A Method for Estimating Stock-Out Based Substitution Rates by Using Point-of-Sale Data

Selçuk Karabati¹, Barış Tan² and Ömer Cem Öztürk³

¹¹,²² Graduate School of Business, Koç University, Rumeli Feneri Yolu, Sariyer, 34450 Istanbul, Turkey
²³ Graduate School of Sciences and Engineering, Koç University, Rumeli Feneri Yolu, Sariyer, 34450 Istanbul, Turkey
¹¹skarabati@ku.edu.tr, ²²btan@ku.edu.tr, ³³oozturk@bus.emory.edu

Empirical studies in retailing suggest that stock-out rates are considerably large in many product categories. Stock-outs result in demand spill-over, or substitution, among items within a product category. Product assortment and inventory management decisions can be improved when the substitution rates are known. In this paper, we present a method to estimate product substitution rates by using only Point-of-Sale (POS) data. Our practical approach clusters POS intervals into states where each state corresponds to a specific substitution scenario. We then consolidate available POS data for each state, and estimate the substitution rates using the consolidated information. We provide an extensive computational analysis of the proposed substitution rate estimation method. The computational analysis and comparisons with an estimation method from the literature show that the proposed estimation method performs satisfactorily with limited information.

Key words: substitution rate estimation; stock-outs; point-of-sale data; retailing

1. Introduction

Understanding how customers respond when they cannot find the product they are looking for is of utter importance for both retailers and producers (Rajaram and Tang (2001), Gruen et al. (2002), Colacchio et al. (2003)). When a customer demanding a particular product cannot find it on the shelf, she may delay her purchase; decide not to buy the item; buy the item at another store, or substitute the item with another product (Campo et al. (2003)). The substitution rate between two products is defined as the probability that a customer arriving with the intention of purchasing a particular product purchases the other one when the product she intends to buy is not available. Recent studies reported that on average 40 percent of consumers purchase a substitute product when they are faced with an out-of-stock situation (Gruen et al. (2002)).

From the producers’ perspective, product substitution may result in a partial loss due to a substitution to a lower priced item (in the case of substitution to the same brand) or, more unfavorably, in a demand shift within the category to competitors’ items (in the case of substitution to a different
brand). Similarly, retailers can improve their profits by taking product substitution into account when making assortment and inventory decisions about a category. More precisely, they can use substitution rates to determine the optimal assortment of products, and the optimal inventory levels of products included in the assortment (see, e.g., van Ryzin and Mahajan (1999), Smith and Agrawal (2000), Netessine and Rudi (2003), Kök and Fisher (2007), Cachon et al. (2005)).

If a product is not carried by the retailer, or, it is carried by the retailer but it is temporarily out of stock, customers cannot find that item in the store. Although both of these situations may lead to product substitution, the first case is due to a strategic assortment decision, and the latter one is due to demand uncertainty and inventory decisions. In this study, we focus on stock-out based substitutions. Understanding customer response in this setting is also instrumental in the management of product assortments. Furthermore, despite the adoption of various initiatives, empirical studies show that the measured stock-out rates in retailing are quite high. Studies reported in Andersen Consulting (1996) and Gruen et al. (2002) show that on average 8.3 percent of the SKUs that a typical retailer carries are out-of-stock at a particular moment in time.

Although inventory and assortment planning under product substitution problems are analyzed extensively in the literature, the number of studies that focus on the estimation of substitution rates is limited. Substitution rates are used as parameters in inventory and assortment planning problems, and the objective of this study is to propose a practical method that can be used to estimate the substitution rates by using information that is readily available to retailers. Since inventory data is reported to be unavailable or highly inaccurate (DeHoratius and Raman (2004)), our proposed method does not rely on it and uses only POS data. Furthermore, no assumptions are made regarding the inventory control policy used by the retailer and the characteristics of the customer demand arrival process.

Our state-space-based method clusters POS intervals into states where each state corresponds to a specific substitution scenario depending on availability of the products under consideration.
We then consolidate available POS data for each state, and estimate the substitution rates using the consolidated information.

The stock-out situations and resulting substitution actions affect the observed sales of a product in such a way that sales information no longer reflects the core demand of the product. Therefore, the estimation of demand from sales information can be considered as a problem of estimation of censored and truncated distributions from a right-censored sample (see, e.g., Greene (1997), Allison (2002)). A number of studies examine estimators for such distributions (see, e.g., Mullahy (1986), Shaw (1988), Grogger and Carson (1991)). However, in order for these estimation methods to perform accurately, the truncated or censored parts should be small. Furthermore these methods do not give information about the substitution structure.

The closest study to the problem analyzed in this paper has been presented by Anupindi et al. (1998). Anupindi et al. (1998) present a maximum likelihood estimation method to estimate the arrival rates and stock-out based substitution rates in a setting with Poisson arrival and the simultaneous replenishment of products. Furthermore, this method requires that inventory data be available all the time. A comparison of our method with the method proposed in Anupindi et al. (1998) shows that our method performs satisfactorily even when the inventory related information is not available.

The main contribution of this paper is the development of a method that can be used under very general conditions to estimate product substitution rates that are required by inventory and assortment planning methods. Our method does not assume any arrival, substitution, or inventory control structure, and it only requires POS data.

The outline of the paper is as follows. In §2, we describe our model and estimation method. In §3, we explain the implementation of our method for two different information availability situations. In §4, we provide a computational evaluation of our method’s performance. In §5 we compare the
performance of our method with that of the maximum likelihood estimation method presented in Anupindi et al. (1998). Finally, we present our concluding remarks in §6.

2. Model

2.1. Model Description

We consider a retailer that stocks and sells $N$ products in a category. Demand rates for these $N$ products are $\lambda_1, \lambda_2, \ldots, \lambda_N$, respectively. We assume that the retailer has access to the POS data for the period $[0, T]$, i.e., information on the number of units sold of each product in each POS interval. Each POS interval is assumed to have a length of $\tau$ unit-time periods.

If customers cannot find their first-choice product on the shelf, the demand for that product will either spill over to another product according to a probabilistic substitution structure or will be lost. With probability $1 - \psi$ customers do not substitute. Given that a customer decides to substitute, the probability of substituting product $j$ for product $i$ is $\beta_{ij}$. Then, for an arriving customer, the probability of substituting product $j$ for product $i$ is $\alpha_{ij} = \psi \beta_{ij}$. If the second-choice product is not available either, then the demand is lost. In other words, we put one substitution attempt restriction, which is also imposed in Anupindi et al. (1998) and Smith and Agrawal (2000).

When the service levels are reasonably high, the impact of the secondary level substitutions may be negligible; therefore, a model with the one substitution restriction can still generate reliable estimates of the primary substitution rates. Kök et al. (2006) state that it is also possible to approximate a multiple substitution attempt model with a single-attempt model by adjusting the parameters.

The most common customer choice model that operates under the one substitution attempt restriction is the Market-Share Based Substitution Model (see Smith and Agrawal (2000), Netessine and Rudi (2003), and Kök and Fisher (2007)). In the market-share based model, the substitute product is chosen according to the substitution probability matrix $\alpha_{ij} = \psi \frac{\lambda_j}{\sum_{i \in N_j} \lambda_i}$, $i, j = 1, 2, \ldots, N$, and $i \neq j$. Smith and Agrawal (2000) define three additional choice models that operate
under the single substitution attempt assumption: Random: $\alpha_{ij} = \frac{6}{N-1}, i,j = 1,2,\ldots,N$, and $i \neq j$; Substitution to Adjacent Product: $\alpha_{1,2} = \alpha_{N,N-1} = \psi$; $\alpha_{i,i-1} = \alpha_{i,i+1} = \frac{\psi}{2}, i = 2,\ldots,N-1$; and Substitution to a Single Item: $\alpha_{ik} = \psi, k = \left\lceil \frac{N}{2} \right\rceil, i = 1,2,\ldots;k-1,k+1,\ldots,N$. Our analysis in §5 shows that the impact of the underlying customer choice model on the estimation quality of the proposed methodology is not significant.

2.2. State-Space-Based Estimation Method (SSBE)

We model the customer arrival process as a discrete time stochastic process. Let $A_i(n)$ be the indicator variable that is equal to 1 if a customer who demands one unit of product $i$ arrives in period $n$ and 0 otherwise. We take the length of each period very short to ensure that the probability of having two arrivals in the same time period is very small, and can be assumed to be zero. The arrival rate of customers in a period demanding product $i$ can be expressed as $\lambda_i = \lim_{T \to \infty} \frac{1}{T} \sum_{n=1}^{T} A_i(n)$.

We can observe, from the POS data, whether product $i$ is sold in period $n$. Let $S_i(n)$ be the sales indicator variable (1 if sales occurs and 0 otherwise) after customer arrival. However, we cannot observe the inventory status which is tracked with indicator variable $I_i(n)$. $I_i(n)$ is 1 if product $i$ is available at the beginning of period $n$ and it is 0 otherwise. The inventory status of all products at the start of period $n$ is $I(n) = (I_1(n),\ldots,I_N(n)), I(n) \in I_S$. Note that $I_S$ has $2^N$ members.

Product $i$, $i = 1,2,\ldots,N$, can be sold only if it is available and demanded by a customer. The demand for product $i$ can be generated by a customer who arrives with the intention to purchase product $i$, and also by a customer who arrives with the intention to purchase product $j$ and switches to product $i$ when product $j$, $j = 1,2,\ldots,N$, she is looking for is not available in the inventory. Then the probability that product $i$ is sold in period $n$ can be written as:

$$P[S_i(n) = 1] = P[A_i(n) = 1, I_i(n) = 1] + \sum_{j \neq i} \alpha_{ij} P[A_j(n) = 1, I_j(n) = 0, I_i(n) = 1]. \quad (2.1)$$
If we condition Equation 2.1 on a particular realization of the inventory indicator state vector \( I(n) = I_0 \), we obtain:

\[
P[S_i(n) = 1 | I(n) = I_0] = P[A_i(n) = 1, I_i(n) = 1 | I(n) = I_0] \\
+ \sum_{j \neq i} \alpha_{ji} P[A_j(n) = 1, I_j(n) = 0, I_i(n) = 1 | I(n) = I_0].
\] (2.2)

Since the arrival process is independent of the inventory status, Equation 2.2 becomes:

\[
P[S_i(n) = 1 | I(n) = I_0] = P[A_i(n) = 1] P[I_i(n) = 1 | I(n) = I_0] \\
+ \sum_{j \neq i} \alpha_{ji} P[A_j(n) = 1] P[I_j(n) = 0, I_i(n) = 1 | I(n) = I_0].
\] (2.3)

Note that, for a given inventory status realization \( I_0 \), \( P[I_i(n) = 1 | I(n) = I_0] \) and \( P[I_j(n) = 0, I_i(n) = 1 | I(n) = I_0] \) are either 1 or 0. Let the indicator variable \( \delta_{i,I_0} \) is defined to be equal to 1 when \( I_i(n) = 1 \) in \( I_0 \) and 0 otherwise. Then, Equation 2.3 can be written as:

\[
P[S_i(n) = 1 | I(n) = I_0] = P[A_i(n) = 1] \delta_{i,I_0} + \sum_{j \neq i} \alpha_{ji} P[A_j(n) = 1] (1 - \delta_{j,I_0}) \delta_{i,I_0}.
\] (2.4)

The above equation in the long run yields

\[
\lim_{T \to \infty} \frac{1}{T} \sum_{n=1}^{T} P[S_i(n) = 1 | I(n) = I_0] = \delta_{i,I_0} \lim_{T \to \infty} \frac{1}{T} \sum_{n=1}^{T} P[A_i(n) = 1] \\
+ \delta_{i,I_0} \sum_{j \neq i} \alpha_{ji} \lim_{T \to \infty} \frac{1}{T} \sum_{n=1}^{T} P[A_j(n) = 1] (1 - \delta_{j,I_0}).
\] (2.5)

Therefore, from the definition of \( \lambda_i, i = 1, 2, \ldots, N \),

\[
\lim_{T \to \infty} \frac{1}{T} \sum_{n=1}^{T} P[S_i(n) = 1 | I(n) = I_0] = \delta_{i,I_0} [\lambda_i + \sum_{j \neq i} \alpha_{ji} \lambda_j (1 - \delta_{j,I_0})].
\] (2.6)

Since, \( E[S_i(n)] = P[S_i(n) = 1] \), the left-hand side of the above equation is the expected sales of product \( i \) per unit time when the inventory indicator state is \( I_0 \), denoted by \( s_{i,I_0} \):

\[
s_{i,I_0} = \lim_{T \to \infty} \frac{1}{T} \sum_{n=1}^{T} E[S_i(n)|I(n) = I_0] = \delta_{i,I_0} [\lambda_i + \sum_{j \neq i} \alpha_{ji} \lambda_j (1 - \delta_{j,I_0})].
\] (2.7)

In Equation 2.7, it is clear that if \( \delta_{i,I_0} = 0 \) then \( s_{i,I_0} = 0 \). On the other hand, if we choose an inventory status realization \( I'_0 \) such that \( \delta_{i,I'_0} = 1 \), Equation 2.7 yields:

\[
s_{i,I'_0} = \lambda_i + \sum_{j \neq i} \alpha_{ji} \lambda_j (1 - \delta_{j,I'_0}).
\] (2.8)
Equation 2.8 is the main equality that relates the substitution probabilities to the sales rates. In our setting, the available data is the consolidated sales figures in each POS interval that includes a fixed number of periods. Although the inventory status may change from 1 to 0, if the inventory is depleted, or from 0 to 1, if a replenishment is made, during a POS interval, for those POS periods where the inventory status does not change, an approximation for Equation 2.8 is obtained by calculating the expected sales rate from the average of sales observed in POS intervals whose inventory indicator state equal to \( I_0' \):

\[
\bar{s}_i I_0' = \lambda_i + \sum_{j \neq i} \alpha_{ji} \lambda_j (1 - \delta_{ji} I_0') + \epsilon_i I_0', \tag{2.9}
\]

where \( \bar{s}_i I_0' \) is the average sales of product \( i \) per unit time of POS intervals with the inventory indicator state equal to \( I_0' \), and \( \epsilon_i I_0' \) is the approximation error.

Our method is based on estimating the unknown parameters by minimizing the sum of errors in a set of equations obtained by using Equation 2.9 for different values of \( I_0' \), where \( I_0' \) is specified by using a labeling procedure that is based on grouping POS intervals according to the observed sales numbers. More specifically, let \( \hat{\lambda}_i \) and \( \hat{\alpha}_{ji} \) be the estimated demand rate of product \( i \) and the substitution probability from product \( j \) to product \( i \), respectively. Then we can write

\[
\bar{s}_i I_0' = \hat{\lambda}_i + \sum_{j \neq i} \hat{\alpha}_{ji} \hat{\lambda}_j (1 - \delta_{ji} I_0') + \epsilon_i I_0', \tag{2.10}
\]

where \( \epsilon_i I_0' \) is the estimation error. Then, we determine \( \hat{\lambda}_i \) and \( \hat{\alpha}_{ji} \) by solving the following optimization problem:

\[
\min \sum_{i=1}^{N} \sum_{I_0' \in I_S} w_{I_0'} \epsilon_i I_0'^2
\]

subject to

\[
\hat{\lambda}_i + \sum_{j \neq i} \hat{\alpha}_{ji} \hat{\lambda}_j (1 - \delta_{ji} I_0') + \epsilon_i I_0' = \bar{s}_i I_0', \quad I_0' \in I_S,
\]

\[
\sum_{j} \hat{\alpha}_{ji} \leq 1, \quad i = 1, 2, \ldots, N,
\]
\[ \lambda_i \geq 0, \quad i = 1, 2, \ldots, N, \]
\[ \alpha_{ij} \geq 0, \quad j, i = 1, 2, \ldots, N, \]

where \( w_{I_0} \) is the weight of observed state \( I_0' \). We note that, when \( m(I_0') \) is defined as the number of observed POS intervals in state \( I_0' \), \( w_{I_0} \) can be computed as \( m(I_0') / \sum_{I \in I_S} m(I) \). For a given product \( l, l = 1, 2, \ldots, N \), there are \( 2^{N-1} \) possible values of \( I_0' \) with \( \delta_{l, I_0} = 1 \). Therefore, we can write \( (2^{N-1})N \) equations to estimate \( N^2 \) unknowns, i.e., \( N \) demand rates and \( N^2 - N \) substitution probabilities.

Since the number of unknowns is much smaller than the number of equations, we do not need to generate all of the equations. When we choose the intervals in which all of the products are available, i.e., \( I_0' = (1, \ldots, 1) \), from Equation 2.10, we obtain \( \pi_{0,(1,1,\ldots,1)} = \lambda_i + e_{i,(1,\ldots,1)} \). Similarly, by using the intervals in which only one product is not available and all the others are available, we generate \( N(N - 1) \) additional equations to estimate all the substitution probabilities and the demand rates. Therefore, our method can be used to estimate the substitution probabilities and the demand rates with \( N + N(N - 1) \) or \( N^2 \) equations. This could be of practical value especially when the number of products in a category is large. In addition, we can use other states such as the ones in which two products are unavailable and all the others are available to improve the estimations. We call this the State-Space-Based Estimation (SSBE) method. Since the inventory status is not observable, in the next section we explain in detail how we estimate the average sales rates from the POS data for a given \( I_0' \).

3. POS Based Estimation Methodology

In this section, we provide a detailed description of the implementation of SSBE. For ease of explanation, we first discuss the case with complete information. Then, we explain SSBE with POS data only.

3.1. SSBE with Complete Information

SSBE first defines a set of sub-states and the related state structure, calculates appropriate statistics, and then finds the estimates of the parameters. In Section 3.3, as an additional step of the
estimation procedure, we will discuss the appropriate metrics to evaluate the accuracy of the estimates.

For exposition simplicity, and without any loss of generality, we assume that the retailer replenishes products periodically (e.g., at the beginning of every week) and that the lead time is zero. We also assume that the category has three items, labeled A, B, and C. In this case, we estimate three demand parameters ($\lambda_A, \lambda_B, \lambda_C$) along with six substitution parameters ($\alpha_{AB}, \alpha_{AC}, \alpha_{BA}, \alpha_{BC}, \alpha_{CA}, \alpha_{CB}$). Furthermore, the initial stocking levels of products are assumed to be available as an input to the estimation procedure.

The POS data, initial stocking levels, and the length of the replenishment cycle are the inputs of the estimation procedure with complete information. We first determine the POS intervals in which a replenishment occurs by using the periodicity of replenishment and the length of a POS interval. Suppose that the retailer operates 10 hours/day, 7 days/week, gives its orders once a week, and the length of a POS interval is equal to one hour. Assuming that the retailer gives an order at the beginning of the first interval (and receives it instantly as lead time is zero), the first replenishment interval is the first POS interval. The second replenishment occurs 70 hours later, which is just the beginning of 71st POS interval, and so on.

After specifying the replenishment intervals, our method determines the last sale points of products in each replenishment cycle. This is equivalent to determining, for each product, the POS interval in which the inventory is depleted.

Next, we label the POS data intervals according to each product’s inventory status by using the sub-state definitions presented in Table 1. In order to label the sub-states, we first determine the amount of the cumulative sales of a product in a replenishment cycle. When this is equal to the initial stock quantity of the product, then the last sale point for this product is the transition interval. The POS intervals before that transition interval until the previous replenishment interval (i.e., beginning of the cycle) are in-stock intervals, and the intervals after the transition interval
until the beginning of next cycle are out-of-stock intervals. On the other hand, if the amount of
the cumulative sales of a product in a cycle is lower than its initial stock quantity all intervals in
that cycle are in-stock intervals.

<table>
<thead>
<tr>
<th>Sub-State</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>Product is on the shelf during the POS interval.</td>
</tr>
<tr>
<td>O</td>
<td>Product has been depleted completely within the POS interval.</td>
</tr>
<tr>
<td>×</td>
<td>Product is unavailable from the beginning of the POS interval.</td>
</tr>
</tbody>
</table>

In SSBE, every POS interval is described by a state, which is a combination of the sub-states
of products in that POS interval. In a 3-product problem, we use the $2^3 - 1 = 7$ states
given in Table 2 to estimate the demand and substitution probabilities.

<table>
<thead>
<tr>
<th>State</th>
<th>Sub-state of product A</th>
<th>Sub-state of product B</th>
<th>Sub-state of product C</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1,1)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(1,1,0)</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>(1,0,1)</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>(1,0,0)</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>(0,1,1)</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(0,1,0)</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>(0,0,1)</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note that states with transition sub-states (i.e., sub-states representing the intervals in which
the beginning inventory of the POS interval is positive but depleted completely) are not taken into
consideration, because substitution effects are not fully observed in these sub-states. In addition,
the state where all sub-states are out-of-stock is not included in the table, because in that case
POS intervals will not provide any information about demand and substitution probabilities (no
sales and substitution can take place in that state). Figure 1 presents an example of the labeling
procedure for weekly POS data with a POS interval length of one hour.

In the last step, we estimate demand rates and substitution probabilities. For the case with three
products, $2^{3-1} \times 3 = 12$ equations can be derived from 2.10. In line with the formulation we have
A Method for Estimating Stock-Out Based Substitution Rates by Using Point-of-Sale Data

Figure 1 Sample labeling for SSBE with complete information.

presented in Section 2.2, these equations and the estimation error minimization problem used to determine \( \hat{\lambda}_s \) and \( \hat{\alpha}_{ij} \) can be explicitly stated as follows:

\[
\min \sum_{i \in \{A,B,C\}} \sum_{I_0 \in I_S} w_i e_i^2 I_0^2
\]

subject to

\[
\varphi_{A,(1,1,1)} = \lambda_A + e_{A,(1,1,1)}
\]

\[
\varphi_{B,(1,1,1)} = \lambda_B + e_{B,(1,1,1)}
\]

\[
\varphi_{C,(1,1,1)} = \lambda_C + e_{C,(1,1,1)}
\]

\[
\varphi_{A,(1,1,0)} = \hat{\alpha}_{CA} \hat{\lambda}_C + \hat{\lambda}_A + e_{A,(1,1,0)}
\]

\[
\varphi_{A,(1,0,1)} = \hat{\alpha}_{BA} \hat{\lambda}_B + \hat{\lambda}_A + e_{A,(1,0,1)}
\]

\[
\varphi_{A,(1,0,0)} = \hat{\alpha}_{BA} \hat{\lambda}_B + \hat{\alpha}_{CA} \hat{\lambda}_C + \hat{\lambda}_A + e_{A,(1,0,0)}
\]

\[
\varphi_{B,(1,1,1)} = \hat{\alpha}_{CB} \hat{\lambda}_C + \hat{\lambda}_B + e_{B,(1,1,1)}
\]

\[
\varphi_{B,(0,1,1)} = \hat{\alpha}_{AB} \hat{\lambda}_A + \hat{\lambda}_B + e_{B,(0,1,1)}
\]

\[
\varphi_{B,(0,1,0)} = \hat{\alpha}_{AB} \hat{\lambda}_A + \hat{\alpha}_{CB} \hat{\lambda}_C + \hat{\lambda}_B + e_{B,(0,1,0)}
\]

\[
\varphi_{C,(1,1,1)} = \hat{\lambda}_B + \hat{\alpha}_{CB} \hat{\lambda}_C + e_{C,(1,1,1)}
\]

\[
\varphi_{C,(0,1,1)} = \hat{\alpha}_{AC} \hat{\lambda}_A + \hat{\lambda}_C + e_{C,(0,1,1)}
\]

\[
\varphi_{C,(0,0,1)} = \hat{\alpha}_{AC} \hat{\lambda}_A + \hat{\alpha}_{BC} \hat{\lambda}_B + \hat{\lambda}_C + e_{C,(0,0,1)}
\]

\[
\sum_{j \in \{A,B,C\}, j \neq i} \hat{\alpha}_{ij} \leq 1, \quad i \in \{A,B,C\},
\]

\[
\hat{\lambda}_i \geq 0, \quad i \in \{A,B,C\},
\]

\[
\hat{\alpha}_{ij} \geq 0, \quad i,j \in \{A,B,C\}.
\]
In this non-linear optimization problem with a weighted-quadratic objective function, $s_{i,I_0}$ is a parameter of the problem, and equal to the average sales of product $i$ in the POS intervals with the inventory indicator state equal to $I_0$. The weights of the error terms in the objective function are taken proportional to observed number of POS interval states that is associated with the equality where the error term is used (see Section 2.2).

In the equalities where the POS state vector is equal to $(1,1,1)$, the error terms are equal to the difference between the estimated demand rates and the average effective demands of products in POS intervals in state $(1,1,1)$. When the POS state vector is equal to $(1,1,0)$, only product $C$ is out-of-stock, the average effective demand of product $A$ ($B$) is made up of its average real demand for a POS interval and $\alpha_{CA}$ ($\alpha_{CB}$) times the average real demand of product $C$. In a similar manner, in the constraints where the POS state vector is equal to $(1,0,1)$ and $(0,1,1)$ substitutions from products $B$ and $A$, respectively, are considered.

In the equality where the POS state vector is equal to $(1,0,0)$, both product $B$ and product $C$ are out-of-stock, the average effective demand of product $A$ will include its average real demand, $\alpha_{BA}$ times average real demand of product $B$, and $\alpha_{CA}$ times average real demand of product $C$. In a similar manner, in the equalities where the POS state vector is equal to $(0,1,0)$ and $(0,0,1)$ substitutions to products $B$ and $C$, respectively, are considered.

3.2. SSBE with POS Data Only

So far we assumed that we had information on the initial stocking levels of products and the replenishment intervals in addition to the POS data. In this section, we present the SSBE method for the case where neither initial stocking levels nor replenishment intervals are known. Therefore, we estimate the states corresponding to each POS interval, and then use these estimated states to estimate demand rates and substitution probabilities.
In this setting we cannot determine the POS intervals in which replenishments take place with certainty. Moreover, since inventory levels are not known, we cannot know exactly which POS intervals correspond to out-of-stock intervals.

Inspired by the standard process control procedures, we propose to estimate the out-of-stock intervals by first calculating the probability of observing a zero sale due to no arrival in a POS interval, and then comparing the probability of observing a group of \( l \) consecutive zero-sales POS intervals with a threshold value which is specified as \( \epsilon \). More precisely, if the probability of observing a group of \( l \) consecutive zero-sales POS intervals is less than a threshold value, then a zero sales group of length \( l \) is labeled as a stock-out group, and the POS interval just before this group is labeled as a transition interval. On the other hand, if the probability of observing a group of \( l \) consecutive zero-sales POS intervals is greater than the threshold, then these zero-sales POS intervals are labeled as in-stock intervals. In order to determine the probability of a zero sale, we first specify the functional form of the distribution of demands, and then estimate their parameters of the distribution through the POS intervals with positive sales.

With a higher value of \( \epsilon \), the probability of labeling an in-stock state as an out-of-stock state is smaller. It is also true that with a higher \( \epsilon \) value fewer states are labeled as out-of-stock, decreasing the number of POS intervals that will be used in estimating the substitution probabilities. Therefore, the estimation quality would be sensitive to the value of \( \epsilon \). In our numerical analysis, however, we observed that the estimation quality is more or less the same in the \( 10^{-5} \leq \epsilon \leq 10^{-3} \) range, and set the value of \( \epsilon \) as \( 10^{-4} \).

To illustrate the labeling procedure of SSBE with POS data only, suppose that the distribution of demand of the product under consideration is specified as Poisson. The methods developed for estimation from truncated data can be used to estimate the rate of the distribution using POS intervals with positive sales. Let \( \hat{\lambda} \) be the estimated rate. Then, the probability of observing zero sale in a POS interval due to no arrival is \( e^{-\hat{\lambda}} \). In that case, if the number of intervals in a group of
consecutive zero-sales POS intervals is less than or equal to the critical number $n$, where $(e^{-\lambda})^n > \epsilon$ and $(e^{-\lambda})^{n+1} \leq \epsilon$, or $-\frac{\ln r}{\lambda} - 1 < n < -\frac{\ln r}{\lambda}$, then they are labeled as in-stock intervals. We also note that, if the demand distribution is assumed to have a different functional form, the assumed distribution’s functional form is used to determine the critical number $n$ in a similar manner.

After estimating the out-of-stock intervals, our method labels the sub-states accordingly. Subsequent steps of SSBE, including the labeling of the states and the estimation of demand rates and substitution probabilities, are the same as those applied in the case with complete information.

### 3.3. Accuracy of the Estimates

The estimation procedure we have presented in Sections 3.1 and 3.2 works with POS data which is generated by customer arrivals, the inventory management system, and the underlying customer choice model or substitution structure. The estimation accuracy of our procedure greatly depends on whether we observe statistically significant number of POS intervals in which substitution takes place. The inventory management system has a central role in this process: consider for example a deterministic setting with two products and substitution probabilities that are equal to 1. If the two products are replenished concurrently, and their order quantities are equal to their individual demands per unit time times a constant, their inventories will be depleted simultaneously. Although the underlying substitution structure has substitution probabilities that are equal to 1, because no substitution could be observed through the POS data, we estimate substitution probabilities as zero. We also note that any procedure that relies only on the POS and/or inventory data to estimate the substitution probabilities faces this limitation.

In Section 4.2, we will study a number of examples to illustrate the relationship between the observed substitution rates and the accuracy of the estimates generated by SSBE with POS data only. The rate at which product $i$ is substituted with product $j$, $\zeta_{ij}$, can be estimated as follows:
Let $I_{ij}$ be the set of inventory status vectors with $I_i = 0$ and $I_j = 1$. Then, $\zeta_{ij}$ is given as

$$
\zeta_{ij} = \hat{\lambda}_i \hat{\alpha}_{ij} \frac{\sum_{I \in I_{ij}} m(I)}{T/\tau}.
$$

We note that, in the above expression, $\sum_{I \in I_{ij}} m(I)$ is the total number of POS intervals where substitution from product $i$ to product $j$ is possible, and $T/\tau$ is the total number of POS intervals. Therefore, the ratio $\frac{\sum_{I \in I_{ij}} m(I)}{T/\tau}$ gives the fraction of time where substitution from product $i$ to product $j$ takes place.

4. Performance of the Estimation Method with Limited Information

In this section we present a computational analysis of SSBE with POS data only. After describing the general setting in which the test problems are created in Section 4.1, we will study the impact of observed substitutions on the accuracy of the estimates in Section 4.2. The performance of the estimation methodology with different number of products is reported in Section 4.3.

4.1. Problem Setting

We evaluate the performance of the estimation method by generating problems for a given customer arrival process, customer choice model, and replenishment policy for a category consisting of up to 10 products. The test problems analyzed in Sections 4 and 5 have been created through a simulation in the MATLAB 7.0 environment, and the corresponding optimization problems have been solved with GAMS IDE 2.0.30.1 using CONOPT as the nonlinear programming solver.

In the test problems, the arrival process of product $i$ is a Poisson random variable with rate $\lambda_i$. If the customer cannot find her favorite product on the shelf the demand for that product may spill over to another product in the category in accordance with the underlying customer choice model. It is also assumed that a fixed-review period order-up-to level $(R, S)$ type inventory control system is employed. According to this system, every $R$ units of time (i.e., at every review point) an order is given to raise inventory position to a predetermined level $S$. We determine the order-up-to levels by using the decision rule for a specified fraction of demand satisfied directly from the
shelf, in the case of a periodic review system with lost sales and normal demand (see Silver et al. (1998) for details). Since demands of products are assumed to be Poisson, we employ the normal approximation for Poisson in computing the service levels.

4.2. Accuracy of the Estimates

In the test problems studied in this section, we consider a setting with general replenishment lead times and policies. In a real retail setting, items can be replenished several times in a week, and replenishment lead times may differ among the items in a category.

We assume that the category has five products, labeled $A$ through $E$. Hourly demand rates for the products $A$, $B$, $C$, $D$, and $E$ are 18, 15, 6, 3, and 3 respectively. The POS interval length is 0.25 hours or 15 minutes. We assume that the retailer operates 10 hours/day and 7 days/weeks. The customer choice process is assumed to be market-share based, resulting in the following substitution matrix:

\[
\begin{array}{cccc}
A & B & C & D \\
A & 0.222 & 0.089 & 0.044 & 0.044 \\
B & 0.240 & 0.080 & 0.040 & 0.040 \\
C & 0.185 & 0.154 & 0.031 & 0.031 \\
D & 0.171 & 0.143 & 0.057 & 0.029 \\
E & 0.171 & 0.143 & 0.057 & 0.029 \\
\end{array}
\]

In estimating the substitution probabilities, we employ SSBE with limited information for 5 products where a threshold value of $10^{-4}$ is used to estimate the out-of-stock POS intervals. Along with the substitution probabilities, for each product, we also compute the rate at which the product is substituted with another product, i.e., $\zeta_{ij}$. For example, if the number of states, or POS intervals, where product $B$ is available and product $A$ is not available is equal to 100 in a 4-week simulation (with a total of 1120 POS intervals), we can then compute the estimated substitution rate from product $B$ to product $A$ as product $B$'s estimated demand times the fraction of POS intervals with substitution from product $B$ to $A$ (i.e., $\frac{100}{1120}$) times the estimated substitution probability from product $B$ to $A$. Through the cases we will present in subsequent cases, we will argue that the estimation quality is very much dependent on the rate at which the product is substituted. When
the substitution rate is relatively higher, we observe more POS intervals with substitution, and estimations are based on statistically more significant observations.

**Case 1: Product B’s Service Level is Lower:**

In the first case we analyze, the total substitution probability, \( \psi \), is equal to 60%, service levels of Products A, C, D, and E are equal to 90%, and the service level of product B is equal to 80%. The replenishment lead time is one day for all products. Orders for products C, D, and E are released once a week (on Wednesday mornings). On the other hand, the retailer orders product A and product B twice per week (on Wednesday and Friday mornings). The system is simulated for 3 months with 100 replications.

In Figure 2a we report the quality of the substitution probability estimates for the ten largest estimated substitutions per hour. We measure the estimation quality as the ratio of the estimated and actual substitution probabilities. For example, the substitution rate from product B to product A is estimated, on average in 100 problems, as 0.331 substitutions per hour (secondary y axis, on the right-hand-side of the graph), the 95% Confidence Interval for the ratio of estimated and actual substitution probabilities is computed as [0.977, 1.004] (primary y axis, on the left-hand-side of the graph).

We note that, since the service level of product B is lower, we observe more substitutions from product B to other products, and corresponding substitution probabilities are estimated reasonably well. The remaining substitution probabilities cannot be estimated with high accuracy, because in a week we observe at most 0.05 substitutions/hour \( \times 70 \) hours/week = 3.5 substitutions between product pairs. For example, only 0.28% of product A’s demand is substituted with product C. It is clear that no statically significant data can be generated when substitution rates are low relative to the demand rates of products. From an inventory management point of view, it can be argued that the estimation of substitution probabilities that have very little impact on the effective demand of a product is not of primary importance.
Case 2: Non-Identical Review Times:

In the second case we consider the problem we presented in Case 1, however change the inventory review times as follows: Product $A$: Wednesday and Friday mornings; Product $B$: Thursday and Saturday mornings; Product $C$: Wednesday morning; Product $D$: Thursday morning; Product $C$: Friday morning. We also change the service level as 95% for all products. The system is simulated for 3 months with 100 replications. Since products are not depleted simultaneously now, in Figure 2b we observe a higher number of substitution instances, and the estimation accuracy is reasonably good when the estimated substitution rate is above 0.1 units per hour.

Case 3: Products $C$, $D$, and $E$’s Service Levels are Lower:

In the third case we again consider the problem we presented in Case 1, however change the service levels as follows: Products $A$ and $B$: 95%; Products $C$, $D$ and $E$: 75%. The system is first simulated for 3 months with 100 replications. We now expect to observe more substitutions from products $C$, $D$ and $E$ to products $A$ and $B$, and as presented in Figure 3a, all substitution probabilities from low service level items to high service level items are accurately estimated. In Figure 3b, we report the results when the system is simulated for 6 months. We note that the estimated substitution rates do not change much, indicating the fact that even when POS data are collected for longer time periods, the accuracy problem for substitution probabilities from high service level items to low service level items will not be eliminated.

4.3. Impact of Number of Products

We note that the discussion we presented in the previous section has shown when substitution probabilities can be estimated accurately depending on the estimated rate of substitution. Our objective now is to analyze the impact of certain factors on the overall performance of the estimation procedure, and to analyze the performance in this and the subsequent sections, we employ three common estimation error measures: mean absolute percentage error ($MAPE$), mean absolute deviation ($MAD$), and maximum absolute deviation ($MAX$). The corresponding error metrics for
substitution probabilities are defined as follows:

$$MAPE = \frac{\sum \sum |\hat{\alpha}_{ij} - \alpha_{ij}|}{N^2 - N} \times 100$$,
$$MAD = \frac{\sum \sum |\hat{\alpha}_{ij} - \alpha_{ij}|}{N^2 - N}$$,
$$MAX = \max_{i,j} |\hat{\alpha}_{ij} - \alpha_{ij}|$$.

We note that these performance measures can be significantly affected by a large error in the estimation of a substitution probability. For example, estimating a non-zero substitution probability
as 0 due to not observing substitutions will yield an error of 100%. For a 3-product case, this single estimation error would increase $MAPE$ by 16.6%.

In Tables 3 and 4, we report the performance of SSBE in 4-, 6-, 8-, and 10-product problems. We assume that the arrival processes are Poisson random variables with equal rates, the total arrival rate is 48 units per hour, and the substitution matrix is market-share based. We generate 100 problems for each number of products for a simulation length of one year and POS interval length of 0.25 hours. We report the average error values, and the bounds of the 95% confidence intervals on the error values for the complete information (SSBE-CI) and limited information (SSBE-POS) cases. In Table 4, where the errors in the estimation of demand rates are reported, the error metrics are defined as follows:

$$MAPE_\lambda = \frac{\sum |\hat{\lambda}_i - \lambda_i| 100}{N}, \quad MAD_\lambda = \frac{\sum |\hat{\lambda}_i - \lambda_i|}{N}, \quad MAX_\lambda = \max_i |\hat{\lambda}_i - \lambda_i|.$$  

When we examine the figures reported in Tables 3 and 4, we observe that the performance of SSBE does not change significantly even when the number of products is quite large.

### Table 3  MAPE, MAD, and MAX with respect to real substitution probabilities.

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>No of Products</th>
<th>SSBE-CI</th>
<th>SSBE-POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>4</td>
<td>39.343 [37.594, 41.093]</td>
<td>38.554 [36.882, 40.225]</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>65.47 [63.920, 67.021]</td>
<td>64.008 [62.459, 65.556]</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>83.293 [81.985, 84.601]</td>
<td>81.339 [80.087, 82.590]</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>96.071 [94.864, 97.278]</td>
<td>95.217 [94.032, 96.403]</td>
</tr>
<tr>
<td>MAD</td>
<td>4</td>
<td>0.052 [0.050, 0.055]</td>
<td>0.051 [0.049, 0.054]</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.052 [0.051, 0.054]</td>
<td>0.051 [0.050, 0.052]</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.048 [0.047, 0.048]</td>
<td>0.046 [0.046, 0.047]</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.043 [0.042, 0.043]</td>
<td>0.042 [0.042, 0.043]</td>
</tr>
<tr>
<td>MAX</td>
<td>4</td>
<td>0.121 [0.116, 0.126]</td>
<td>0.12 [0.115, 0.125]</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.132 [0.125, 0.140]</td>
<td>0.127 [0.120, 0.135]</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.151 [0.144, 0.159]</td>
<td>0.146 [0.139, 0.153]</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.178 [0.169, 0.186]</td>
<td>0.177 [0.169, 0.184]</td>
</tr>
</tbody>
</table>
Table 4  \( MAPE_{\lambda}, MAD_{\lambda}, \) and \( MAX_{\lambda} \) with respect to real demand rates.

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>No of Products</th>
<th>SSBE-CI [\lambda]</th>
<th>SSBE-POS [\lambda]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>4</td>
<td>0.455 [0.421, 0.490]</td>
<td>0.454 [0.421, 0.488]</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.552 [0.517, 0.587]</td>
<td>0.55 [0.515, 0.586]</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.683 [0.648, 0.718]</td>
<td>0.688 [0.653, 0.723]</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.728 [0.695, 0.761]</td>
<td>0.727 [0.692, 0.761]</td>
</tr>
<tr>
<td>MAD</td>
<td>4</td>
<td>0.055 [0.050, 0.059]</td>
<td>0.055 [0.050, 0.059]</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.044 [0.041, 0.046]</td>
<td>0.044 [0.041, 0.046]</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.041 [0.039, 0.043]</td>
<td>0.041 [0.039, 0.043]</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.035 [0.034, 0.036]</td>
<td>0.035 [0.034, 0.036]</td>
</tr>
<tr>
<td>MAX</td>
<td>4</td>
<td>0.1 [0.093, 0.107]</td>
<td>0.1 [0.093, 0.107]</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.091 [0.085, 0.097]</td>
<td>0.09 [0.084, 0.096]</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.091 [0.086, 0.096]</td>
<td>0.092 [0.086, 0.097]</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.085 [0.080, 0.090]</td>
<td>0.085 [0.080, 0.090]</td>
</tr>
</tbody>
</table>

5. Comparison of SSBE with Maximum Likelihood Estimation (MLE)

Anupindi et al. (1998) develop a MLE procedure to obtain estimates of both demand and substitution parameters. MLE with perpetual system data requires detailed transaction data for each sale. More precisely, data corresponding to the time of a stock-out, the identity of the product that runs out, and the cumulative sales of other products up to that time are known in each replenishment cycle. We generated 100 replications of the problem data for each parameter set, and compared the performance of SSBE with complete information to that of MLE with perpetual system data over the same set of problems.

In our setting we allow substitution among products, therefore the service level used in calculating re-order levels will not probably be realized. However, in comparing the performance our estimation method with that of the Maximum Likelihood Estimation (MLE) of Anupindi et al. (1998) (see Section 5), the same demand realizations and re-order levels are used. Therefore, this will not pose any problem regarding the comparison.

We first compare performances under the market-share based substitution matrix. Table 5 provides detailed information about error values obtained in a 3-product category with three different
combinations of demand values (\(\lambda_A = 3.3, \lambda_B = 3.3, \lambda_C = 3.3; \lambda_A = 2.5, \lambda_B = 2.5, \lambda_C = 5\); and \(\lambda_A = 1, \lambda_B = 3, \lambda_C = 6\)), three different POS interval lengths (0.01, 0.1, and 0.25 hours), three different probability of substitutions (0.2, 0.4, and 0.6), and two different simulation lengths (3 and 6 months). In Table 5, we report the errors with respect to the real substitution matrix. Since the errors with respect to the realized substitution matrices are very similar, they are not reported here. The reader is referred to Öztürk (2006) for a detailed comparison with respect to both real and realized substitution matrices.

Analysis of the results indicates that \(MAD\) and \(MAX\) values are increasing in the probability of substitution, whereas \(MAPE\) values are decreasing. The results also suggest that the error values are generally decreasing in demand rate homogeneity, and in the simulation length for both of the methods. The analysis of the figures reveals that the POS interval length has a very slight impact on the error values generated by SSBE, which is very favorable for the practical use of the method.

We also note that \(MAPE\) values seem to be very high, and this is due to the fact that even small deviations from small-valued substitution probabilities result in high percentage errors. For instance, consider the case where \(\lambda_A = 1, \lambda_B = 3, \lambda_C = 6\) and the probability of substitution is 0.2. In this particular instance, \(\alpha_{AB}\) is equal to 0.0286. If the method estimates this substitution rate as 0.06 then the absolute percentage error for this substitution rate will be approximately 110 percent. Therefore, even if \(MAPE\) values are very high, especially when the probability of substitution is low, the absolute deviations are relatively smaller.

When the performances of the two methods are compared, it is observed that they perform similarly when the probability of substitution is medium. When the probability of substitution is small, SSBE performs better, and MLE generates smaller mean error values when the probability of substitution is large. The poorer performance of SSBE when the substitution probability is large appears to be related to the decrease in the numbers of observed substitution states. We note that,
when the substitution probability is high, the inventory is depleted much faster, and we observe fewer POS intervals in which substitutions take place.

In Table 6, we report performance of SSBE under random, adjacent, and one-item substitution models. We note that, because some substitution probabilities can be equal to zero under adjacent and one-item substitution models, \textit{MAPE} values can be undefined, and we report the \textit{MAD} and

<table>
<thead>
<tr>
<th>Probability of Substitution</th>
<th>Demand (units/hour)</th>
<th>Estimation Model</th>
<th>3-Months</th>
<th>6-Months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MAPE</td>
<td>MAD</td>
<td>MAX</td>
</tr>
<tr>
<td>0.2</td>
<td>[3,3,3,3,3]</td>
<td>MLE</td>
<td>96.677</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.01</td>
<td>83.093</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.1</td>
<td>83.754</td>
<td>0.084</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.25</td>
<td>86.302</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>[2,5,2,5,5]</td>
<td>MLE</td>
<td>104.150</td>
<td>0.100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.01</td>
<td>89.455</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.1</td>
<td>87.841</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.25</td>
<td>88.107</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>[1,3,6]</td>
<td>MLE</td>
<td>140.542</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.01</td>
<td>114.980</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.1</td>
<td>113.591</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.25</td>
<td>115.347</td>
<td>0.108</td>
</tr>
<tr>
<td>0.4</td>
<td>[3,3,3,3,3]</td>
<td>MLE</td>
<td>71.615</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.01</td>
<td>66.184</td>
<td>0.130</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.1</td>
<td>66.378</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.25</td>
<td>67.315</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td>[2,5,2,5,5]</td>
<td>MLE</td>
<td>85.673</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.01</td>
<td>79.318</td>
<td>0.140</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.1</td>
<td>80.466</td>
<td>0.143</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.25</td>
<td>81.541</td>
<td>0.144</td>
</tr>
<tr>
<td>0.6</td>
<td>[3,3,3,3,3]</td>
<td>MLE</td>
<td>47.857</td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.01</td>
<td>48.135</td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.1</td>
<td>48.420</td>
<td>0.145</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.25</td>
<td>49.472</td>
<td>0.148</td>
</tr>
<tr>
<td></td>
<td>[2,5,2,5,5]</td>
<td>MLE</td>
<td>51.999</td>
<td>0.148</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.01</td>
<td>52.750</td>
<td>0.152</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.1</td>
<td>53.274</td>
<td>0.153</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.25</td>
<td>52.304</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>[1,3,6]</td>
<td>MLE</td>
<td>59.753</td>
<td>0.153</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.01</td>
<td>59.908</td>
<td>0.154</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSBE-0.25</td>
<td>60.321</td>
<td>0.157</td>
</tr>
</tbody>
</table>

Table 5 Comparison of performances with respect to real substitution probabilities: SSBE vs. MLE.
MAX values only. The results reported in Table 6 indicate that the method’s performance depends more on the value of substitution probability $\psi$ rather than the customer choice model.

Table 6  Comparison of performances under different substitution structures.

<table>
<thead>
<tr>
<th>Substitution Structure</th>
<th>Probability of Substitution</th>
<th>Demand (units/hour)</th>
<th>MAD</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>MLE</td>
<td>SSBE</td>
</tr>
<tr>
<td>Adjacent</td>
<td>0.2</td>
<td>[3.3,3.3,3.3]</td>
<td>0.064</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.5,2.5,5]</td>
<td>0.072</td>
<td>0.064</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1,3,6]</td>
<td>0.082</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>[3.3,3.3,3.3]</td>
<td>0.08</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.5,2.5,5]</td>
<td>0.076</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1,3,6]</td>
<td>0.097</td>
<td>0.102</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>[3.3,3.3,3.3]</td>
<td>0.081</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.5,2.5,5]</td>
<td>0.083</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1,3,6]</td>
<td>0.098</td>
<td>0.115</td>
</tr>
<tr>
<td>One Item</td>
<td>0.2</td>
<td>[3.3,3.3,3.3]</td>
<td>0.051</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.5,2.5,5]</td>
<td>0.05</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1,3,6]</td>
<td>0.07</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>[3.3,3.3,3.3]</td>
<td>0.053</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.5,2.5,5]</td>
<td>0.06</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1,3,6]</td>
<td>0.07</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>[3.3,3.3,3.3]</td>
<td>0.055</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.5,2.5,5]</td>
<td>0.06</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1,3,6]</td>
<td>0.071</td>
<td>0.068</td>
</tr>
<tr>
<td>Random</td>
<td>0.2</td>
<td>[3.3,3.3,3.3]</td>
<td>0.073</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.5,2.5,5]</td>
<td>0.079</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1,3,6]</td>
<td>0.093</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>[3.3,3.3,3.3]</td>
<td>0.094</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.5,2.5,5]</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1,3,6]</td>
<td>0.112</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>[3.3,3.3,3.3]</td>
<td>0.097</td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.5,2.5,5]</td>
<td>0.105</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1,3,6]</td>
<td>0.119</td>
<td>0.135</td>
</tr>
</tbody>
</table>

In summary, SSBE and MLE exhibit similar behavior to changes in problem parameters, and yield more or less the same error values in most of the cases.
6. Conclusion

In this paper, we presented a practical method that can be applied under very general conditions to estimate product substitution probabilities in a category by using only POS data.

Our estimation method is based on a set of sub-states and a related state structure. First sub-states are determined given the corresponding amount of sales of each product. Subsequently, states are labeled for different combinations of sub-states. After the identification of states, an optimization problem is solved to estimate demand rates and product substitution probabilities.

Our computational analysis has shown that the accuracy of the estimates is reasonably good when the estimated substitution rate per unit time is not very low.

The results of our computational study indicate that, in terms of $MAPE$, $MAD$, and $MAX$, SSBE generally performs better than MLE when the probability of substitution is low and performs worse when it is high. For a medium probability of substitution their performances are close. Moreover, we observed that the two methods behave similarly when problem parameters’ values change (except the POS interval length, which does not affect MLE estimations).

Based on our extensive computational analysis, we propose SSBE as a practical estimation method to determine the stock-out based substitution probabilities by using only the POS data.

Acknowledgments

The authors thank the anonymous referees, the anonymous associate editor, and the Department Editor for their valuable comments and suggestions. This research was sponsored by TÜBİTAK grant 106K175.
References