores carry frozen vegetables not only from this one supplier, but also from a competing supplier.

Summary statistics. Figure (2) is a histogram plot of sales across all stores, products, and days.[1] While the mode of sales is 2 and the mean is 5.05 units, the maximum of observed sales is 64 units of one product in one store on one day, consistent with a category that exhibits high positive demand shocks that could lead to stockouts. Mean daily sales vary between stores and products. Tables (1) and (2) give mean daily sales per store and per product, respectively. Mean daily sales per store range from 2.18 to 7.58. Mean daily sales per product range from 2.11 for Garlic Peas and Mushrooms to 15.34 for Sweet Corn.

<< Insert Figure (2) about here >>

<< Insert Tables (1) and (2) about here >>

Stores receive shipments in multiples of 12, i.e. one case consists of 12 units. Summing over all stores and products, we observe shipments on about one out of three days (36.4%). The vast majority of shipments (89.3%) consist of only one case per product. Another 7.7% are two-case shipments. The largest observed shipments consists of 6 cases. Not surprisingly, average shipments per store and average shipments per product (in units) closely resemble the numbers reported for average sales, as shown in Tables (1) and (2), respectively.

Data preparation. For each day, we observe the number of units sold as well as the dollar revenue per product and store. Thus, we calculate prices as dollar revenue divided by the number of units sold. For days with no units sold, this is undefined. However, since prices rarely change in the course of the time-series for a given product at a given store, we can fill in the missing price data with great confidence.

Moreover, we are missing data for a total of 329 individual store-product-day points (distributed over 15 of the 169 store-product time-series). In addition to this, we also exclude some of the data for the following reason. The model assumes that the shipments follow a pull strategy, i.e. products are ordered when inventory runs low. This is a reasonable assumption for most of the data. However, for part of the data, the shipments are likely the result of a push strategy, i.e. the national headquarters decided to send shipments irrespective of current inventory levels, presumably for some kind of display or promotion. In particular, this happens for five products (across all ten stores) towards the end of the time series. During this period, shipment volumes increase significantly (up to 63 cases per shipment), and consequently occur less frequently. The rare shipments in this period occur on the same day for all stores. Since the inventory model informs the timing and amount of shrinkage by linking the probability of orders to the unobserved inventory and shrinkage, the independence of the shipments from inventory in this period defeats the purpose of the inventory model. Thus, we restrict ourselves to the more common case of a pull strategy and exclude the data from the period of the push strategy. For two products, this affects the last month of the data; for the other three, we must disregard the last two months. In total, this affects 2,704 of the 30,589 store-product-day combinations in our data.

Similar to the promotions at the end of the data, for some stores and products the beginning of the time-series seems to be the week after some sort of promotions just ended. This is evident as significant amounts of inventory are sold off during that time. In this period, no shipments are received during that period despite positive sales. Yet, we do not exclude this data, as those observations still accord with our inventory model in that the lack of shipments in this period is due to high inventory levels (as it would be when following a pull-strategy), not to independence from inventory levels. Except for one product at one store, sales for these periods do not differ from the rest of the data, suggesting that the promotion had already ended. Figure (3) shows the sales pattern for the exception. Since it is obvious that the sales in the first week are noticeably higher than for the remainder of the time despite zero shipments during those days, we assume that this store ran the promotion one week longer than the other stores. We therefore define a dummy variable for promotion that equals one for this one week, store, and product, and zero for all other days, stores, and products.

<< Insert Figure (3) about here >>

Model

Sales. We start the applied version of the model by specifying the distributional assumptions on the two distributions combining to generate observed sales (cf. equation (1)). Sales are commonly modeled using the Poisson distribution (e.g., Anupindi et al. 1998; Neelamegham and Chintagunta 1999). However, as we model the observed sales as a mixture of two unobserved distributions, we cannot test a priori whether a Poisson distribution, in particular its property that the variance is equal to the mean, is appropriate for D_{svt} and SO_{svt}. We therefore choose to model both D_{svt} and SO_{svt} by the more flexible Conway-Maxwell-Poisson (CMP) distribution (Conway and Maxwell 1961; Shmueli et al. 2005). The CMP distribution has two parameters: the first parameter, $\lambda \geq 0$, can be interpreted as a measure of central tendency, while the second parameter, $\nu \geq 0$, governs the level of dispersion. If the decay parameter ν is 1, the CMP reduces to the Poisson distribution; if $0 \leq \nu < 1$ ($\nu > 1$) the CMP distribution has thicker (thinner) tails compared to the Poisson, modeling overdispersed (underdispersed) data relative to the Poisson. Thus, the CMP distribution allows for modeling both over- *and* under-dispersed data; the posterior distribution for the ν parameter indicates whether the commonly used Poisson distribution would have been appropriate.

We then let the unobserved demand be given by

$$D_{svt} \sim \text{CMP}(\lambda^D_{svt}, \nu^D) \tag{5}$$

where λ^D_{svt} is a function of the observed demand shifters $\mathbf{X_{svt}}$ and the unobserved demand shocks γ_{svt}. In particular, $\mathbf{X_{svt}}$ includes variations in price as well as seasonality, weekend, holiday, and promotion effects. We denote these by $PRICE_{svt}$, $SEASON_t$, WE_t, ID_t (Independence Day), LD_t (Labor Day), TH_t (Thanksgiving), and $PROMO_{svt}$, respectively. The measure of central tendency for the D_{svt} distribution is then given by

$$\begin{aligned} \log(\lambda^D_{svt})|\vec{\beta_{sv}}, \gamma_{svt} &= \beta_{0sv} + \beta_{1sv}\log(PRICE_{svt}) + \beta_{2sv}WE_t + \beta_{3sv}\log(SEASON_t) \\ &\quad + \beta_{4sv}ID_t + \beta_{5sv}LD_t + \beta_{6sv}TH_t + \beta_{7sv}PROMO_{svt} + \gamma_{svt} \end{aligned} \tag{6}$$

where

$$\beta_{msv} \sim \text{N}(\bar{\beta}_{msv}, \sigma_m) \text{ for } m = 0, 1, ..., 6 \tag{7}$$

and are mutually independent. We do not impose a hierarchy on β_{7sv} since there is only one store-product for which this parameter is meaningful.

As prices do not change at all for some products at some stores, we normalize the $\log(PRICE_{svt})$ such that its mean is equal to zero for each store-product time series to avoid the problem of separate identification of β_{0sv} and β_{1sv} in those cases. Similarly, the mean of $\log(SEASON_t)$ is normalized to 0, where $SEASON_t$ is a sine curve with wavelength of one year fitted to the average sales per day across all stores and products (excluding Thanksgiving Day as well as the five products for which we excluded the end of the time-series, as they would have biased the estimate for seasonality due to its extraordinarily low and extraordinarily high sales,respectively). $\log(SEASON_t)$ then ranges from -.11 to .22. WE_t, ID_t, LD_t, and TH_t are dummy variables taking the value 1 if day t was a weekend day, Independence Day, Labor Day, or Thanksgiving Day, respectively, and 0 otherwise. Figure (4) shows the average sales per day with the fitted sine curve. Moreover, the weekend spikes in sales amount are obvious, as is the expected low number of sales on Thanksgiving. The impacts of Independence Day and Labor Day on average sales are not as strong. $PROMO_{svt}$ is the dummy variable discussed in the section on data preparation to capture the effect of the promotion.

<< Insert Figure (4) about here >>

If within-brand/across-product substitution effects were expected, equation (6) could be extended to include incorporate those effects. However, the category in our application is known to be one of the least brand-loyal in the U.S. (Mitchell 2011). The grocery chain in our data does carry a second brand of frozen vegetables, making substitution more likely to be between-brand/same-product (leading to the discrepancy of interests between supplier and retailer discussed in the introduction). Since we cannot model between-brand substitution with our data, we do not incorporate substitution effects in our model.

For the model of the unobserved demand shocks, γ_{svt}, we use the same multivariate normal distribution for each daily vector of the 170 store-product shocks, $\vec{\gamma}_t$. We let the variance of this distribution be given by a Kronecker product of two matrices S and V, a 10 by 10 and a 17 by 17 matrix, respectively.[2] This achieves two goals: (1) It reduces the dimensionality, so rather than having to work with a 170 by 170 covariance matrix, we can work with two smaller matrices. And (2) it allows us to introduce some meaningful structure into the random shocks. In particular, S represents covariances across stores, i.e., it captures effects that may affect one or more products similarly across several stores like national advertising campaigns. Similarly, V captures covariances across products, i.e., effects which affect some or all products within the same store. Depending on the product category, these correlations may be caused by differences in weather conditions at the different stores, or other local events like community celebrations. Taking the Kronecker product of these two matrices parsimoniously combines the two sources of covariance for each store-product combination. We thus have

$$\vec{\gamma}_t|\vec{\omega}, S, V \sim \text{MVN}(\vec{\omega}, S \otimes V) \quad \forall t \tag{8}$$

A natural choice for modeling the observed sales in the presence of a stockout would be to let them equal demand minus some lost sales (e.g., following a binomial distribution conditional on demand). If demand were Poisson distributed (i.e., $\nu^D = 1$), this would result in a Poisson distribution for SO_{svt} with parameter $(1-p)\lambda^D_{svt}$ (where p is the probability of the binomial distribution for lost sales). For $\nu^D \neq 1$, this relationship still holds approximately. For computational speed, we therefore choose to model SO_{svt} directly with a CMP distribution exhibiting the same level of dispersion as D_{svt}. In addition, sales should be higher despite a stockout on days on which demand was higher. We therefore also link the first parameters of the SO_{svt} and the D_{svt} distributions by a reduction parameter δ_{sv}. Taken together, we get

$$SO_{svt} \sim \text{CMP}(\delta_{sv}\lambda^D_{svt}, \nu^D) \tag{9}$$

The final piece of equation (1), the probability of a stockout on a given day, is described in equation (2). We use a Beta distribution for function f in equation (2), i.e.

$$\rho_{svt}|a_{svt}, b_\rho \sim \text{Beta}(a_{svt}, b_\rho) \tag{10}$$

where

$$\log(a_{svt}) = a_{0s} + a_{1s}\mathcal{I}_{\{INV_{svt}=0\}} + a_{2s}\gamma_{svt} \tag{11}$$

and

$$a_{0s} \sim \mathrm{N}(\bar{a}_0, \tau_0) \tag{12}$$

Inventory. We now turn our attention to inventory, orders, and shrinkage. Rather than specifying g, the joint distribution for starting inventory and daily shrinkage, directly, we define it by a set of marginal distributions combined with a set of constraints (equations (17) and (18)) to ensure non-negativity of inventory.

The marginal distribution for starting inventory is simply given by a prior based on conversations with several store managers (see next subsection). The situation is slightly more complicated for the marginal distributions of daily shrinkage.

Careful inspection of the data suggests that shrinkage seems to occur on both the unit- and the case-level. While unit shrinkage (e.g., due to theft or misplaced items) is the probably more commonly thought of type of shrinkage, Figure (5) shows a likely instance of case shrinkage. Figure (5) is a time-series of the observed part of inventory for one product at one store; it suggests that a case may have disappeared around the apparent jump at day 100 (and potentially even a second case later in the time series). Causes for case shrinkage may be improperly recorded deliveries or incorrect handling of a case at the point of receipt or restocking (causing a whole case to thaw and go bad). We believe that unit shrinkage is more likely to be caused by the consumer and case shrinkage is more likely to be caused by store employees.

<< Insert Figure (5) about here >>

Total shrinkage on a given day is then given by $SHRINK_{svt} = S_{u,svt} + M \cdot S_{c,svt}$, where the subscripts u and c refer to unit- and case-wise shrinkage, respectively, and $M = 12$ is the number of units per case.

As (unit and case) shrinkage is also discrete and non-negative for all days, the Poisson (or the NDB or the geometric) distribution again would be a standard choice for the marginal distribution. However, due to the non-negativity constraint on inventory, the distribution of shrinkage is actually truncated. Typically, truncation leads to proportionally increased probabilities for the non-truncated values. Since the truncation differs from day to day, depending on the remainder of the time-series, the probabilities of the non-truncated values would be increased by different amounts for different days. This would lead to the paradoxical result that low amounts of shrinkage are more likely on days with low inventory compared to days with high inventory. In order to avoid that, we could implement a zero-inflated truncated distribution, i.e. the sum of the truncated probabilities is added to the probability of zero shrinkage. This keeps the absolute probabilities of positive amounts of shrinkage independent of the available inventory; in turn, the probability of whether or not shrinkage occurs at all depends on inventory, with shrinkage being less likely for lower amounts of inventory. Therefore, we choose to implement a slightly more complex distribution for shrinkage.

In particular, we model the marginal distribution for both unit and case shrinkage in two parts. The first part, a Bernoulli distribution, models the occurrence of shrinkage. Shrinkage is likely to be very rare. Yet, when it occurs it is caused by some event (e.g., a thief being in the store or a store employee being distracted) that may affect more than just one unit or case. The second part, a CMP-binomial distribution, models the amount of shrinkage, given that shrinkage occurred. The CMP-binomial distribution can be derived from the CMP distribution and can be viewed as the result of correlated Bernoulli draws (rather than independent Bernoulli draws which would lead to the standard binomial distribution; see Shmueli et al. 2005). It has three parameters: the number of draws, the probability of a success, and a parameter governing the association between the draws. If this third parameter is less than 1, the Bernoulli draws are positively correlated; if it is greater than 1, the Bernoulli draws are negatively correlated. If it is 1, the Bernoulli draws are independent, and the CMP-binomial distribution reduces to the standard binomial distribution.

The Bernoulli distribution keeps the probability of shrinkage independent of inventory, while the CMP-binomial distribution keeps the (marginal) probability of each individual available unit to be shrunk constant. The relative and absolute probabilities of different amounts of shrinkage then vary with inventory. We think that it is more appropriate in this context to keep (marginal) individual probabilities constant than to keep either relative or absolute probabilities of the overall amount of shrinkage constant.

We thus have

$$S_{c,svt}|b_{c,s}, n_{c,svt}, p_{c,s}, \nu_c \sim \text{Bernoulli}(b_{c,s}) \cdot (\text{CMP-Binomial}(n_{c,svt}, p_{c,s}, \nu_c) + 1) \tag{13}$$

$$S_{u,svt}|b_{u,s}, n_{u,svt}, p_{u,s}, \nu_u \sim \text{Bernoulli}(b_{u,s}) \cdot (\text{CMP-Binomial}(n_{u,svt}, p_{u,s}, \nu_u) + 1) \tag{14}$$

where the two Bernoullis are correlated with η_s. This correlation allows for the possibility that one event can cause both unit and case shrinkage simultaneously. The number of draws for the CMP-binomial are given by

$$n_{c,svt} = \max(0, \text{floor}(\frac{I\hat{N}V_{svt}}{M}) - 1) \tag{15}$$

$$n_{u,svt} = \min(M - 2, I\hat{N}V_{svt} - 1) \tag{16}$$

where $I\hat{N}V_{svt}$ is the estimate of inventory before the occurrence of shrinkage for this day. The -1 in equation (15) is due to the +1 in equation (13), i.e. the fact that if shrinkage occurs, at least 1 case is shrunk. Similarly, the -2 (and the -1) in equation (16) is partly due to the +1 in equation (14), but also to the fact that, for simplicity and identification, we assume that the maximum amount of unit shrinkage is $M - 1$ (otherwise it will be counted as case shrinkage).

Note that the parameters governing the probability of shrinkage occurring are defined on a store-level. Thus, the model allows separating out stores that may have greater issues with shrinkage.

These marginal distributions could result in negative inventory on some days. We therefore impose the following constraint on the resulting joint distribution to ensure non-negative inventory for all days:

$$\sum_{1}^{t}(S_{u,svt} + M \cdot S_{c,svt} + SALES_{svt}) \leq INV_{0,sv} + \sum_{1}^{t} SHIP_{svt} \quad \forall s, v, t \tag{17}$$

However, for some days, we impose a stronger constraint, reflecting the belief that inventory was very likely equal to zero on those days. This addition to the model again was motivated by inspection of the observed part of inventory. See Figure (6) for an example of a sales pattern and the corresponding observed part of inventory. The combination of zero sales and the flat line in observed inventory from day 15 to 25 strongly suggests that inventory in this store was zero on those days (and the store managers agreed). Yet, since the observed part of inventory is one unit lower than the flat line on day 7, this implies that there must have been at least one unit and/or one case (if starting inventory was higher than in our plot) of shrinkage between day 8 and day 15. Thus, this stronger constraint can greatly help the estimation of the occurrence of shrinkage. We determine the days for which to use the stronger constraint by the following fairly strict criteria:

1. The future observed part of inventory must be non-decreasing.

2. We observe at least two weeks of future data.

3. The observed part of inventory is not higher than on the previous day.

4. We observe at least two consecutive days with zero sales.

<< Insert Figure (6) about here >>

Condition (1) is a necessary condition for zero inventory on a given day. Condition (2) increases our confidence in asserting zero inventory for a certain day by making sure that we observe enough future data to make condition (1) meaningful. Condition (3) rules out days on which we know a shipment was received that may satisfy the other conditions. And finally, condition (4) further tightens the criteria by requiring that the effect persists for at least two days, thereby resulting in a visible flat line in a plot of the observed part of inventory, to make it less likely that the observed pattern is just a coincidence of one day with zero sales and zero shipments. Based on these criteria, we have 173 days for which to include the stronger constraint. Referring back to Figure (5), notice that the flat line before day 100 also satisfies the above criteria (as sales are zero for three consecutive days), implying that shrinkage in the preceding days is very likely.

Let all days satisfying the above criteria for store s and product v be denoted by T^*_{sv}. The stronger constraint for those days still makes sure that inventory was non-negative, but it also reflects the belief that inventory most likely was zero on those days:

$$\sum_{1}^{T^*_{sv}}(S_{u,svt}+M\cdot S_{c,svt}+SALES_{svt})\begin{cases} < INV_{0,sv}+\sum_1^{T^*_{sv}} SHIP_{svt} \text{ with prob. } 1-\pi_{svT^*_{sv}} \\ = INV_{0,sv}+\sum_1^{T^*_{sv}} SHIP_{svt} \text{ with prob. } \pi_{svT^*_{sv}} \end{cases} \quad \forall T^*_{sv} \text{ and } \forall s,v \tag{18}$$

The top of the right hand side of equation (18) is the same non-negativity constraint as in equation (17), while the lower part results in zero inventory on T^*_{sv}.

As discussed in the general model, we can use the placement of orders for further information on the exact timing of shrinkage. (Moreover, modeling the ordering process is necessary to estimate the missing data.) From talking to several local managers of the national grocery chain, we know that the lag between the placing of the order and the receipt of the shipment is typically two days. Therefore, the number of cases ordered on a given day is

$$ORDER_{svt}=\frac{SHIP_{sv,t+2}}{M} \tag{19}$$

From inspection of the data, it is obvious that the distribution of orders is highly underdispersed relative to a Poisson distribution. We therefore again use the more flexible CMP distribution and let the distribution of orders be given by

$$ORDER_{svt}|\lambda^O_{svt}, \nu^O \sim \text{CMP}(\lambda^O_{svt}, \nu^O) \tag{20}$$

As argued above, λ^O_{svt} then depends on inventory, providing information on shrinkage. In particular, we include a variable calculated from today's inventory plus tomorrow's shipment (which is already ordered and therefore known) relative to the average sales (denoted by $\overline{SALES_{sv}}$) in the two days needed for the shipment to arrive. Moreover, λ^O_{svt} also depends on the lagged dummies for the seasonal trend and for weekends (as defined above), as higher orders should be expected when the store expects a demand spike two days later. The changes of expected demand level and their effect on orders is captured by these two variables. Thus we have

$$\log(\lambda^O_{svt}) = \kappa_{0sv} - \kappa_{1sv}\frac{INV_{svt} + SHIP_{sv,t+1}}{2 \cdot \overline{SALES_{sv}}} + \kappa_{2sv}\log(SEASON_{sv,t+2}) + \kappa_{3sv}WE_{sv,t+2} \tag{21}$$

and

$$\kappa_{lsv} \sim \text{N}(\bar{\kappa}_l, \varsigma_l) \quad \text{for } l = 0, 2, 3 \tag{22}$$

and

$$\kappa_{1sv} \sim \text{Gamma}(\bar{\kappa}_1, \varsigma_1) \tag{23}$$

Notice the gamma-hierarchy on κ_{1sv}; this ensures that the effect of inventory on orders is negative, as it should be. See the Limitations section below for a discussion of this restriction.

Priors. Table (3) gives an overview of the priors we choose for the parameters.

<< Insert Table (3) about here >>

<u>Sales:</u> For $\bar{\beta}_m$ and σ_m (for $m = 0, 1, ..., 6$) we specify normal distributions with mean 0 and variance 100 and an inverse gamma distribution with parameters 2.5 and 2.5, respectively, as uninformative priors. The prior for β_7 for the one store-product combination with the promotion at the beginning of the time-series is also a normal distribution with mean 0 and variance 100. Since $\exp(\gamma_{svt})$ has a multiplicative effect on λ_{svt}, we can reasonably expect that it is closely clustered around 1. In equation (8), we specify a multivariate normal distribution for γ_{svt} which is equivalent to a multivariate log-normal distribution for $\exp(\gamma_{svt})$. Since the mode of a log-normal distribution is equal to the exponent of its mean minus its variance, we choose a prior such that, in expectation, the mean is equal to the variance. In addition, the variance should be fairly small since very large shocks are unlikely. We thus let the prior for $\vec{\omega}$ be a multivariate normal distribution with mean $.0\vec{0}1$ and a variance conditional on S and V given by the $.001 S \otimes V$. S and V are distributed according to inverse Wishart distributions that combine, in expectation, to values of .001 on the diagonal of their Kronecker product. We let the parameters for the inverse Wishart distribution for S be given by 150 and $4 \cdot R_S$, where R_S is a correlation matrix with .2 for all off-diagonal elements. Similarly, for V the parameters of the inverse Wishart distribution are 150 and $4 \cdot R_V$, where R_V is a correlation matrix with .1 for all off-diagonal elements. The difference in off-diagonal elements reflects our belief that shocks affecting sales of the same product across stores are slightly more likely than shocks that affect different products in individual stores. Finally, while this choice of priors for $\vec{\omega}$, S, and V does not retain conjugacy for S and V, it still allows for simple Gibbs drawings of S conditional on V and vice versa.

As prior for ν^D, we choose a gamma distribution with parameters 2 and 1, resulting in a mode of 1 (which would reduce the CMP distribution to a Poisson distribution) but also allowing over- and under-dispersion of the data. For δ_{sv} for all s and all v, we choose a beta distribution with parameters 2 and 2.5. Thus, high values for δ_{sv} are unlikely, fitting the idea that there should be an impact on sales for an empty shelf to be counted as a stockout.

Finally, for the priors related to ρ_{svt}, the probability of stockouts, we let b_ρ follow a fairly diffuse gamma distribution with parameters 2 and 1. Since Gruen and Corsten (2008) and Andersen Consulting (1996) find that average rates of stockouts should not be significantly greater than 10%, we choose a prior for $\bar{a}_0$ that, while allowing for some uncertainty, ensures that $\exp(a_{0sv})$ is small relative to b_ρ, namely a normal distribution with mean -2.5 and variance .5. The variance τ_0 follows an inverse gamma distribution with parameters 4 and 2. For a_{1sv} and a_{2sv} (for all s and v) we choose gamma distributions, as both should definitely have a positive effect. For a_{1sv}, we choose parameters 30 and .125; for a_{2sv}, we choose parameters 10 and 3 (recall that all γ_{svt} will be fairly small in absolute value).

Inventory: We believe that shrinkage should be fairly rare, with unit shrinkage being somewhat more likely than case shrinkage. Thus, we let the priors for $b_{u,s}$ and $b_{c,s}$ be beta distributions with parameters $(1, 20)$ and $(1, 30)$, respectively. We use a uniform unconditional prior for the correlation between unit and case shrinkage. (Note that conditional on $b_{u,s}$ and $b_{c,s}$ the correlation may not be able to take all values from -1 to +1.) For the CMP-binomial distribution, we let the probability for each unit or case to be shrunk, i.e. $p_{u,s}$ and $p_{c,s}$, be distributed according to a beta distribution with parameters 1 and 20 (all of the above specifications are independently for all s). For the measure of correlation between the Bernoulli draws, we use a gamma distribution with parameters 11 and .1 for both ν_u and ν_c. Thus, a standard binomial distribution (i.e., no correlation between the Bernoulli draws) is the prior mode ($\nu = 1$), but the prior also allows for positive and negative correlation between the Bernoulli draws.

As mentioned above, the criteria for the days on which we impose the stronger constraint on the joint distribution for starting inventory and shrinkage are fairly strict, i.e. we are very confident that inventory was very likely zero. Moreover, this probability of zero inventory, $\pi_{svT^*_{sv}}$, is set against the very small probabilities of shrinkage which also means that it needs to be fairly large to be effective. Finally, the more days we observe after a day satisfying the the criteria, the more confident we are that inventory actually was zero on that day. These reasons lead to a beta distribution with parameters $d_{svT^*_{sv}}$ and 1 as a prior, where $d_{svT^*_{sv}}$ is the number of days observed after T^*_{sv} until the next missing data point or the end of the data.

We use the same uninformative priors for $\bar{\kappa}_n$ and ς_n for $n = 0, 2, 3$ as for $\bar{\beta}_m$ and σ_m in the sales model, namely a normal distribution with mean 0 and variance 100 and an inverse gamma distribution with parameters 2.5 and 2.5, respectively. For the gamma-hierarchy on κ_{1sv}, we specify two gamma priors, with $\bar{\kappa}_1$ and ς_2 both following a gamma distribution with parameters 4 and .5. For ν^O we use a gamma prior with parameters 2 and 2, reflecting the finding from inspection of the data that orders seem to be underdispersed relative to a Poisson distribution.

Finally, based on discussions with store managers, we choose a prior for starting inventory with most of its mass around the size of one case, but also allowing for some variation around it. Rather than using a Poisson distribution or the NBD, we choose a truncated and discretized normal distribution. Since we observe the tail ends of a promotion for some products with inventory being sold off (see section on data preparation), starting inventory has to be in the tail of the marginal distribution in those cases in order to assure non-negative inventory for all other days. We prefer the discretized normal distribution to the Poisson and the NBD as its tail decreases more quickly, thereby still providing meaningful information for starting inventory in those instances. We discretize the normal distribution by letting the probability for integer x be given by the difference between the normal cdf at $x + 0.5$ and the normal cdf at $x - 0.5$. Let $\mathrm{DT}_y\mathrm{N}$ denote this discretized normal distribution, truncated from below at y. The prior for starting inventory (for all stores and products) then is $\mathrm{DT}_0\mathrm{N}(M, (.75M)^2)$.

ESTIMATION

The model defined in equations ((5)) to ((23)) is estimated using an MCMC algorithm (Casella and George 1992; Gelfand and Smith 1990). We define a binary auxiliary variable z_{svt} to denote the occurrence of a stockout (Tanner and Wong 1987). Note that with several parameters having to be estimated for all store-product-day combinations (ρ_{svt}, z_{svt}, γ_{svt}, $S_{u,svt}$, $S_{c,svt}$) we have far more than 150,000 parameters to estimate! This is feasible thanks to the hierarchical structure of the model, resembling a frequentist random effects model. After a burn-in period of 5,000 draws, we use 45,000 draws for the estimation. Convergence was checked by the Heidelberger and Welch convergence diagnostic as implemented in the **boa** package for R (Heidelberger and Welch 1983; Smith 2007).

The hierarchical parameters in equations (7), (12), and (22) are estimated in a Gibbs-sampler fashion, as they have either conjugate priors, or priors leading to a closed-form conditional distribution (for S and V). Most of the remaining parameters are estimated with a Metropolis-Hastings sampler.

However, the inventory trajectory, i.e. starting inventory, shrinkage, and missing data, are estimated using a particle filter (see Arulampam et al. (2002) for a tutorial). In essence, we construct a discrete representation of the posterior distribution conditional on the current draw for all other parameters and then take a random draw from it. We implement this approach because traditional Metropolis-like methods showed very slow mixing behavior and/or took very long to compute. The particle filter results in a new draw from the conditional posterior joint distribution for starting inventory, shrinkage, and the missing sales and shipment data in every step of the MCMC. Thus, we can simply use it as a drawing method in our Gibbs-sampler estimation.

The discrete representation of the posterior is achieved by recursively filtering a set of n so-called particles according to importance weights updated in a Bayesian fashion. For our application, this means we start by drawing the starting inventory for all particles and update the weights accordingly, then we draw shrinkage (and the potentially missing data) for day 1 and update the weights, then repeat those steps for day 2, and so on. The updating of the weights follows a similar rationale as is implemented in the Metropolis-Hastings sampler. Let $\phi(\cdot)$ denote the proposal density from which shrinkage and, if missing, the missing data (or starting inventory for initializing the particles) are drawn. Let $\psi_t(S_{u,svt}, S_{c,svt}, ORDER_{svt}, SALES_{svt}|\cdot)$ denote the probability of observing $S_{u,svt}$, $S_{c,svt}$, $ORDER_{svt}$, and $SALES_{svt}$ conditional on the other model parameters and the inventory trajectory of that particular particle up to t, as defined by the model. Then the weight for particle k after drawing day t is given by updating its weight after day $t-1$ by

$$w_{k,t} = w_{k,t-1}\frac{\psi_t(\cdot|\cdot)}{\phi(\cdot)} \tag{24}$$

Thus, the greater the probability of the data as defined by the model, the greater the particle's weight; at the same time, the higher the probability of drawing the certain values, the lower the weight. This is very similar in spirit to the Metropolis-Hastings algorithm. The discrete approximation to the posterior distribution is then given by normalizing the weights such that they sum to 1.

In order to avoid degeneracy of the algorithm (i.e., the problem that only one particle may have significant weight compared to all the others) we implement a sequential importance sampling (SIS) particle filter with resampling (Carpenter et al. 1999). In the SIS with resampling algorithm, after each updating of the weights an estimate of the effective sample size (i.e., the number of particles with non-neglegible relative weights), N_{eff}, is calculated by

$$N_{eff} = \frac{(\sum_k w_{k,t})^2}{\sum_k w_{k,t}^2} \tag{25}$$

(Liu and Chen 1998). If N_{eff} falls below a pre-specified threshold, a new set of n particles is resampled randomly from the previous set of particles proportionally to their weights (also, the weights are reset to $1/n$ for all particles after resampling). We implement this approach with $n = 150$ particles and resample whenever $N_{eff} < 15$.

In a final step, we randomly draw one of the particles representing the approximation to the posterior distribution.

RESULTS

The results not only confirm that the suspicion of the supplier was at least partially justified, but also show that each stockout has an immense impact on the supplier's profit.[3] The average rate of stockouts across stores, products, and days estimated by the model is 6.37%.[4] This rate is below the average rate of about 8% reported by previous studies (Gruen and Corsten 2008; Andersen Consulting 1996); however, those studies are based on all instances when no units are on the shelf, whereas our model counts only the instances in which sales are actually lost. More interestingly, though, we also find a wide range of stockout rates across stores. Seven of the ten stores actually have an average stockout rate of less than 6%, the best being only 3.94%; in contrast, two stores have mean stockout rates of over 10% (10.62% and 11.79%)! This clearly suggests that these stores have significant potential for improvement.

From the posterior estimates for δ_{sv} we can derive how much of the unobserved demand is lost on average for the supplier due to a stockout. On average, δ_{sv} is .44, translating to a decrease in expected sales of 60 to 80% for each stockout. The posterior means for individual store-product combinations range from .05 to .76. This means that in the worst case expected sales decrease by 98%, but even in the best case expected sales drop by 39% due to a stockout![5]

This strong impact is due to two factors: First, we defined stockouts such that at least one sale must be lost due to the empty shelf. For the slowly moving products with on average only two or three sales per day, one lost sale obviously is a large percentage lost in sales. However, this reasoning does not explain the high percentages for faster moving products.

Second, then, is the finding that more positive demand shocks lead to a higher chance of a stockout (a_{2s} from 6.99 to 17.98). Thus, unobserved demand tends to be particularly high on days with a stockout, which in turn implies that more sales are lost due to a stockout. This endogeneity of demand and the occurrence of a stockout is not unexpected, but it lends additional urgency to the supplier's quest to avoid stockouts.

Moreover, as to be expected, the probability of a stockout increases significantly if inventory is 0 on a given day (a_{1s} from 1.37 to 1.89). Yet, the average stockout rate for days with zero inventory is only 34.96%. In addition to our definition of stockouts as instances with lost sales, this is mainly due to the fact that our inventory measure is only a snapshot at midnight every day. Thus, on a day with zero inventory in our estimation the store may have had inventory (and therefore sales) until just before midnight. We are confident that zero inventory would have a stronger impact on the probability of stockouts if data on shorter time intervals were available. However, since we have only daily data, we cannot confirm this conjecture.

Somewhat surprisingly, only 4.46% of the stockouts are store stockouts, while the remainder are shelf stockouts. This rate of store stockouts is significantly lower than what has been reported in the literature so far (Gruen and Corsten 2008). However, this may be explained by the short lead times (relative to case-size and average daily sales). Given that daily sales are only a fraction of case-size for most products, a store should generally be able to order and receive more units before running out of units completely, thereby avoiding store stockouts.

To gain more understanding of what makes a product more likely to stock out, we look at the effect of the coefficient of variation of sales (defined as the ratio of the variance of sales to the mean of sales). We find that it is related to both the amount of stockouts as well as the ratio of store to shelf stockouts. Not surprisingly, the higher the relative variance, the higher the stockout rate (correlation coefficient = .343). Obviously, it is easier to avoid stockouts when sales are more regular. Similarly, the ratio of store to shelf stockouts increases with higher relative variance of sales (correlation coefficient = .299), i.e. a larger share of the stockouts are not due to poor shelf replenishment, but to the lack of units in the store. Due to the higher relative variance, it is more likely that an unexpected spike in demand occurs that depletes inventory before new shipments can arrive. This supports the argument above that, absent a large positive demand shock, lead times are short enough to ensure stock in the store.

To illustrate the results of the model, we refer back to the product and store whose observed part of inventory was depicted in Figure (5). Figure (7) shows the corresponding sales pattern and the daily mean probability of stockout as estimated by the model. As a simple inspection of the sales pattern may have already suggested, the model finds that a stockout was most likely during the few days before day 100 with several zero sales days in a row (not surprisingly, the model also estimates a high probability of zero inventory for those days, leading to store stockouts). The results are also reasonable in that in general stockouts are more likely for days with lower sales. However, the model has better discriminating abilities than these simple intuitions. For instance, the estimated probabilities of stockouts for days with zero sales is on average 60%, but ranges from only 15% to an almost certain 95%. For days with sales of one unit, stockouts are significantly less likely (25% on average), yet specific days still have probabilities of stockouts of over 75%. Even days with two or more sales may still have probabilities of over 25%, i.e. on some days with more sales stockouts are estimated to be much more likely than on some days with no sales at all.

<< Insert Figure (7) about here >>

Shrinkage

While shrinkage is of less interest to the supplier, the results provide some valuable insights for the retailer's store managers. As with stockouts, we find a fairly wide variation of shrinkage levels across stores. Posterior means for $b_{u,s}$ range from .26% to 5.65%, i.e. from less than one incident of unit-shrinkage per year to more than 20 incidents per year. Similarly, for $b_{c,s}$ the means range from .10% to .56%; so while case shrinkage is reasonably unlikely for all stores, it is more than five times as likely in some stores than in others! Moreover, the posterior means for $b_{u,s}$ and $b_{c,s}$ are strongly correlated (correlation coefficient = .61), showing that the same stores that have problems with unit shrinkage also have problems with case shrinkage. The correlations between case and unit shrinkage range from .0666 to .3195.

Finally, to illustrate the estimation results for shrinkage provided by the particle filter, we again refer back to the store and product depicted in Figure (5). Remember that we believed that true inventory was likely zero at the flat line before day 100, implying at least some sort of shrinkage. While the particle filter estimates essentially no case shrinkage up to the flat line, the probability of case shrinkage increases to almost 30% on day 98 and to about 20% on day 99. After that, the probability for case shrinkage decreases sharply again, yet remains slightly elevated for the next 10 days relative to the rest of the days. Overall, we estimate at least one case of shrinkage for this particular store and product in more than 80% of all iterations of the MCMC. About 30% of the draws result in two cases of shrinkage. Similarly, the probability of unit shrinkage increases slightly before day 98 and then spikes very clearly on day 98. This shows how the stronger constraint on inventory can help with the estimation of the timing of shrinkage.

Similarly, for the example given in Figure (6) we expected to observe shrinkage between day 8 and the flat line starting at day 15. In almost 70% of the draws of the MCMC we observe unit and/or case shrinkage in that one week, and in more than 85% of the draws in the time from day 8 to day 20 (as the first few instances of zero sales may also have been coincidences).

Limitations

As mentioned above, the inventory model assumes that orders are the result of a pull-strategy, i.e., the amount ordered is inversely related to the current inventory. For this reason, we excluded some data for which this assumption was clearly invalid (see section on data preparation), and then estimated the missing data. However, for a few of those store-product combinations, the assumption seems not to have been satisfied even before this apparent change in ordering strategy. For those instances, the correlations between the relative inventory measure we use in equation (21) and the placed orders are close to zero, i.e. orders are independent of inventory (if lead time was indeed two days). With orders being independent of inventory, there is nothing in the model to prevent very high levels of inventory (since κ_{1sv} will tend towards zero). In general, this should not be a problem since the other parameters in equation (21) should be calibrated to result only in appropriately low amounts of orders. However, keep in mind that (1) the data we excluded was at the end of the time-series and (2) that the variable for the seasonality increases with time but shows relatively little variation earlier in the time-series (see Figure (4)). Due to the little variation of seasonality for the days with non-excluded data, a wide range of values for the coefficient associated with seasonality in equation (21) (i.e., κ_{2sv}) can fit the observed data; yet, highly positive values lead to high estimated orders for the excluded data. Once high orders are estimated for the excluded data (i.e., at the end of the time-series), this reinforces the positive estimates for κ_{2sv}. With orders not depending on inventory levels, we therefore estimate inventory levels for the last few days in those instances up to double of what we observe earlier in the time series, which seems unlikely.

This issue is also the reason for specifying a gamma hierarchy for κ_{1sv} in equation (23). This ensures that the effect of inventory on orders can at least not be positive. Otherwise, the effect described above could spiral even further out of control. In particular, the higher estimated orders despite high inventory levels would lead to positive estimates for κ_{1sv}. However, once κ_{1sv} is positive, the higher inventory levels result in higher estimates for orders, which results in higher inventory levels, etc.

Thus, it is critical for the performance of the model that the ordering process in fact follows the assumed pull-strategy. Moreover, knowing the exact lead times between orders and shipments is crucial. For instance, if lead times for these store-product combinations were not two always days, as claimed by the store managers, that might cause this issue even despite a pull-strategy being in place. Thus, we believe that with more accurate information, this problem could be avoided. Moreover, note that this is not an issue with complete data, i.e., if there is no missing data to be inferred.

DISCUSSION

Based on the results of our model, the concern of the supplier who provided us with the data seems justified. Even though stockouts are relatively infrequent, stockouts tend to occur when demand is unusually high, and the model estimates reveal large losses for the supplier. As for stockout rates, they are significantly higher than what is acceptable to a supplier at least for two of the stores. Both the very small amount of store stockouts and the fact that other stores have significantly lower stockout rates suggest that this is a store management (i.e., restocking) issue rather than a problem with timely re-ordering or with the supply chain. This is useful information to the supplier, as now he can address those stores specifically.

The model is also useful to the regional or national management level of the grocery chain, as it now can distinguish well-managed from not so well-managed stores. Choosing the right amount of inventory is a tricky decision; more inventory leads to higher inventory holding costs, but lower inventory leads to higher chance of stockouts. Foo (2007) reports that companies may classify products into different tiers and set target service levels (or alternatively target stockout rates) for each of them. Obviously, managers choose those levels they consider to be profit maximizing. Therefore, any deviation (no matter in which direction!) is detrimental to profits. With our model, management can check whether the individual retail stores adhere to the pre-specified values.

In further research, it would be interesting to conduct the same analysis for different products in the same stores to see whether the stockout problems in the less well-managed stores persist throughout categories. Alternatively, it could be due to the misalignment of interests between suppliers and retailers mentioned in the introduction, i.e. the retailer has less incentives to avoid stockouts if there is an alternative brand and switching costs are low. In this case (1) the stockout rates for the alternative brand should be low (or stockouts of the two brands at least not overlap) and (2) stockout rates for categories with high brand loyalty should be low.

References

Andersen Consulting (1996). *Where to Look for Incremental Sales Gains: The Retail Problem of Out-of-Stock Merchandise*. The Coca-Cola Research Council, Atlanta, GA.

Anderson, E. T., Fitzsimons, G. J., and Simester, D. (2006). Measuring and mitigating the costs of stockouts. *Management Science*, 52(11):1751–1763.

Anupindi, R., Dada, M., and Gupta, S. (1998). Estimation of consumer demand with stock-out based substitution: An application to vending machine products. *Marketing Science*, 17(4):406–423.

Arulampam, S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–191.

Bijmolt, T. H. A., Heerde, H. J. V., and Pieters, R. G. M. (2005). New empirical generalizations on the determinants of price elasticity. *Journal of Marketing Research*, 42:141–156.

Campo, K., Gijsbrechts, E., and Nisol, P. (2003). The impact of retailer stockouts on whether, how much, and what to buy. *International Journal of Research in Marketing*, 20:273–286.

Carpenter, J., Clifford, P., and Fearnhead, P. (1999). Improved particle filters for nonlinear problems. *Institution of Electrical Engineers Proceedings - Radar, Sonar Navigation*, 146(1):2–7.

Casella, G. and George, E. I. (1992). Explaining the Gibbs sampler. *American Statistician*, 49(4):327–335.

Conway, R. W. and Maxwell, W. L. (1961). A queuing model with state dependent service rate. *Journal of Industrial Engineering*, 12:132–146.

Foo, Y. N. (2007). A diagnostic analysis of retail out-of-stocks. Master's thesis, Massachusetts Institute of Technology,.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85:398–409.

Gruen, T. W. and Corsten, D. S. (2008). *A Comprehensice Guide to Retail Out-of-Stock Reduction in the Fast-Moving Consumer Goods Industry*. Grocery Manufacturers of America, Food Marketing Institute, National Association of Chain Drug Stores, The Procter & Gamble Company, University of Colorado at Colorado Springs.

Gruen, T. W., Corsten, D. S., and Bharadwaj, S. (2002). *Retail Out-of-Stocks: A Worldwide Examination of Extent, Causes, and Consumer Responses*. Grocery Manufacturers of America, Washington, DC.

Heidelberger, P. and Welch, P. (1983). Simulation run length control in the presence of an initial transient. *Operations Research*, 31:1109–1144.

Kalyanam, K., Borle, S., and Boatwright, P. (2007). Deconstructing each item's category contribution. *Marketing Science*, 26(3):327–341.

Liu, J. S. and Chen, R. (1998). Sequential monte carlo methods for dynamical systems. *Journal of the American Statistical Association*, 93:1032–1044.

Mitchell, R. (2011). Private eyes. *Refrigerated & Frozen Food Retailer*, June:8.

Musalem, A., Olivares, M., Bradlow, E. T., Terwiesch, C., and Corsten, D. S. (2010). Structural estimation of the effect of out-of-stocks. *Management Science*, 56(7):1180–1197.

Neelamegham, R. and Chintagunta, P. (1999). A bayesian model to forecast new product performance in domestic and international markets. *Marketing Science*, 18(2):115–136.

Shmueli, G., Minka, T. P., Kadane, J. B., Borle, S., and Boatwright, P. (2005). A useful distribution for fitting discrete data: Revivel of the Conway-Maxwell-Poisson distribution. *Journal of the Royal Statistical Society, Series C*, 54(1):127–142.

Smith, B. J. (2007). boa: An R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software*, 21(11):1–37.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398):528–540.

FOOTNOTES

[1] All statistics reported in this section exclude certain day-store-product combinations, as explained in the next subsection.

[2] This implies that these shocks are also estimated for the missing store-product combination. We therefore treat the complete time series as missing data to be estimated in the MCMC.

[3] We focus on the most interesting results concerning stockouts and shrinkage in the main body of the paper. See the appendix for results of the remaining parameters.

[4] For the stockout rates, we report the means of the auxiliary variable z_{svt} indicating the occurrence of a stockout, rather than the mean of ρ_{svt}; however, results are very close.

[5] We compute these values using the average of the unobserved demand (on a non-holiday weekday). Due to the non-linear nature of the CMP distribution, these values are only approximate at other levels of demand.

Table 1: Mean Daily Sales and Shipments by Store

Store	1	2	3	4	5	6	7	8	9	10
Sales	5.36	2.49	6.02	5.70	5.67	2.18	4.09	7.58	6.92	4.62
Shipments	5.28	2.50	6.07	5.68	5.70	2.10	4.00	7.48	6.89	4.62

Table 2: Mean Daily Sales and Shipments by Product

Product	Sales	Ship.	Product	Sales	Ship.
Sweet Corn	15.34	15.27	Asparagus and Gold&White Corn	3.86	3.95
Broccoli Cut	8.43	8.31	Brocc./Carrots/Peas/Chestnuts	3.82	3.61
Mixed Vegetables	7.92	7.94	Whole Green Beans	3.63	3.56
Sweet Peas	7.71	7.72	Baby Brussel Sprouts	3.56	3.43
Brocc./Caulifl./Carrots	6.58	6.61	Corn on the Cob	3.17	3.14
Green Beans	5.87	5.82	South-West Corn	2.55	2.61
Broccoli and Cauliflower	4.76	4.79	Garlic Cauliflower	2.38	2.40
Broccoli Florets	4.28	4.08	Garlic Peas and Mushrooms	2.11	2.15
Asian Medley	4.19	4.27			

Table 3: Overview of Prior Specifications

Sales		Inventory	
$\bar{\beta}_m$ for $m = 0, 1, ..., 6$	N(0,100)	$b_{u,s}$ $\forall s$	Beta(1,20)
σ_m for $m = 0, 1, ..., 6$	Inv. Gamma(2.5,2.5)	$b_{c,s}$ $\forall s$	Beta(1,30)
β_7	N(0,100)	η_s $\forall s$	Uniform(-1,1)
$\vec{\omega}\|S, V$	MVN$(.\vec{00}1, .001 S \otimes V)$	$p_{u,s}$ $\forall s$	Beta(1,20)
S	Inv. Wish.$(150, 4 \cdot R_S)$	$p_{c,s}$ $\forall s$	Beta(1,20)
V	Inv. Wish.$(150, 4 \cdot R_V)$	ν_u	Gamma(11,.1)
ν^D	Gamma(2,1)	ν_c	Gamma(11,.1)
δ_{sv} $\forall s, v$	Beta(2,2.5)	$\pi_{svT^*_{sv}}$ $\forall T^*_{sv}$ and $\forall s, v$	Beta$(d_{svT^*}, 1)$
$\bar{a}_0$	N(-2.5,.5)	$\bar{\kappa}_l$ for $l = 0, 2, 3$	N(0,100)
τ_0	Inv. Gamma(4,2)	ς_l for $l = 0, 2, 3$	Inv. Gamma(2.5,2.5)
a_{1sv} $\forall s, v$	Gamma(30,.125)	$\bar{\kappa}_1$	Gamma(4,.5)
a_{2sv} $\forall s, v$	Gamma(10,3)	ς_1	Gamma(4,.5)
b_ρ	Gamma(2,1)	ν^O	Gamma(2,2)
		$INV_{0,sv}$	DT$_0$N$(M, (.75M)^2)$ $\forall s, v$

Table 4: Posterior Means and Asymptotic Standard Errors (in Parentheses) for Sales and Orders

Sales		**Inventory**	
ν^D	.5300 (.00076)	ν^O	3.8515 (.00309)
$\bar{\beta}_0$	.6156 (.00134)	$\bar{\kappa}_0$	2.7659 (.00951)
$\bar{\beta}_1$	-.3111 (.00329)	$\bar{\kappa}_1$	3.0549 (.03166)
$\bar{\beta}_2$	.2573 (.00028)	$\bar{\kappa}_2$	5.0405 (.00939)
$\bar{\beta}_3$	.5474 (.00159)	$\bar{\kappa}_3$	.9624 (.00094)
$\bar{\beta}_4$	-.2651 (.00132)		
$\bar{\beta}_5$	.0849 (.00102)		
$\bar{\beta}_6$	-1.2731 (.00502)		
σ_0	.1847 (.00034)	ς_0	2.5931 (.03322)
σ_1	.3877 (.00397)	ς_1	.4301 (.00443)
σ_2	.0355 (.00003)	ς_2	29.9161 (.16736)
σ_3	.6248 (.00173)	ς_3	.1348 (.00038)
σ_4	.1408 (.00055)		
σ_5	.1120 (.00037)		
σ_6	.3391 (.00389)		

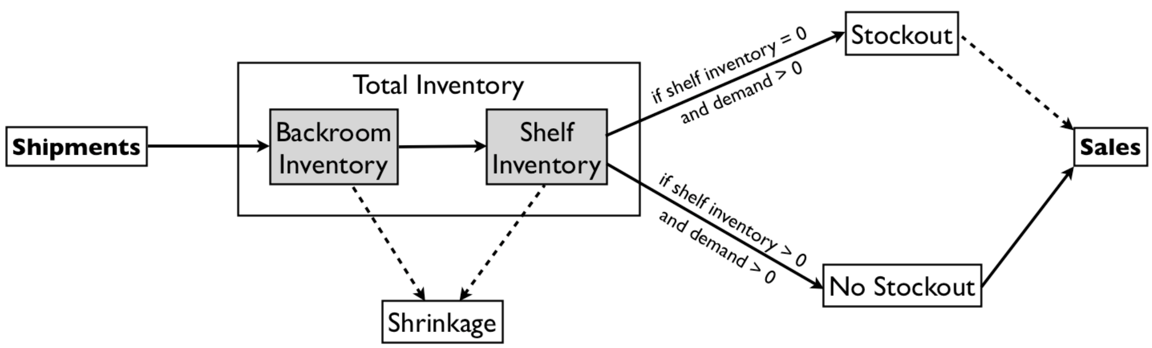

Figure 1: Overview of the Data Generating Process (Observed Data in Bold)

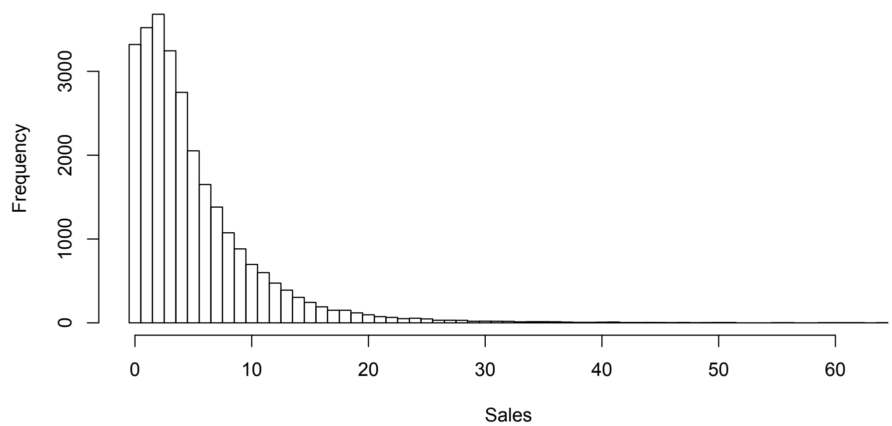

Figure 2: Histogram of Sales

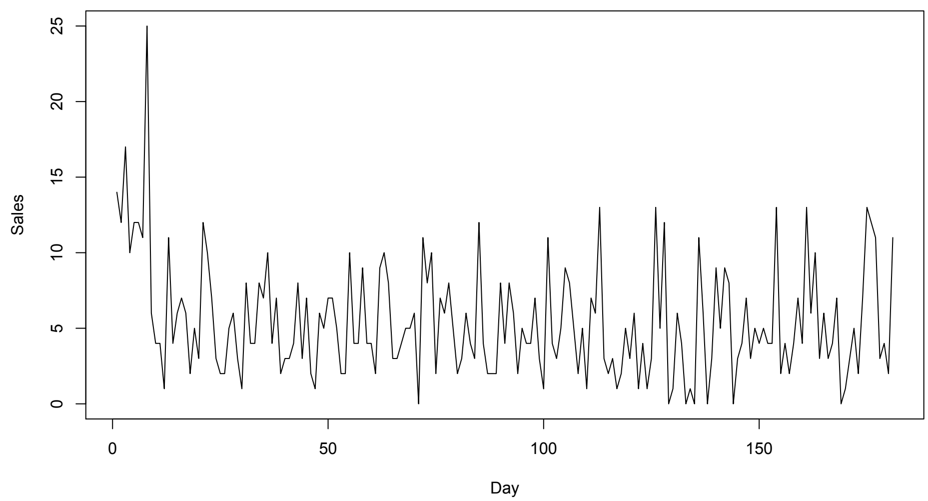

Figure 3: Sales Pattern for the Product with Promotion at the Beginning of the Time-Series

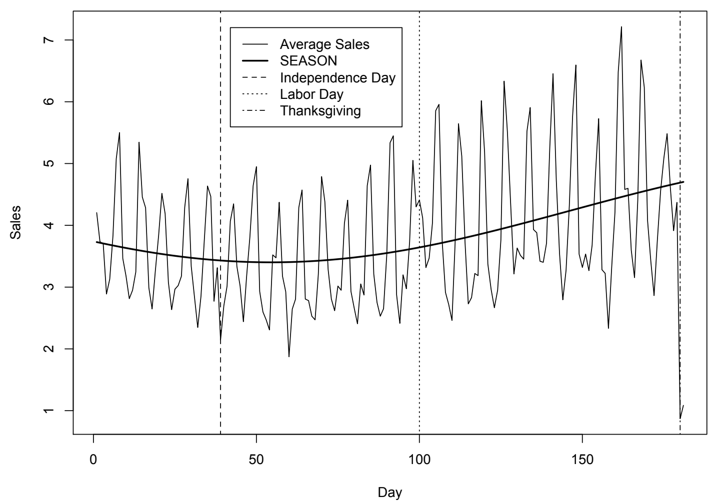

Figure 4: Trend of Average Sales Volume

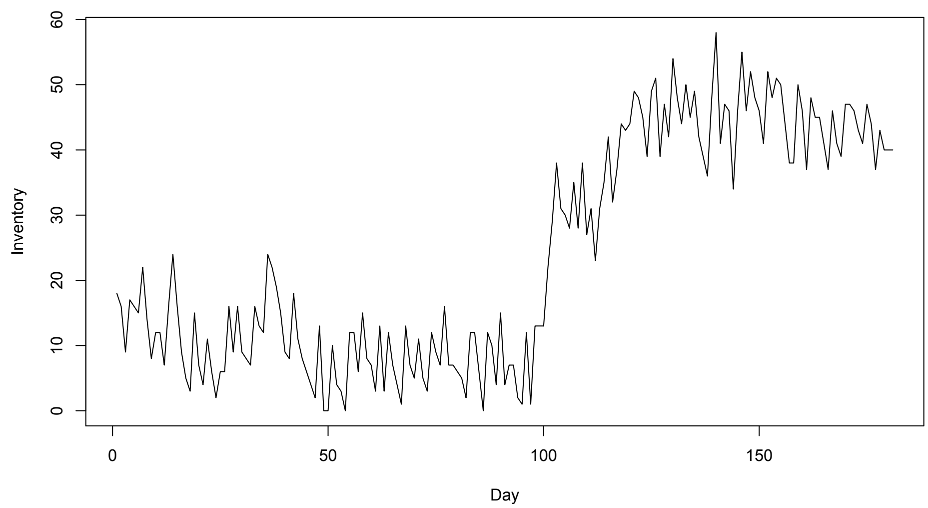

Figure 5: Observed Part of Inventory for one Product at one Store

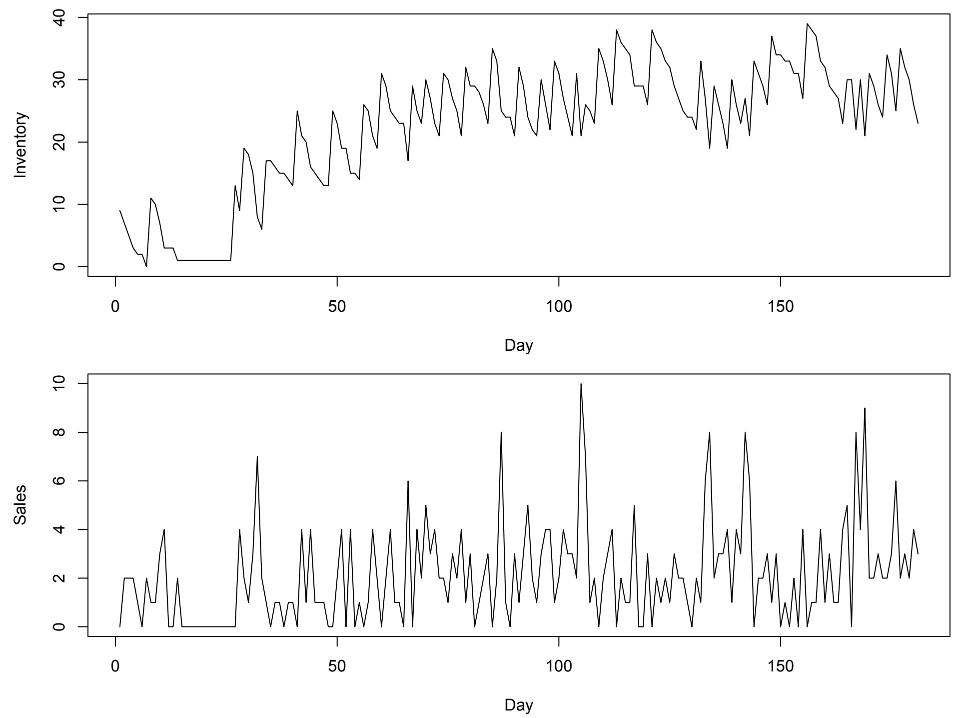

Figure 6: Observed Part of Inventory and Sales for one Product at one Store

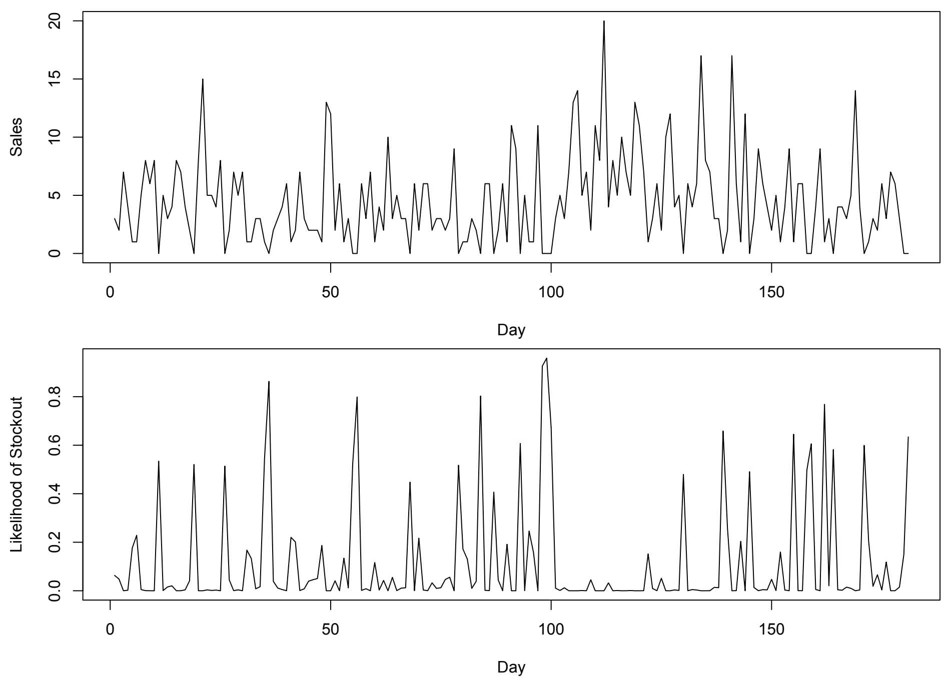

Figure 7: Sales Pattern and Estimated Likelihood of Stockouts for one Product at one Store

APPENDIX

Remaining Results

Table (4) presents the means and standard deviations for the hyper-parameters for the CMP distribution for sales without a stockout and for the CMP distribution for orders, as well as for ν^D and ν^O.

<< Insert Table (4) about here >>

Sales. First note that the CMP distribution for sales is highly overdispersed, as indicated by the mean of ν^D being .53. Thus, a model using the standard Poisson distribution for sales would have been strongly misspecified. This overdispersion also means that small changes in λ^D_{svt} lead to larger changes in the mean of the demand distribution, a fact to keep in mind when analyzing the remaining parameters.

The results for $\bar{\beta}_1$ to $\bar{\beta}_6$ are as expected. Lower prices lead to higher demand ($\bar{\beta}_1 = -.31$; values in parentheses refer to the mean); the relatively high variance across stores and products for the price parameter ($\sigma_1 = .39$) is due to the problem that for many store-product combinations prices change only very rarely, if ever. This translates to a price elasticity of -.4 to -.6 (depending on the amount of average sales, due to the non-linearity of the CMP distribution for demand). This elasticity is well within the range of commonly observed price elasticities, though - not surprisingly given the category and low prices - towards the more inelastic side (see Bijmolt et al. (2005) for a meta-analysis). Also, demand is 40-60% higher on weekends ($\bar{\beta}_2 = .26$) and increases as the season progresses towards the winter time ($\bar{\beta}_3 = .55$). This translates to an elasticity of demand with respect to season of about .7 to 1 (however, given the wide range of β_3 ($\sigma_3 = .62$) this may vary more for individual store-product combinations). It is reassuring, though, that the elasticities are close (on average) to 1, which would be the expected average value given how $SEASON$ was defined.

The holidays have different effects on demand: On Independence Day, demand is 30-40%

lower on average ($\bar{\beta}_4 = -.27$); Labor Day, in contrast, is probably regarded as an extra weekend day to do shopping by many people, leading to a weak (10-20%) positive effect on demand ($\bar{\beta}_5 = .08$; the lower bound of the 95% highest probability density interval is .008). Finally, Thanksgiving has a very strong (80-90%) negative effect since most likely the stores were closed for at least part of the day ($\bar{\beta}_6 = -1.27$).

The posterior means of the means for the multivariate normal distribution for the random shocks ($\vec{\omega}$) range from .00079 to .00153, with an overall mean of .00104. The variances (i.e., the diagonal of $S \otimes V$) have posterior means ranging from .00090 to .00137, with an overall mean of .00109. Thus, variances are very close to the means, leading to modes around one for $\exp(\gamma_{svt})$.

Finally, the posterior means for $\bar{a}_0$ and b_ρ are -4.46 and .19, respectively, resulting in a strongly U-shaped beta distribution for the probability of stockouts.

Inventory. In contrast to the distribution for sales, the CMP distribution for orders is strongly underdispersed relative to a Poisson distribution ($\nu^O = 3.85$), as expected from inspection of the data. As a result, the changes in λ^O_{svt} have to be fairly large even for small changes in the mean probability of orders. Thus, the size of the $\bar{\kappa}$ parameters and the $\bar{\beta}$ parameters cannot be directly compared. But again, as expected, the probability of orders increases with upcoming weekends ($\bar{\kappa}_3 = .96$), as well as with the season ($\bar{\kappa}_2$=5.04). However, the effect of the season is not very consistent across stores and products ($\varsigma_2 = 29.92$); the 95% credible interval of κ_{2sv} is strictly positive for only 82 of the 170 store-product combinations, and even strictly negative for seven of them. The posterior means for the hierarchy on κ_{1sv}, the effect of inventory relative to expected sales, are $\bar{\kappa}_1 = 3.05$ and $\varsigma_1 = .43$. Combining them to get the resulting gamma distribution at each iteration and then averaging over them, the posterior mean and variance of the distribution are 1.29 and .56, respectively.

The parameters for the CMP-binomial distributions for shrinkage given that an incident occurs are as follows: First, the probability of shrinkage for one case/unit is almost independent of any other case/unit, as evidenced by the posterior means for ν_u and ν_c being .9971 and 1.0044, respectively. The marginal probability of each available unit to be shrunk, $p_{u,s}$, ranges from 1.11% to 3.42% across stores (remember that the +1 in equation (14) implies that at least one unit is shrunk if an incident happens). Similarly, the marginal probability of each available case to be shrunk ranges from 4.32% to 5.31%. In comparing those rates, one should keep in mind that at any day there are less cases that can be shrunk than individual units, i.e. these values do not necessarily imply that shrinkage of two cases is more likely than shrinkage of two units, conditional on occurrence of the respective type of shrinkage.