Fordham University
Department of Economics
Discussion Paper Series

Improved Baseline Sales

Kurt A. Jetta
TABS Group

Erick W. Rengifo
Fordham University, Department of Economics

Discussion Paper No: 2009-02
February 2009

Department of Economics
Fordham University
441 E Fordham Rd, Dealy Hall
Bronx, NY 10458
(718) 817-4048

IMPROVED BASELINE SALES

Kurt A. Jetta[1] **Erick W. Rengifo**[2]

December 2008

Abstract

This paper develops a more accurate and robust baseline sales (sales in the absence of price promotion) using Dynamic Linear Models and a Multiple Structural Change Model (DLM/MSCM). We first discuss the value of utilizing aggregated (chain-level) vs. disaggregated (store-level) point-of-sale (POS) data to estimate baseline sales and measure promotional effectiveness. We then discuss the practical advantage of the DLM/MSCM modeling approach using aggregated data, and we propose two tests to determine the superiority of a particular baseline estimate: the minimization of weekly sales volatility and the existence of no correlation with promotional activities in these estimates. Finally, we test this baseline against the industry standard ones on the two measures of performance. Our tests find the DLM/MSCM baseline sales to be superior to the existing log-linear models by reducing the weekly baseline sales volatility by over 80% and by being uncorrelated to promotional activities.

JEL Classification codes:

Keywords: Dynamic linear Models, Multiple Structural Change Model, Consumer Packaged Goods, Marketing, Sales, Promotions. Baseline Sales.

[1] President and CEO, Tabs Group. 2 Corporate Drive, Suite 254, Shelton CT 06484, USA. Phone +1 (203) 925 9157, fax: +1 (203) 929 3643, e-mail: kurtjetta@tabsgroup.com.

[2] Assistant Professor, Department of Economics at Fordham University. 441 East Fordham road, Dealy Hall E513. Bronx, NY 10458, USA. Phone +1 (718) 817 4061, fax: (718) 817 3518, e-mail: rengifomina@fordham.edu

1 Introduction

In the United States, the Consumer Packaged Goods industry (CPG) accounts for over $500 Billion in annual retail sales according to A.C. Nielsen, and at least twice that world-wide. It is well-documented that retailer price promotions (defined as a temporary reduction in retailer price for a specific set of products for a specific set of time) account for the largest share of CPG firms' marketing budget (Cannondale 2007), and that percentage has grown consistently over time. Industry estimates peg the amount of annual spending on retailer price promotions at about $50-75B annually in the U.S. (about 15-20% of factory sales according to Accenture) and over $100B worldwide.

The CPG industry has among the most extensive information infrastructures of any industry. Most U.S. retail outlets are able to track the sales of virtually every product that is sold in the store with the use of scanners. These scanners can read the Universal Product Code (UPC) on each product, and the UPC is matched to information that describes dozens of characteristics about the product: manufacturer, brand, product type, flavor, weight, count size, and so on. The in-store scanner data is augmented by household-level scanning data (aka panel data) from over 100,000 U.S. households. This panel data is used to generate even more granular information on the consumer purchasing process. Two major firms, Information Resources (aka IRI) and Nielsen, have created a multi-billion dollar industry by collecting much of this information and selling it to manufacturers, retailers and other interested parties.

Armed with this information, manufacturers, retailers and academics have developed extraordinarily detailed models to measure the effectiveness of promotions and other marketing tactics like consumer advertising, price changes and public relations. A common denominator of all these models is that in order to determine the effectiveness of a given marketing tactic one needs to determine first the benchmark baseline sales level, i.e. the expected sales in the absence of a particular marketing variable like price promotion. It is worth noting that the baseline sales are simply the counterfactual of sales activity in the hypothetical case of no promotions for a period of time.

In this paper we propose a new model to estimate baseline sales and compare it to the two models that are considered to be the industry and academic standard: Scan*Pro (Wittink et al (1988)) and PromotionScan (Abraham and Lodish (1993)). Scan*Pro and

PromotionScan were developed in conjunction with Nielsen and IRI. Both are log-linear models that provide estimates of baseline sales and sales response as a function of specific retailer promotional tactics like price discounts, feature ads and displays. According to Bucklin and Gupta (1999) and Hanssens, Parsons and Schultz (2000), both models are fundamentally similar.

While there have been no formal academic challenges to the validity in the model, there are obvious data limitations in terms of quality and availability. The use of disaggregated data could potentially have measurement errors. Moreover, it is generally recognized by CPG practitioners and consultants that the baseline sales generated by these models are flawed[3] in that they yield "phantom" spikes. That is they show increases in baseline sales exactly concurrent with promotional activity when the expectation is that no such spike should occur. In Figure (1) we show one example of regular phantom spikes. Later in this paper we explain the reasons why baseline sales are supposed to be relatively stable estimates over time and why baseline estimates should be uncorrelated with promotional activity.

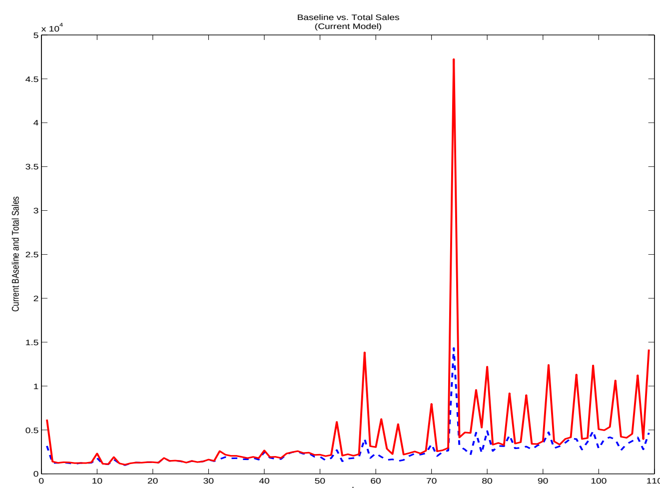

FIGURE 1: This figure presents the sales (solid line) and the actual estimate of the baseline sales generated using Scan*Pro (dashed line) for an adult personal care product.

3 Judgment is based on dozens of formal and information interviews of practitioners in CPG. Several of the interviewees are available to discuss their assessments upon request.

The contribution of this paper is the introduction of a methodology that leads to a more robust, less costly and more accurate estimate of baseline sales. This contribution to the existing literature is important since any measure of promotion performance depends directly on the baseline sales estimate. Flaws in the existing baseline model understate the incremental sales impact of price promotions and overstate the overall level of baseline sales. We implement the new baseline model using two econometric techniques: the Dynamic Linear Model (DLM) based on Ataman, Mela and Van Heerde (2007) which we improve with the use of the inclusion of endogenous dummy variable to flag promotional activity (Jetta – 2008) as well as a Multiple Structural Change Model (MSCM) of Bai and Perron (2003).

On the empirical side, this paper proposes to make several important contributions to the field. First, a better baseline estimate will help managers make better spending decisions on their promotion budgets. Second, the baseline method will be extendable to a broader section of retailers to include Club Stores and Category Killers (like Home Depot and Staples). Third, the baseline model can reach thousands of small-to-mid sized manufacturers that cannot afford the significant investment required to purchase baseline estimates from the major syndicated data suppliers. Fourth, the DLM-MSCM is new in marketing applications and this paper adds these useful tools in econometric analysis to the body of knowledge in marketing research.

The rest of this paper is structured as follows. Section 2 discusses the use of aggregated versus disaggregated data, describes the actual baseline model and its flaws and presents some desirable properties that any baseline sales should have. Section 3 presents the econometric techniques used in constructing the new sales baseline. Section 4 shows the empirical results obtained and, Section 5 concludes and presents future research ideas.

2. The Baseline Sales

Fundamental to the analysis of any marketing tactic is the concept of baseline sales. In order to determine if a causal variable generated some effect on sales, the analyst needs a reasonable estimate of what sales would have been without the existence of the causal variable (the counterfactual). Therefore, baseline sales is defined as an estimate of sales in the absence of specific promotional activity for a specific product and for a determined period of time.

In this section we present a brief discussion about the use of aggregated (chain or market-level) versus disaggregated (store-level) data, and we point out the reasons of why one should prefer to work with aggregated data. Later we present the actual baseline model and its flaws. Finally, we introduce two tests that a desirable baseline sale should satisfy.

2.1 Aggregated vs. Disaggregated Data for Baseline Modeling

Starting with Wittink et al (1988) and Abraham and Lodish (1993), the use of disaggregated data standard was established and the research paradigm that only disaggregated data should be used for marketing model estimation persists to this day. The research on the issue – Christen et al (1997), Foekens et al (1994), Van Heerde, Leeflang and Wittink (2002) - maintained that there was a significant risk of parameter estimation bias by using aggregated data in non-linear models (Scan*Pro and PromotionScan are log linear). This bias would imply, for example, that the estimate of the percentage increase in sales from display activity using aggregated data might be overstated.

These authors concede however, that the use of aggregated data holds several appealing properties in the areas of cost, availability, modeling flexibility, processing time and overall compliance and acceptance by practitioners. Christen et al (1997) suggest a debiasing procedure that can be used for market-level data. They mentioned nothing, however, on the more important issue of chain-level aggregation. Therefore, there has been almost no use of the debiasing procedure in other literature, and the conventional wisdom remains that disaggregated data is always optimal for modeling.

A further discussion of the practical shortcomings of disaggregated data is in order. Most importantly, disaggregated data is not aligned with the standard of management accountability, which is aggregated data either at the chain or market-level. The research tools have not been developed to predict and explain results at that level.

While there is clearly a role for the use of disaggregated data, the initial discovery process should occur at the group (aggregated) level to determine the total effects of programs in which managers are most interested. Unfortunately, the disaggregation paradigm means that most promotional researchers have overlooked aggregated effects entirely.

The second major limitation of disaggregated data is in the area of cost and availability. Currently, very few parties have access to this data. In the academic world, there are just two databases - University of Chicago Dominick's Database and the Stanford Basket Dataset – with this information. This static universe of data available for research limits the opportunity to check the robustness of existing results and to test new hypotheses. From both a commercial and academic standpoint, there are significant processing constraints to modeling store-level data unless there is a costly computer hardware investment in processing the massive database. This constraint is why most econometric models in the literature are built on databases of 30 stores or less. Even then, the Dynamic Linear Model used by Ataman, Mela and Van Heerde (2007) took several weeks to process a 30 store database. This precludes any meaningful analysis of a 6000+ store chain like CVS for all but the most powerful of hardware and software.

From a modeling standpoint, disaggregated data only has a marginal advantage vs. aggregated data. Van Heerde, Leeflang and Wittink (2002) state that the primary reason for using disaggregated data is to ensure that there is no estimation bias of parameters when the independent variables are heterogeneous. Accordingly, as long as marketing activity is implemented homogenously there is very little risk of biased estimation. Furthermore, even with heterogeneous marketing activity the magnitude of the bias depends on the percentage of stores promoted: the bias decreases as the percentage of stores promoted becomes larger (Van Heerde el al (2002)). From a practical standpoint, most chains execute Ads and Price Reductions, homogeneously. That is, every store within a chain receives the same marketing stimulus. For example, for the Adult personal care category studied in this paper, 86% of the 34008 observations with some level of Feature activity had ACV (All Commodity Volume) percentages of 80% or more. This observation is particularly valid for the US and Canada.[4]

In summary, disaggregated data contains severe practical and quantitative limitations that preclude it from being the sole or even primary data source for marketing research. Particularly given the homogeneity of most marketing stimulus, aggregated chain-level data should be appropriate for most applications.

2.2 The Existing Baseline

Bucklin and Gupta (1999) point out that many practitioners believe the baseline measure to be an actual number, when, in fact, it is a modeled measure. A modeled measure presents difficulties in determining whether the measure is accurate, since there is never any actual data to validate against. The first benchmark of measurement is intuition and judgment. That is, does the baseline appear to measure sales in the absence of promotion? In discussions with dozens of practitioners over the years, many expect baseline sales to be relatively stable, similar to sales trends they see during sustained periods without sales promotion.

A Baseline sales estimate can range in sophistication from the back-of-the-envelope "guess" to complex, econometric models that require a lot of data input and computer processing power. In the CPG industry, the two industry standard models are Scan*Pro (Wittink et al - 1988) and PromotionScan (Abraham and Lodish - 1993). Both are log-linear models that are fundamentally similar (Bucklin and Gupta (1999); Hanssens, Parsons and Schultz (2000)). Both models regress the log of unit sales against log price and dummy variables for other promotional effects like display or feature activity.

Van Heerde, Leeflang and Wittink (2002) lay out the original version of the Scan*Pro model. This model is non-linear, hence the authors' concern about parameter bias. Taking the natural log of this model provides the opportunity to conduct normal OLS regression on the data. The authors imply that the model is simple, as it was the first

[4] It is important to note that some industry experts feel that feature advertising in Europe is implemented heterogeneously by several major retailers and disaggregated data would more appropriate in those instances.

step in an evolutionary model building process. Later, they expand this model by introducing dynamics either through time varying parameters or via the inclusion of leads and lags. Their model also incorporates cross-brand promotional effects from numerous brands. In the latest published version of this model, the dependent variable can be a function of hundreds of independent variables once all the cross brand and timing variables are considered.

Van Heerde, Leeflang and Wittink (2002) present a baseline sales based on their extended model where they include four weeks of leads and lags to the original model in order to "accommodate the illusive post-promotion dip".[5] Their estimated baseline sales also show several sharp dips and spikes. Within a 10 week period the baseline deviates by +/- 12% around the median level for the period. They contend that this dynamic effect is "consistent with expectation" since promotional lifts tend to reduce post-period baseline sales. However, they do not mention any explanation for the "phantom" spikes observed in their baseline sales.

Just as there are significant practical limitations to disaggregated POS data, there are major practical disadvantages of the log-linear models, regardless of data source. Collection of these causal inputs requires a multi-million dollar data-gathering infrastructure (Sources: IRI Annual Report, 2003; Nielsen Annual Report, 1999). This presents a very high barrier to categories and retailers not currently included in that infrastructure. Retailers not included in this infrastructure account for over 60% of the retailers in the top 50 in the U.S according to MVI, including the largest one, Wal-Mart. Similarly, there are many products with sales outside of the IRI/Nielsen infrastructure that cannot be modeled using the existing modeling standard.

Other models have been used in the literature to calculate baseline sales, but none has been offered as a formal alternative to the industry-standard log linear models. Nijs et al (2001) and Pauwels et al (2002) both developed baseline models using a VARx where baseline sales are implied from the sales forecast for time t. They then use Impulse Response functions for each promotion to gauge the incremental effects for periods t, t+1,

[5] The authors call it the "illusive post-promotion dip" because they acknowledge that the dip, which is widely accepted to be true, is rarely evident from inspection of aggregated POS data. For more details on the issue, see Jetta (2008).

t+2, and so on. Ataman, Mela and Van Heerde (2007) used a Dynamic Linear Model (DLM) to estimate Baseline sales in a model for decomposing the effects of various marketing mix elements in new brands. Both of these models are confined to specific academic applications.

2.3 Validity Standards for Baseline Sales

After observing sustained periods of no promotional activity for certain brands, we propose that baseline sales should be relatively stable over time, in the absence of any major structural shift in sales (e.g. increased retail distribution). However, we observe baseline sales from the syndicated data supplier that are vastly different from the expectation of a stable baseline over time. Figure (2) shows an example of this lack of stability (i.e. volatility).

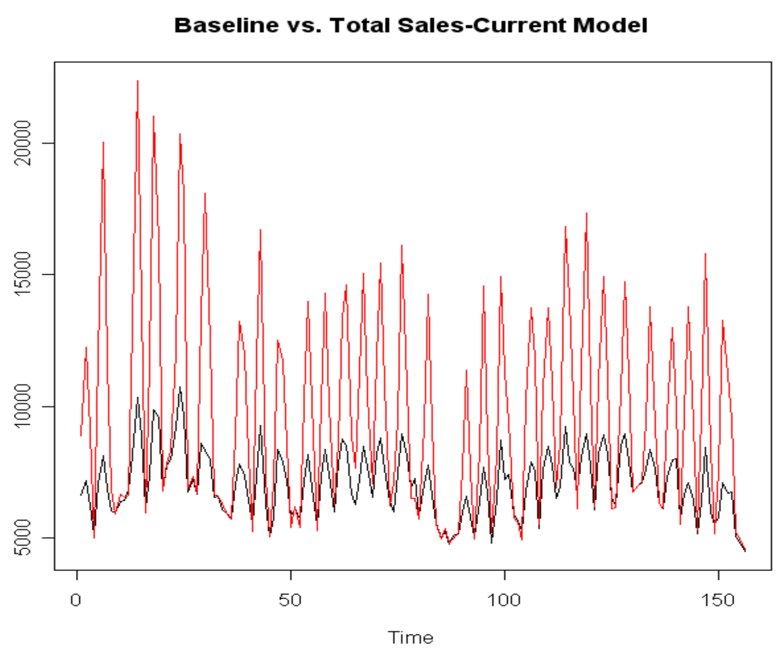

FIGURE 2: This figure presents the sales (solid line) and the actual estimate of the baseline sales generated using PromotionScan (dashed line) for a frozen food product.

Additionally, we see that this baseline sales exhibits "phantom spikes" concurrent with promotional activity. There is no reason to expect high correlation between the expected sales in the absence of promotions with "promotional activity" except in cases where manufacturers consistently execute other marketing programs not tracked in the data gathering process (e.g. FSI's or single-week TV Advertising) during those promotion weeks. Instances where manufacturers are able to do this *consistently* are rare. If we assume that the baseline sales are correct it would be better for sales managers not to do *any* promotional activities since just by not doing them, the sales would "naturally" increase.

We propose four tests that will prove the proposition that a superior baseline model should demonstrate minimal weekly volatility and be uncorrelated with promotional activity. We first test whether non-promotion weeks do, in fact, have less sales variability than those weeks with promotional activity for a specific retail chain. We then establish a test to determine which baseline model (log-linear vs. DLM/MSCM) has lower variability. Finally each model will be tested for evidence of correlation between promotional activity and baseline sales. A superior baseline model implies that promotions and baseline sales should be contemporaneously uncorrelated. Note that in each test we are ensuring that the equality is always placed in the null hypothesis. This increases the power of the test, that is, the likelihood of rejecting a null hypothesis that should be rejected.

We formally define the following null hypotheses:

<u>H0</u>: *Weekly sales variability during non-promoted weeks is equal to or greater than weekly sales variability of non-promoted weeks.*

$\sigma_{L(NONPROMO)} \geq \sigma_{L(PROMO)}$,

where, $\sigma_{L(NONPROMO)}$ ($\sigma_{L(PROMO)}$) is the standard deviation of the natural log differences of sales during non promotion (promotion) weeks. <u>It is expected that the null hypothesis will be rejected.</u>

<u>H1</u>: *The proposed baseline model has volatility that is equal to or greater than the existing baseline model.*

$\sigma_{L(Bn)} \geq \sigma_{L(Be)}$,

Where, $\sigma_{L(Bn)}$ ($\sigma_{L(Be)}$) is the standard deviation of the natural log differences of sales for the new (existing) model baseline model. It is expected that the null hypothesis will be rejected.

H2: *The current baseline sales estimate is not contemporaneously correlated with promotional activity.*

Correlation(B_{ert},φ_{rt}) = 0,

Where B_{ert} is the baseline sales estimate for the existing model (e) in retailer (r) at time (t). It is expected that the null hypothesis will be rejected.

H3: *The new baseline sales estimated is not contemporaneously correlated with promotional activity.*

Correlation(B_{nrt},φ_{rt}) = 0,

Where B_{nrt} is the baseline sales estimate for the new model (n) in retailer (r) at time (t). It is expected that the null hypothesis will be accepted.

3. Econometric implementation

This section presents the econometric techniques used to create the new baseline sales. The first subsection introduces the Dynamic Linear Model (DLM) used by Ataman, Mela and Van Heerde (2007) as well as the endogenous dummy variable innovation proposed by Jetta (2008). The second subsection presents the technique developed by Bai and Perron (2003) to detect multiple structural changes (MSCM). Our basic modeling strategy is to first flag weeks with high promotional activity with a dummy variable, then we detect structural changes in the data, and then we apply piece-wise DLM to the regimes that were found using the former technique while recursively enhancing the dummy variable estimation.

3.1 Dynamic Linear Models

DLM is a modeling technique pioneered by West, Harrison and Migon (1985) to address time series problems. The technique uses a recursive Bayesian approach to provide probability parameters to each observation in a time series. Each parameter estimated from the observation equation is based on the conditional probabilities of the state equations of prior periods. As more time periods are added to the model, the parameters are recursively refined to minimize the forecasting error.

From a marketing modeling perspective, Ataman, Mela and Van Heerde (2007) offer the following advantages of DLM: first, it has greater statistical efficiency with parameter evolution and explanation in one step; second, there is no need for pre-steps (like unit root testing) or assumptions on the distribution of error terms. This gives DLM an advantage over Kalman Filters, which require the assumption of normally distributed error terms; third, parameters update immediately as new data becomes available; fourth, missing data is accommodated relatively easy by using estimates from prior periods for imputation in the missing data; fifth, the technique allows for subjective information. Prior expectations can be overridden to accommodate anomalies in the data and, sixth, the model accommodates longitudinal as well as cross-sectional heterogeneity.

The disadvantages of DLM involve issues related to the implementation of the model rather than any statistical weakness. Specifically DLM's can be extremely processing intensive, where models can take days – even weeks – to run. Another minor disadvantage is that few software packages include a DLM.[6]

While Ataman, Mela and Van Heerde (2007) have not offered their baseline model as superior to the log-linear models, it has several elements which suggest a DLM is a better alternative to the current standard. Some of these elements include: First, it does not rely on an expensive infrastructure to gather causal measurements. Second, the model does not need to include any independent variables; therefore it is independent of potential data collection issues associated with the causal inputs. Its primary benefit is that it can be applied to any retailer that has scanner data without requiring a major data collection infrastructure. Third, the baseline estimates appear to be free of the volatility and phantom spikes we see in the log-linear models.

[6] This paper uses Matlab to program the DLM. The Multiple Structural Change model was developed using Gauss.

The Ataman, Mela and Van Heerde (2007) DLM model is:

$$Sales_t = \alpha_t + \beta_t * PIndex_t + v_t \qquad (1)$$

where,

$$\alpha_t = \lambda * \alpha_{t-1} + \omega_t \qquad (2)$$

$$\beta_t = \beta_{t-1} + \varepsilon_t \qquad (3)$$

Observed sales for a given week are a function of a dynamic baseline component (α) at time t and a dynamic promotion response evolution defined by β at time t. It is the evolution of the baseline that Abraham and Lodish (1993) identify as the first step in promotional response analysis. It is evident that this is a more parsimonious model than the log-models used in the industry. This model leaves open the possibility of additional exogenous variables, but as a first-generation model it is much simpler.

In general equations (1) and (2) can be written as:

$$Sales_t = \begin{bmatrix} 1 & PIndex_t \end{bmatrix} \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} + v_t$$

or

$$y_t = F_t'\theta_t + v_t \qquad (4)$$

and,

$$\begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} \omega_{1t} \\ \omega_{2t} \end{bmatrix} \qquad (5)$$

The core DLM equations are then:

$$(y_t \mid \theta_t) \sim \mathrm{N}(F_t'\theta_t, V_t) \qquad (6)$$

$$(\theta_t \mid \theta_{t-1}) \sim \mathrm{N}(\mathrm{G}_t'\theta_{t-1}, W_t) \qquad (7)$$

$$(\theta_0 \mid I_0) \sim \mathrm{N}(\mathrm{m}_0, \mathrm{C}_0) \qquad (8)$$

where I_0 is the initial prior information at time 0, including F_t, G_t, V_t and Wt. Moreover, at any future time t the available information set is $I_t = \{Y_t, I_{t-1}\}$. The posterior for some mean m_{t-1} and variance matrix C_{t-1} is given by:

$$(\theta_{t-1} \mid I_{t-1}) \sim N(m_{t-1}, C_{t-1}) \tag{9}$$

It can be shown that the prior at time t is:

$$(\theta_t \mid I_{t-1}) \sim N(a_t, R_t) \tag{10}$$

where, $a_t = G_t\, m_{t-1}$ and $R_t = G_t\, C_{t-1}\, G_t^{'} + W_t$. With this, the one step ahead forecast is given by:

$$(y_t \mid I_{t-1}) \sim N(f_t, Q_t) \tag{11}$$

with $f_t = F_t^{'}\, a_t$ and $Q_t = F_t^{'}\, R_t\, F_t + V_t$. Thus, the posterior at time t is:

$$(\theta_t \mid I_t) \sim N(m_t, C_t) \tag{12}$$

where $m_t = a_t + A_t\, e_t$, $C_t = R_t - A_t\, Q_t\, A_t^{'}$, $A_t = R_t\, F_t\, Q_t^{-1}$ and, $e_t = y_t - f_t$

DLM belongs to the Bayesian-type models and uses the Gibbs sampling. Gibbs sampling is an algorithm to generate a sequence of samples from the joint probability distribution of two or more random variables (Casella and George (1992)). In the case of the baseline model, each sequence generates samples for the Observation and the State Equations. By filtering both back and forth, we reduce the forecasting errors.

While Equations (1) to (3) provide a good starting point for a general baseline model, the inclusion of the Price Index variable presents some potential problems. The price is prone to measurement error: some retailers deduct promotion discounts off of the entire shopping order and do not assign them to a specific product. Additionally, many promotional vehicles like In-Ad Coupons, Rebates and Loyalty Card discounts have a history of tracking difficulties. From a retailer perspective, some stores may lower prices on a local basis for competitive reasons without the typical promotional support like shelf tags. Other retailers have non-traditional methods of handling Buy One/Get One Free consumer deals. Often both items will be scanned at full revenue with some other code denoting the BOGO offer. Even though, there is no evidence of systematic problems with price tracking in the syndicated data (except for a few isolated retailers) with so many potential shortcomings even for Own Brand promotion response, using the Price Index as an exogenous variable does not appear to be optimal.

Based on the previous shortcomings of the Price index as an appropriate explanatory variable, Jetta (2008) introduces an estimation technique for a dummy variable (φ) to account for promotional weeks, i.e. it takes on the value of 1 if it is a promotion week and 0 otherwise. Whereas the Price Index variable and other explanatory variables have problems with respect to acquisition costs and availability beyond CPG products carried in Food/Drug/Mass (ex Wal-Mart), the dummy variable technique has no such problems. It eliminates the data acquisition costs just by using weekly unit sales and eliminating all other causal inputs. Additionally, the measure is available to all retailers where scanner data is available.

This dummy variable (φ) is calibrated to flag any observation week where there is an abnormal deviation in weekly sales change or where the absolute sales level is significantly above the overall average. Our model runs through several ordinary least squares iterations to refine this variable in order to minimize the model standard error and maximizing its coefficient of determination. This variable is further refined during the DLM processing stage. In the appendix, we provide a complete description of the method.

Accordingly, the DLM model that we are going to use has the following observation equation:

$$Sales_t = \alpha_t + \beta_t\varphi_t + \gamma_{ti}I_t + v_t \qquad (13)$$

From here, the sales are a function of the dynamic baseline sales (α_t), Promotional activity (P_t) and other explanatory variables (I_t). By construction our baseline sales captures the unit sales in absence of promotions. The respective state equations are:

$$\alpha_t = \lambda_t * \alpha_{t-1} + \omega_{1t} \qquad (14)$$

$$\beta_t = \delta_t * \beta_{t-1} + \omega_{2t} \qquad (15)$$

$$\gamma_t = \rho_t * \gamma_{t-1} + \omega_{3t} \qquad (16)$$

Equation (14) presents the baseline evolution lift parameter. In Equation (15) we replace the price index as an explanatory variable with Promo dummy variable φ. This equation shows the dynamics for the lift parameter (β) and permits to test for promotional wear out effects over time. Equation (16) could potentially include category-specific dummies to control for seasonality, for example.

3.2 Multiple Structural Change Model

A weakness of the DLM introduced previously is that is not able to capture structural changes in sales. Structural changes occur when the demand of a given product increases by facts not directly related to promotions. From a practical standpoint, these structural changes are usually related to major increases or decreases in item-level distribution for a promoted brand. The main distinction between and structural change and a promotion shift is that the first implies a permanent shift meanwhile the later implies a temporal one. In order to capture this behavior, we complement the DLM with a technique proposed by Bai and Perron (2003) that allows us to capture multiple structural changes that could potentially be present.

Bai and Perron (2003) define a multiple linear regression with n breaks (n+1 regimes) as follows:

$$y_t = x_t^{'}\beta + z_t^{'}\gamma_j + \mu_t \qquad t=T_{j-1}+1, \ldots , T_j; \qquad j=1,\ldots,n+1 \qquad (13)$$

where y_t is the dependent variable observed at time t, x_t and z_t are vectors of covariates, (p x 1) and (q x 1), respectively. The vectors of covariates are β and γ_j. μ_t is the disturbance term at time t. The break points are identified by $T_1, \ldots , T_n$ and are treated as unknown variables. The unknown regression coefficients are estimated together with the break points when T observations on (y_t, x_t, z_t) are available. As the authors mention, this is a partial structural change model since β is not subject to changes and it is estimated for the complete sample. Setting p=0, gives rise to the pure structural model.

The estimation method is based on the least squares principle. For each n-partition, the associated least-squares estimates of β and γ_j are obtained by minimizing the sum of squared residuals:

$$(\mathrm{Y} - \mathrm{X}\beta - \mathrm{Z}\gamma)'(\mathrm{Y} - \mathrm{X}\,\beta - \mathrm{Z}\gamma) = \sum_{i=1}^{n+1} \sum_{t=T_{i-1}+1}^{Ti} [y_t - x'_t\beta - z'_t\gamma_i]^2 \qquad (14)$$

The authors showed that the break point estimators are global minimizers of the objective function.[7] For the estimation procedure they propose the use of an algorithm based on a dynamic programming principle that allows the computation of estimates of the break points as global minimizers of the sum of squared residuals.[8]

In this paper we propose to use the DLM and complement it with the multiple structural change model (MSCM). The main idea is first to detect structural changes in the data. Once this step is done, apply piece-wise DLM to the resulting regimes.

4. Empirical Application

In this section we compute our new baseline sales using the econometric techniques described above. For this application, we use aggregated data for adult personal care products and frozen foods. In this section we also present the Jetta (2008) technique, using ordinary least squares, to identify promotion weeks. This dummy variable will then be used as an independent variable in the model. After this, we present our new baseline sales and test it in the framework of the hypotheses presented in Section 2.

4.1 Data description

We use aggregated data at the retail chain level for two categories: adult personal care product and frozen foods. The data was gathered at weekly frequency from each of the major syndicated data suppliers (one category from IRI and one category from

---

[7] Following the authors, computing the estimates ($\hat{\beta}(\{T_j\})$ and $\hat{\gamma}(\{T_j\})$) on the n-partition (T_j) and substituting these in the objective function, the resulting sum of the squared residuals ($S_T(T_1, \ldots, T_n)$) and the estimated breakpoints ($\hat{T}_1, \ldots, \hat{T}_n$) are such that $(\hat{T}_1, \ldots, \hat{T}_n) = \arg\min_{T_1, \ldots, T_n} S_T(T_1, \ldots, T_n)$, where the minimization is taken over all partitions $(T_1, \ldots, T_n)$, such that $T_1, \ldots, T_n \geq q$.

[8] For a detailed exposition of the method we refer the reader to Bain and Perron (2003).

Nielsen). The data spans from 4/30/2006 to 5/25/2008 (109 weeks) and from 8/27/2005 to 1/21/2008 (125 weeks) for adult personal care products and frozen foods, respectively. We present the basic analytical grouping as a Dataclass, which are all weekly observations for a specific category, within a specific retailer for a specific brand. We have 312 Dataclasses in adult personal care products and 247 Dataclasses in frozen foods. The specifications of the data are as follows:

4.2 The endogenous Promo dummy variable

We propose the endogenous PROMO dummy variable (φ) as an alternative to the costly acquisition of the Percentage of Units on Any Promotion (PUAP) measure provided by the data supplier, a measure that is currently the industry accepted standard for detecting the presence of meaningful promotional activity. This dummy variable is calibrated to flag any observation week where there is an abnormal deviation in weekly sales change or where the absolute sales level is significantly above the overall average. The model runs through several iterations to refine this variable in order to minimize the model standard error or to maximize its coefficient of determination (R^2).[9]

To measure the accuracy of our method we compare this PROMO dummy to the PUAP provided by the data supplier. We run simple regressions where the data provided by the data supplier was treated as the dependent variable and the PROMO dummy variable as the explanatory one. In the case of the adult personal care product (frozen food) the average R^2 is 96% (94%), the coefficients were significant at the 5% significance level in all the cases and, also in all the cases the p-values of the F-test were smaller than 0.05. The minimum R^2 for adult personal care product (frozen foods) is 89% (91%), showing a robust result.

We can thus observe that there is a high level of convergence between the two measures of promotional activity, meaning that our estimated PROMO variable tracks very closely the syndicated values. With more than 90% accuracy in capturing high levels of promotional activity, the endogenous promotional calculation provides a viable substitute for the expensive causal measure infrastructure. So with confidence in the

[9] For a detailed description of the method we refer the reader to read Appendix A.

validity of the estimated promotional dummy, φ, we can proceed with our new baseline model.

4.3 Stationarity of the data

Even though the DLM model does not require stationarity of the time series data, it is a necessary condition for the multiple structural change model of Bai and Perron (2003). Each dataclass was tested for both level and trend stationarity. In total there were 559 Dataclasses across two categories (312 adult personal care product and 247 frozen foods). We conduct the Augmented Dickey-Fuller Unit Root test on each dataclass to test for stationarity in levels and considering a deterministic trend. The unit root results show that retail sales data is a trend stationary process, as 95.4% of the dataclasses did not have a unit root.[10]

4.4 The new baseline sales

We estimate our new baseline sales using a Dynamic Linear Model (DLM) where we model the unit sales as a function of a constant (the baseline) and a dummy variable that captures promotional activity.[11] However, as mentioned before, this model alone is not able to capture permanent changes in the unit sales. We improve this initial model using the Multiple Structural model of Bai and Perron (2003) to capture this feature of the data.

In this section we present and test the new baseline sales and compare it with the existing ones. The tests are performed under the hypotheses stated in Section 2. That is, an improved baseline estimate should exhibit low week-to-week variability and no contemporaneous correlation with promotional activity.

First, we test H0 where the null hypothesis is that weekly sales variability for low promotion weeks is the equal to or greater than the weekly sales variability for high promotion weeks. To test this hypothesis, each weekly observation within each dataclass

[10] Due to space limitation we do not write all the results. However, they are available upon request. For similar results we refer the reader to Jetta (2008).
[11] The model is also able to capture seasonality patterns with the inclusion of an additional variable in the main DLM model.

was divided into one of four quartiles based on the PUAP that week: class 1 (0-25.0%), class 2 (25.1-50.0%), class 3 (50.1-75.0%) and class 4 (75.1% - 100%). The PUAP is a measure directly pulled from the data supplier with no other manipulation to the figures. Table 1 provides the analysis of variance results by category.

Chain	(%) Unit on Promotion	Adult personal care			Frozen foods		
		Obs.	Standard deviation	p-value	Obs.	Standard deviation	Std Error
Quartile I	0-25%	15,478	9,849.0		18,685	1,972.7	
Quartile II	25.1-50%	3,755	11,554.0	0.00	4,774	2,566.3	0.01
Quartile III	50.1-75%	2,935	14,937.1	0.00	3,871	2,532.9	0.01
Quartile IV	75.1%-100%	5,244	15,157.6	0.00	3,545	2,968.4	0.00

Table 1: This table presents the standard deviation of unit sales according to the PUAP on a given week (broken into quartiles from lowest to highest values) and the p-values of the F-test for equality of variances. The null hypothesis, in all cases, is that the variances of any Quartile relative to Quartile I are equal. The results show that the standard deviations of all classes respect to class 1 are significantly different at the 5% confidence level.

Table 1 shows that in all the cases and for both categories (adult personal care products and frozen foods) the F-test for equality of variances is rejected at a 5% significance level. Moreover the variance of unit sales during low-promotion weeks is significantly different (and low) compared to highly promoted weeks (class 3 and 4). This result goes in hand with our empirical observation and is the basis for testing hypothesis H1.

In order to test hypothesis H1 we compare the variance of our new baseline sales with the existing industry standard models. The null hypothesis (H1) is that the new baseline sales volatility is equal to the volatility of the industry standard one. Figure 2 shows the results of our comparison using a histogram. This figure depicts the difference of the standard deviation in the log difference of our proposed baseline sales minus the standard deviation in the log difference of the existing baseline sales. We can see a dramatic reduction (over 80%, on average) in the variability in weekly baseline sales estimate using our new baseline model.

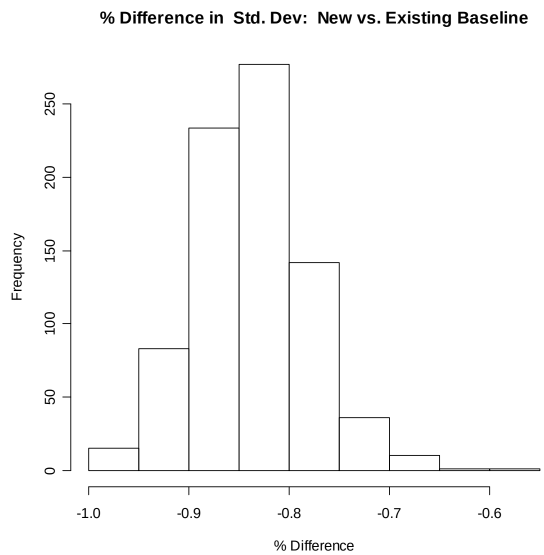

Figure 2: This histogram presents the frequencies of the difference of variability between our new baseline sales and the existing baseline sales model.

We also performed a test based on the four data classes described before, i.e. we tested the null hypothesis of equal variances of the new baseline sales and the Scan*Pro one per data class. In these cases 100% of the variances were significantly different at the 5%. Moreover, this difference is significantly larger in heavily promoted weeks (Class 4).[12] In conclusion, we reject the null hypothesis that the volatility of the new baseline is greater than or equal to the existing baseline.

The final two tests measure the existence of contemporaneous correlation between promotional activity and baseline sales. By inspection it appears that the major source of the variability in the existing baseline measures is due to this correlation, which we have referred to as the "phantom spike." H2 will test whether this "phantom spike" exists on a consistent basis for all Dataclasses.

We perform this test using the natural log differences in the weekly baseline sales for all weeks. We code each weekly observation as: first week of promotion and other

[12] Due to space restrictions we do not present the results here. They are available upon request.

promotion week. We make this distinction because the first week of promotion in a multi-week promotion usually exhibits a much high sales increase (lift) than subsequent weeks. The test will capture whether there may be a specific class of promotional observations where this correlation exists. Of course, this will not be an issue in single-week promotions.

Table 2 presents the average, maximum and minimum simple correlation between each individual baseline sales observation and its respective promotional activity, which is defined as $\varphi=1$. Moreover, this table also presents the pooled p-value of the t-test under the null of significant correlation. We construct the pool by using all available data (regardless of product) in a given group and performed the t-test for significant correlation.[13]

Week	**Scan*Pro**		**PromotionScan**	
	Av.corr/max/min	**p-value**	**Av.corr/max/min**	**p-value**
First-week promo	0.81/0.96/0.75	0.41	0.85/0.99/0.72	0.46
Other-week promo	0.74/0.88/0.69	0.37	0.79/0.92/0.74	0.42

Table 2: This table presents the average, the maximum and minimum correlation between the industry baseline sales (Scan*Pro and PromotionScan) and promotional activity. The data was divided into two groups: first-week of promotion and other promotion week. This table also presents the pooled p-value of the t-test under the null of significant correlation. The results show that there is significant correlation between the existing baseline models (Scan*Pro and PromotionScan) and promotional activity for both groups.

Table 2 presents the results for H2. We created two groups and analyzed the correlation between the respective baseline sales and promotional activity at the individual product level and at the aggregated level (regardless of product).

Observing the results in this table, we can appreciate the high level of significant positive correlation between both of the existing baseline sales and the promotional

[13] We also performed the t-test at the individual level. Results show that approximately than 95% of the individual products in the adult personal care products baseline sales are significantly not correlated with promotional activity for the first group (First-week promo) and 93% are significantly not correlated with promotional activities for the last group (Other-week promo).

activities. This feature is not desirable, since there is no reason to expect and increase in sales in the absence of any promotional activity. This high, positive and significant correlation of the existing baseline sales with promotion activities is evident by observing the existence of spikes and dips throughout, as shown in Figure 1.

Week	**Adult personal care**		**Frozen foods**	
	Av.corr/max/min	**p-value**	**Av.corr/max/min**	**p-value**
First-week promo	0.12/0.15/0.05	0.03	0.09/0.14/0.07	0.01
Other-week promo	0.08/0.12/0.05	0.02	0.06/0.11/0.03	0.00

Table 3: This table presents the average, the maximum and minimum correlation between the baseline sales and promotional activity for each individual product in each of the categories that we analyze (adult personal care and frozen foods). The data was divided into two groups: first-week of promotion and other promotion week. This table also presents the pooled p-value of the t-test under the null of significant correlation. The results show that there is no significant correlation between our new baseline and promotional activity for both groups (First-week promo and other-week promo).

We next performed the same test for the DLM/MSMC model. From Table 3, it is clear that that there is no significant correlation between our new baseline sales and promotional activity at the 5% significance level for both categories under analysis.[14]

Based on the results presented here, we reject H2 that the correlation between promotional activity and baseline sales for the existing model is zero. Conversely, we fail to reject the H3 that the correlation between promotional activity and the next baseline estimate is zero.

Finally, we also present similar information to the one presented in Tables 2 and 3 using a graphical comparisons. In Figure 3 we can see: a) Actual sales vs. new baseline (upper left); b) Existing baseline vs. new baseline (upper right); c. Actual sales vs. existing baseline (lower left), and d) actual sales vs. fitted sales from DLM. This figure represents one Dataclass from the frozen foods category.

[14] We refer the reader to Jetta (2008) for another test based on a linear regression analysis.

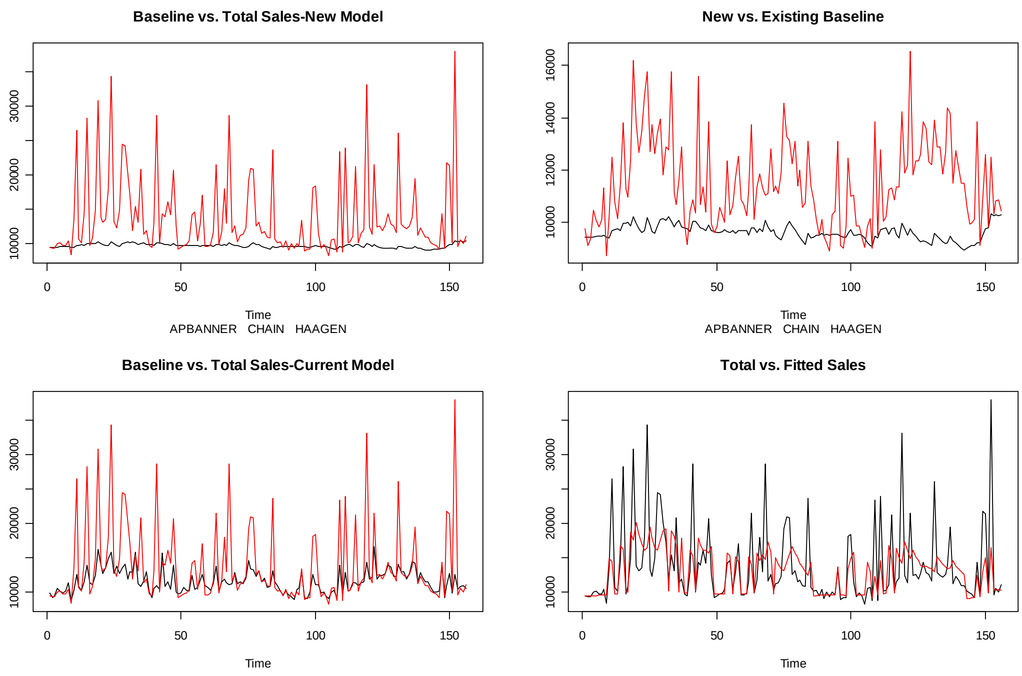

FIGURE 3: This figure presents the following plots: a. Actual sales vs. new baseline (upper left); b. Existing baseline vs. new baseline (upper right); c. Actual sales vs. existing baseline (lower left), and d. actual sales vs. fitted sales from DLM. This figure corresponds to a product that belongs to the frozen foods category.

We can see visually what was proven analytically: the new baseline is less volatile than the new one (upper right panel), and that the existing baseline sales highly covariate with actual promotion sales (lower left panel). This figure also presents the fitted unit sales created using *only* the DLM model. The main problem with our DLM specification is that is not able to capture structural changes (see upper left of Figure 4) when a product experiences a permanent change in its units sold due to changes in distribution, for example. In order to deal with this issue we propose to complement the DLM with the multiple structural change model (MSCM) of Bai and Perron (2003). The upper left graph in Figure 5 shows the results for the same product but with the baseline model computed using the DLM-MSCM.

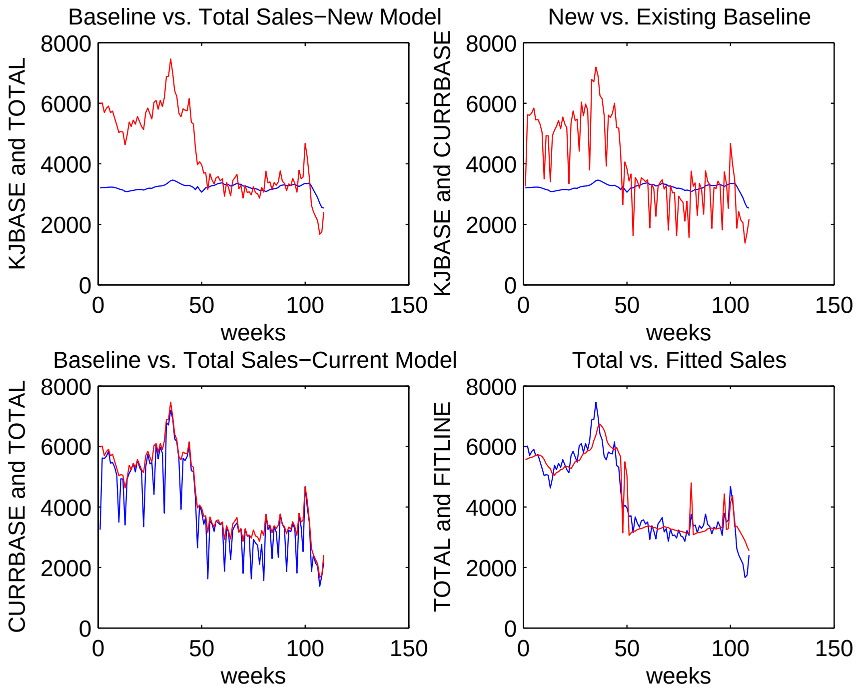

FIGURE 4: This figure presents the following plots: a. Actual sales vs. new baseline (upper left); b. Existing baseline vs. new baseline (upper right); c. Actual sales vs. existing baseline (lower left), and d. actual sales vs. fitted sales from DLM. This figure corresponds to a product that belongs to the adult personal care category. Observe that our baseline computed using only the DLM model is not able to capture structural changes that can be present in the data (upper right).

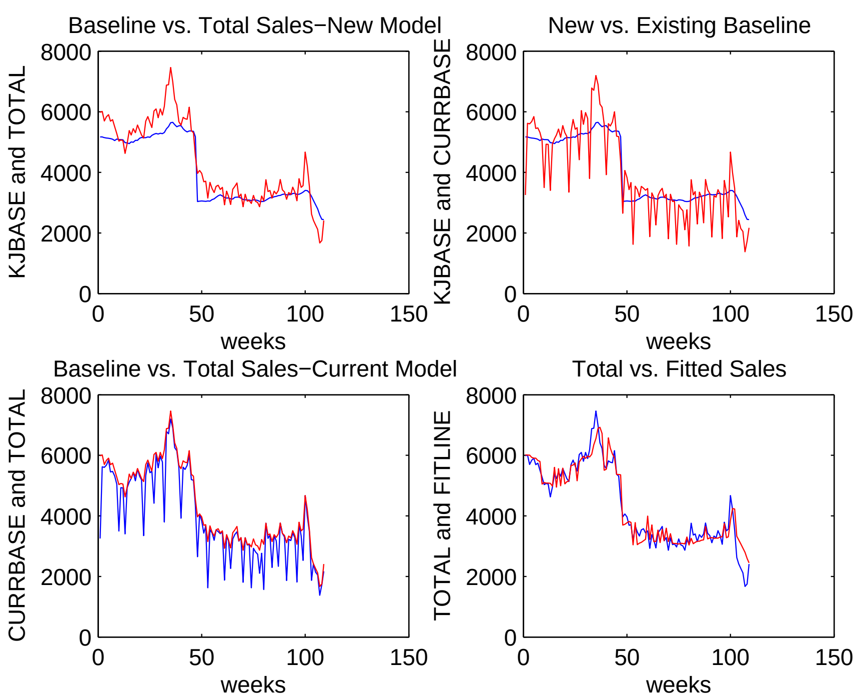

FIGURE 5: This figure presents the following plots: a. Actual sales vs. new baseline (upper left); b. Existing baseline vs. new baseline (upper right); c. Actual sales vs. existing baseline (lower left), and d. actual sales vs. fitted sales from DLM. This figure corresponds to a product that belongs to the adult personal care category and the industry baseline used corresponds to the Scan*Pro. Observe that our baseline computed using the DLM-MSCM model is now able to capture structural changes that can be present in the data (upper right).

Comparing both figures, it is clear that our new baseline model, implemented by the use of the multiple structural change model of Bai and Perron (2003) and the DLM model is able to capture that characteristic of the data quite accurately.[15] It is worth noting that the average R^2 of the fitted values (see lower left panel of Figures 4 and 5) without (with) the MSCM is 0.71 (0.86). This high average R^2 without the MSCM could imply that just by using the DLM we can have a significantly good fit of the data. However, this high R^2 (0.71) is due to the fact that for the data that we have available almost 90% of the Dataclasses do not present structural changes in unit sales. However, if we only consider the remaining 10% of Dataclasses (i.e. products with structural breaks in unit sales) the average R^2 using *only* the DLM model drops to 0.52. In the same vain, if we just consider this group with structural breaks and compute the baseline sales with the DLM-MSCM the R^2 jumps to 0.92. Thus, we dramatically improve the fit of our data using the mixture of these two models.

5. Conclusions and future research

In this paper we introduce a new technique to compute the baseline sales. This new technique consists on the use of a Dynamic Linear Model DLM (based on Ataman, Mela and Van Heerde (2007)) that is enhanced with an algorithm to flag promotional activity endogenously (Jetta – 2008) and, further, complemented with a Multiple Structural Change Model (MSCM) proposed by Bai and Perron (2003). This new baseline sale has many highly desirable properties for being considered as the expected sales in the absence of promotional activities: it has low variability and it is almost not correlated with promotional activities. Moreover, this new baseline model is able to capture structural changes that could be present in certain products after controlling for seasonality and other predictable patterns.

[15] Additional figures and proofs are available upon request.

We checked these desirable properties not only for our new proposed baseline model but also for the actual industry standard, the Scan*Pro and PromotionScan. Our findings show that the industry benchmarks lack of both these properties: they have high volatility and they are highly correlated with promotional activities.

We presented an empirical application studying two main categories: adult personal care products and frozen foods. Using aggregated data of 312 dataclasses in adult personal care products and 247 for frozen foods, we show how our baseline sale is able to capture a more reliable expectation about the sales in the absence of promotional activities, after controlling for seasonality. We can also observe from our results that the new model strategy perfectly capture structural changes in the data.

In summary, these tests provide compelling evidence of the superiority of the baseline sales using the DLM-MSCM compared to the existing industry standards. First, we demonstrated that non-promotion week shows lower level of sales variability than promotion weeks, particularly for chain-level data. Next, we showed that the new DLM-MSCM greatly reduced the level of volatility in the weekly baseline estimates. On average, the reduction in variability is around 80%. Finally, we demonstrated that the existing baseline sales based on a log linear model exhibited high correlation with promotion weeks (+75%) for Chain-level data. Meanwhile, our new DLM-MSCM has no significant correlation with promotional activities. Moreover, our baseline model is able to capture structural changes present in some products. In addition to the quantitative benefits of this model, it also has the advantage of not being reliant on an expensive data-gathering infrastructure for causal measures and it can be extended to any retailer and trade class which gathers weekly point-of-sale data.

Two potential limitations on this research are that it reflects only two categories, and that the research was done only in the US market. Future research involves testing this model for other CPG categories in order to generalize the results. Additionally, this model should be conducted in European markets where there is a belief by some that homogeneity of retailer promotional stimulus cannot be assumed for some chains. One other limitation is that this model does not purport to provide accurate estimates of the lift parameters associated with specific promotional tactics (e.g. the change is sales of a 20%

off promotion vs. 40% off). However, any model that seeks to quantify those lift parameters must first ensure that the baseline sales estimate is accurate. This research has demonstrated a new approach that greatly improves the baseline model accuracy.

APPENDIX 1

IMPROVED BASELINE ALGORITHM

1. Data Setup

    1.1. Assign a dataclass for each brand/geography combination

    1.2. Input exogenous variables (if needed)

    1.3. Read in and clean up unit sales.

    1.4. Within each dataclass create additional measures

        1.4.1. LnSales for each week

        1.4.2. Diff(LnSales) for each week (Week 1=0)

2. Generate Promotional Dummy Variables

    2.1. *Iteration #1 (Iterate within Dataclass)*

        2.1.1. Flag any week as PROMO (=1) where either (LnSales > Avg(LnSales) + SD(LnSales)) or (LnDiff > SD(LnDiff)

        2.1.2. Flag any week as POST-PROMO where 2.1.1. criteria not met, but Week t-1 is a PROMO (=1).

        2.1.3. Flag any week as NON-PROMO where neither 2.1.1. or 2.1.2. are valid.

        2.1.4. Run Regression of LnDiff against PROMO, POST-PROMO and exogenous variables

        2.1.5. Capture Residuals of the Regression

        2.1.6. Calculate SD (Std Dev) of Residual of each Factor (PROMO, POST-PROMO, NON-PROMO from 2.1.1.-2.1.3.)

        2.1.7. Calculate mean(LnSales) by Factor

        2.1.8. Calculate and Capture Model Error (Std Dev of Residuals)

2.1.9. Flag all observations where Residuals are +/- 1 SD of the applicable factor for that observation

2.2. *Additional Iterations*

2.2.1. If observation not flagged from 2.1.9. test, carry over Factor from last iteration.

2.2.2. For POST-PROMO and NON-PROMO factor observations, if Resid > 1 SD of Factor, then change Factor to PROMO

2.2.3. Capture POST-PROMO observations similar to 2.1.2.

2.2.4. Repeat 2.1.4 to 2.1.9. for all iterations.

2.3. Capture initial parameter estimates for Dynamic Model

2.3.1. After all iterations complete, identify the iteration for each dataclass with the minimum Standard Error

2.3.2. Keep the PROMO values associated with the minimum Standard Error iteration.

3. Generate Dynamic Parameter Estimates for the Observation Equation (each Dataclass)

3.1. Regress LnSales against PROMO (from 2.3.2.) and any exogenous variables and Capture key estimators.

3.1.1. Capture the Coefficients and Std Error estimates for Alpha (Intercept), Beta (Lift) and Gamma (exog parameter)

3.1.2. Capture the model Standard Error

3.2. Initialize the DLM model

3.2.1. Initialize the State Mean estimates with the coefficients from 3.1.1. Call this vector θ_t.

3.2.2. Initialize the State Variance estimates with the Variances from 3.1.2. (SD^2) . Call this vector ω_t.

3.2.3. Initialize the model variance with the Variance from 3.1.2. ($\sigma_t^2 = SE^2$)

3.3. *For first iteration, update weekly parameter estimates*

3.3.1. Update LnSales forecast ($E(S_t) = \theta_t \times X_t$)

3.3.2. Update model variance (Model + State var)

3.3.3. Calculate adjustment factor to apply to θ_t and ω_t. This will be the ratio of the State Variance vs. Model Variance. Call this A.

3.3.4. Update θ. Prior State Mean + (A * [S_t – $E(S_t)$])

3.3.5. Update ω. Prior State Var – (A^2 * σ_t^2)

3.4. *Additional DLM iterations (Gibbs Samples)*

3.4.1. Use parameter estimates from prior model

3.4.2. If PROMO = 0 (nonpromo) and Forecast deviation > 1.5 SD for σ_t then PROMO = 1

3.4.3. If PROMO = 1 (promoted) and Forecast deviation < 1.5 SD for σ_t then PROMO = 0

3.4.4. Rerun 3.3.1.-3.3.5.

Identify the iteration with the minimum variance and select the parameter estimates from that model for θ, PROMO and E(S).

BIBLIOGRAPHY

Abraham, Magid and Leonard Lodish (1993) “An Implemented System for Improving Promotion Productivity Using Store Scanner Data,” *Marketing Science*, **12**(3). 248-269

Ataman, Berk, Carl Mela, and Harald Van Heerde (2007), “Building Brands,” *Working Paper, Marketing Dynamics Conference, University of Groningen*.

Bai, J. and P. Perron (2003), “Computation and Analysis of Multiple Structural Change Models”, Journal of Applied Econometrics, (18), 1–22.

Bucklin, Randolph and Sunil Gupta (1999), “Commercial Use of UPC Scanner Data: Industry and Academic Perspectives,” *Marketing Science*, Vol. 18 (3), 247-273.

Casella, George and Edward I. George (1992) "Explaining the Gibbs sampler," *The American Statistician*, Vol. 46, 167-174.

Christen, Markus, Sachin Gupta, John C. Porter, Richard Staelin, and Dick R. Wittink (1997), “Using Market-Level Data to Understand Promotion Effects in a Nonlinear Model,” *Journal of Marketing Research*, Vol. XXXIV (Aug), 322-334.

Foekens, Eijte W., Peter S.H. Leeflang, and Dick R. Wittink (1994), “A Comparison and Exploration of the Forecasting Accuracy of Nonlinear Models at Different Levels of Aggregation,” *International Journal of Forecasting*, 10, 245-61.

Hanssens, Dominique, Leonard J. Parsons, and Randall L. Schultz (2000), Market Response Models: Econometric and time series analysis. International Series in Quantitative Marketing, Norwell, MA.

Jetta, Kurt A. (2008),, “A Theory of Retailer Price Promotions Using Economic Foundations: It’s All Incremental, *Ph.D. Dissertation,* Fordham University, Department of Economics.

Nijs, Vincent R., Marnik G. Dekimpe, Jan-Benedic E.M. Steenkamp, and Dominique Hanssens (2001), “The category-demand effects of price promotions”, *Marketing Science*, Vol. 20 No. 1, pp. 1-22.

Van Heerde, Harald J., Peter S.H. Leeflang, and Dick R. Wittink (2002), “How Promotions Work: Scan*Pro-Based Evolutionary Model Building,” *Schmalenbach Business Review*, Vol 54 (Jul), 198-220.

West, Mike, Jeff Harrison and Helio Migon (1985), “Dynamic Generalized Linear Models and Bayesian Forecasting,” American *Statistical Association*, Vol. 80, 77-83.

Wittink, Dick R., Michael Addona, William Hawkes, and John C. Porter (1988), “SCAN*PRO: The Estimation, Validation and Use of Promotional Effects Based on Scanner Data,” working paper, ACNielsen.