Journal of Business Research 69 (2016) 4394–4402

Contents lists available at ScienceDirect

Journal of Business Research

A new approach to measuring retail promotion effectiveness: A case of store traffic

Leonardo D. Epstein [a,*], Arturo A. Flores [b], Ronald C. Goodstein [c], Sandra J. Milberg [d]

[a] School of Business and Economics, Universidad de los Andes, San Carlos de Apoquindo 2200, Las Condes, Santiago, Chile
[b] Falabella Inversiones Financieras, Moneda 980, piso 20, Santiago, Santiago, Chile
[c] McDonough School of Business, Georgetown University, 37th and O Streets, NW, Washington, DC 20057, United States
[d] School of Business, Universidad Adolfo Ibáñez, Diagonal Las Torres 2640, Peñalolen, Santiago, Chile

ARTICLE INFO

Article history:
Received 1 February 2015
Received in revised form 1 December 2015
Accepted 1 March 2016
Available online 14 May 2016

Keywords:
Sales promotions
Promotional effectiveness
Store traffic
Baseline prediction
Video data

ABSTRACT

This article presents an approach to measure the effect of promotions on customer traffic, a measure of effectiveness of retail promotions important both to managers and scholars. This effect is unobservable because one lacks baseline measures of traffic: one cannot measure traffic simultaneously with and without a promotion. Hence, the assessment of this effect remains a challenge. However, adoption of imaging and other forms of electronic monitoring allows retailers to collect traffic data, until recently unavailable, on the behavior of both actual and potential customers. Making these data useful for business decisions requires new analytical methods. The approach of this research is novel in two ways: First, a counterfactual argument is the foundation to predict the baseline series. Second, the approach defines predicted residuals to estimate hourly-specific effects on traffic, which one computes after the promotion. Computing the baseline predictions uses a Poisson model with effect-parameters such as time of the day, day of the week, week of the month, secular trends and others, to capture sources of systematic variability. The article illustrates the use of simple plots to visualize and communicate the evolution of the promotion effects. An illustration uses data from Skillup-Chile, an imaging and analytics company.

1. Introduction

Managers in retail businesses invest in marketing or sales promotion campaigns. Often the results of such investments, returns for instance, are uncertain. Furthermore, even if an investment realizes a positive return, the fractions of this return one attributes to the campaign and to the variability of customers' behavior, respectively, are uncertain: decisions to enter the store, decisions to purchase, and the amounts they spend on their purchases are random and may lead to a positive return by chance. The effects of variable behavior, with either favorable or adverse effects, do not necessarily have an association with the campaign. This article puts forward a broadly applicable approach to make inferences about performance measures of marketing or promotion campaigns.

Retailers collect data to monitor and improve financial and operational efficiencies and in particular, to assess the performance of promotions. Such data are often available from cash-registers, for instance the times of sales and their corresponding amounts. The record of the times of sales and their corresponding values is a marked point process. Increasingly available video technologies allow retailers to collect data not available from other sources. For instance, Skillup-Chile provides a technology that places cameras at store entrances that transmit the images to a system that records the times customers walk into the store. The record of these arrival times is a point process.

The conversion of a store visitor into a customer that makes a purchase makes the point process of arrivals of visitors into the marked point process of sales. Lam, Vandenbosch, Hulland, and Pearce (2001) describe in detail the process a person follows when she walks through the store front until she makes a purchase, if she does so. A manager who plans a marketing/promotion campaign (MPC) attempts to increase at least one of the following: the traffic, the conversion rate, or the amounts of sales. For the sake of brevity, and because the areas of stochastic processes or time series use the term intervention to designate events or shocks that alter the behavior of a process, this article refers to an MPC as an intervention and to the time between the start and the end of the intervention, as the intervention interval.

A fundamental obstacle to examine measures of performance of MPCs is the absence of a control store that provides baseline measures to assess performance changes. To describe this obstacle more concretely, suppose the manager plans an intervention for the interval November 6–November 20, and records the monthly number of customers that enter the store, and that he would like to use the number of additional customers the intervention brings to the store as a performance or

* Corresponding author.
E-mail addresses: lepstein@uandes.cl (L.D. Epstein), aandflores@falabella.cl (A.A. Flores), goodster@georgetown.edu (R.C. Goodstein), sandra.milberg@uai.cl (S.J. Milberg).

http://dx.doi.org/10.1016/j.jbusres.2016.03.062
0148-2963/

effectiveness measure. But, if the interest of the manager is to use a different measure of performance, for instance the return of his investment in the intervention, then he can modify the approach of this article to make the assessment with his measure of interest.

For the intervention period the manager can compute neither the number of additional customers nor the return on his investment. Though, he knows the number of customers that arrive to the store, he cannot compute the number of additional customers the intervention brings. Similarly, he can compute the profits for the intervention period, but cannot compute the additional profits that the intervention brings. Clearly then, a manager cannot compute measures of performance of the intervention because he is missing the baseline measure of performance.

This article addresses the problem that the missing baseline poses to compute the effect of an intervention on performance: for the missing baseline the approach computes predictions with observations outside the intervention period only. In this fashion, one may interpret these predictions as the number of customers one would observe if the intervention did not occur. This description simplifies the approach for the sake of clarity. A more technical description, specific to traffic data is as follows: the data are a time series of counts of arrivals of people during one hour time windows, an arrival process. The approach uses the history of this arrival process before the intervention period to build the prediction distribution of the observations for traffic during the intervention period. After the intervention, the actual observations become available. Hence, with the observations and the predictions available for the intervention period the approach computes the differences between them, that is, for the promotion period the approach computes differences "observed minus predicted traffic." These differences measure the effects of interest. One should stress that one may conduct an evaluation at any time during, or after, the intervention. Call t_I the time of the evaluation. Then one uses the data prior to the intervention to make predictions until t_I. Since the predictions are random variables, the resulting measure of performance itself is a random variable, that this article calls π. The distribution of π is useful to compute probabilities that reflect the uncertainty of the MPC. Some of these are the probability of a positive performance and the probability that the performance exceeds a target value a manager specifies prior to the intervention. Plots of this distribution help visualize and communicate the results of the analysis.

A concrete analysis illustrates the approach. The data are the number of customers that arrive at a retail store in one-hour time-bands between September 2, 2011 and November 30, 2011, with a sales promotion campaign taking place between November 6 and November 20. This analysis uses data from Skillup-Chile, a technology company specializing in technologies for video transmission and analytics. The illustration models the counts with a conditional Poisson regression model (Dobson, 2008). Model fitting, analyses and graphs use the R-package (R Core Team, 2013). The model includes covariates for hourly time-bands, day of the week, week of the month, month and an additional function of time to account for secular trends. An autoregressive term models autocorrelation. This illustration provides inferences about the effect on traffic of a promotion that offers a 20% discount to customers who are members of a loyalty program of a major local newspaper.

Research on methods to evaluate promotion effects on traffic with direct connections to this article is scant, but some articles highlight the importance of the method this article proposes. This method is useful to examine the effectiveness in larger shopping environments such as malls. To gauge the effectiveness of sales promotions several key measures have been employed, including store traffic, sales, and profitability (e.g. Doyle and Saunders, 1986). Parsons (2003) indicates that malls use two key performance indicators: sales and visits, precisely the outcome that concerns this paper. Parsons (2003) stresses distinctions between sale drivers and visit drivers and points to possible combinations that would be most profitable. Parsons (2003) further points to research gaps to connect promotional activities with the responses they attempt to induce. More specifically, he stresses that studies on the effectiveness of price-based promotions in shopping malls are scarce. Ando (2008) attempts to address the question of measuring the "effectiveness of marketing activities, the baseline sales, and the effects of controllable/uncontrollable business factors…". Ando (2008) provides the same point-wise baseline, namely fitted values, but models explicitly the effect of "execution of promotion." Leone (1983) focuses on sales response model building. His analyses concentrate on coefficients and discuss distinctions between econometric and time-series approaches. Through multivariate time-series analysis, he explores the competitive environment of an industry where advertising is the main source of competition.

The approach in this article measures the predicted-effect in the scale of the response of interest, that is, number of visits to a store, and in this fashion sidesteps the use of more complex models that incorporate a "marketing activity" covariate. In this approach the prediction distribution provides a baseline to compute effects directly on the response of interest.

2. Methodology

For concreteness, this section describes the method in the context of a specific example but, with adaptations, the method is applicable to a broad class of situations that center on the effects of MPC on a performance measure. In this section the data consist of the hourly counts of people that enter the store (dependent variable) with corresponding values of the following independent variables: time-band, that is one-hour time intervals, day of the week, week within a month, month, and a covariate t that represents consecutive days from the start of the study, $t = 1$. The study period has $T = 90$ calendar days, from September 2, 2011 to November 30, 2011. Fig. 1 shows with a blue line the observed time series of the number of visits in each one-hour time-band.

No data is available for three holidays, September 18 and 19, and October 31, as the store did not open. The data for the first three days of the time series are reliable but incomplete. The reason is that the data collection starts with the installation of the camera, September 2, and works irregularly for three days.

The model uses first-levels of factors for baselines. The first day in the data set, September 2, is a Tuesday. The analysis is easier to conduct if one uses Monday, September 1 as baseline. Only to enable this baseline, the analysis imputes the data for September 1. The imputation of one day of data does not affect the conclusions.

The model uses an auto-regressive Poisson regression model for the time series of counts. The description of the model requires the definitions of the next four factors and one covariate,

1. $B(\cdot)$ the factor representing the time-band that starts at hour of the day $(b-1)$ and finishes at b, $b = 11, 12, \dots, 22$. The level b is in the 1–24 scale. The store opens for business at 10 (10 am) and closes at 22 (10 pm),
2. $D(\cdot)$ the factor with levels $d = m, tu, w, th, f, sa, su$, representing the day of the week,
3. $W(\cdot)$ the factor with levels $w = 1, 2, \dots, 5$, representing the weeks in a month and,
4. $M(\cdot)$ the factor with levels m = sep, oct, nov, representing the corresponding months.
5. Let t be an index for the days with values $t = 1, \dots, T$. with $T = 90$.

Let $Y_{B(b),D(d),W(w),M(m),t}$ be the number of people that enter the store during the band that the covariate t and the combination of factors $B(b), D(d), W(w), M(m)$, define. Kondo and Kitagawa (2000) use similar parametrizations for time series analyses of daily scanner sales.

An initial Poisson regression model with counts $Y_{B(b),D(d),W(w),M(m),t}$ as responses has the following predictors: the covariate t and the factors, $B()$ time-band, $D()$ day of the week, $W()$ week of the month, and $M()$ month. Analyses incorporate the covariate t^2 to account for possible quadratic trends beyond linear trends.

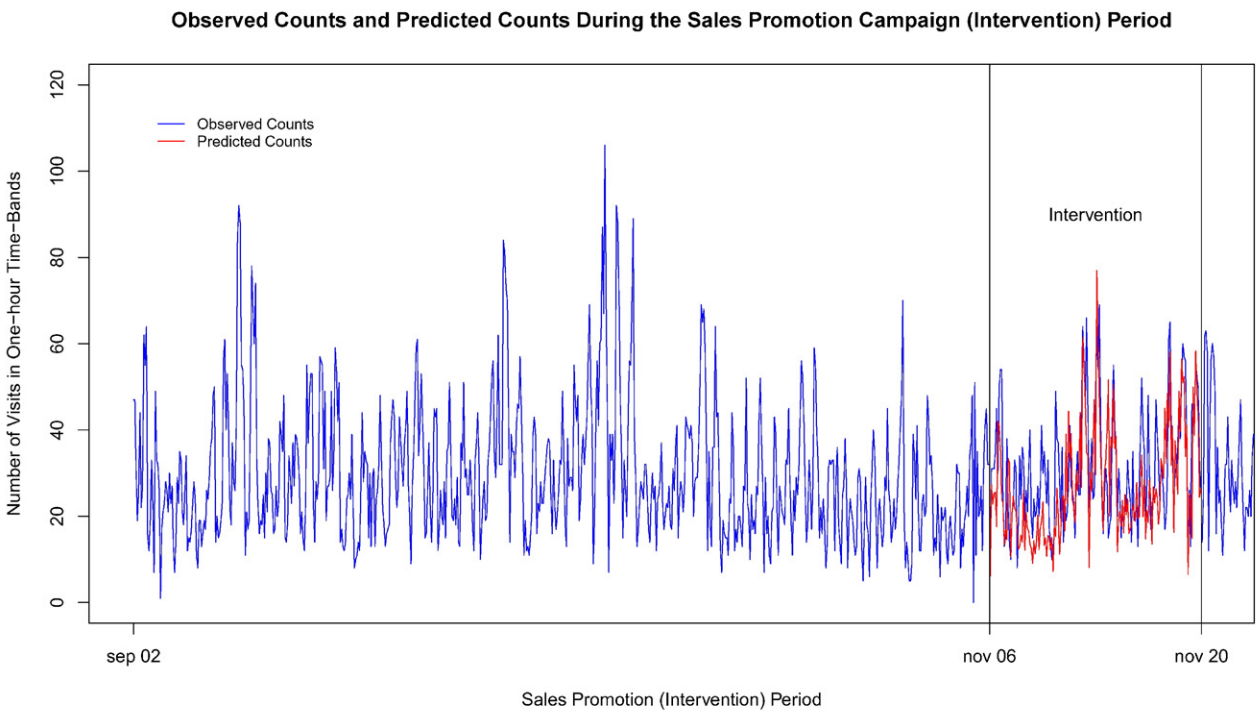

Fig. 1. Observed and predicted counts during the promotion (intervention) period.

A description of the model building process follows. To simplify formulae, the notation omits sub-indexes in the response variable Y and in the Pearson residual r. The model building process starts with a log-linear model with main effects only, $\ln E(Y) = \beta_0 + \beta_{B(\cdot)} + \beta_{D(\cdot)} + \beta_{W(\cdot)} + \beta_{M(\cdot)} + \beta_t t$, and proceeds sequentially with the addition of interactions and evaluations of their significance. This process leads to the model,

$$\ln E(Y) = \beta_0' + \beta_{B(\cdot)} + \beta_{D(\cdot)} + \beta_{W(\cdot)} + \beta_{M(\cdot)} + \beta_{t^2} t^2 + \beta_{D\times M(\cdot,\cdot)} + \beta_{D\times W(\cdot,\cdot)} + \beta_{B\times D(\cdot,\cdot)} + \beta_{D\times t^2(\cdot)} + \beta_{B\times M(\cdot,\cdot)} + \beta_{B\times W(\cdot,\cdot)} + \beta_{W\times t^2(\cdot)} + \beta_{B\times t^2(\cdot)}.$$

Standard diagnostics with Pearson residuals indicate that this linear parameterization for $\ln E(Y)$ is adequate. Usual Poisson regression models are appropriate only for independent observations. In this analysis however, the data set is a time series and therefore correlations are likely among neighboring counts. Diagnostics with Pearson residuals, $r = (Y - Y)/\sqrt{Y}$ where Y is a fitted value, assess the adequacy of a model for independent observations. However, a plot of auto-correlations, Fig. 2, suggests that consecutive observations may be correlated.

For this reason, the analysis modifies the model to include the covariate $\ln Y_{B(b-1),D(d),W(w),M(m),t}$ in the parameterization for $\ln E(Y_{B(b),D(d),W(w),M(m),t})$. This model uses the logarithm for the covariate, $\ln Y_{B(b-1),D(d),W(w),M(m),t}$, to match the scale of the parameterization for $\ln E(Y)$. The next step in the model building process examines the significance of interactions of factors with the lag-covariate $\ln Y_{B(b-1),D(d),W(w),M(m),t}$, or simply $\ln Y_{B(b-1)}$, and retains the final model,

$$\ln E(Y) = \beta_0' + \beta_{B(\cdot)} + \beta_{D(\cdot)} + \beta_{W(\cdot)} + \beta_{M(\cdot)} + \beta_{t^2} t^2 + \beta_{y_{t-1}} \ln y_{t-1} + \beta_{D\times M(\cdot,\cdot)} + \beta_{D\times W(\cdot,\cdot)} + \beta_{B\times D(\cdot,\cdot)} + \beta_{D\times t^2(\cdot)} + \beta_{D\times B(b-1)} + \beta_{B\times M(\cdot,\cdot)} + \beta_{M\times B(b-1)} + \beta_{B\times W(\cdot,\cdot)} + \beta_{W\times t^2(\cdot)} + \beta_{W\times B(b-1)} + \beta_{B\times t^2(\cdot)} + \beta_{B\times B(b-1)}.$$

In this model, the sub-index $B(b-1)$ in a parameter, $\beta_{M\times B(b-1)}$ for instance, indicates that the parameter is the interaction between the factor M and $\ln Y_{B(b-1),D(d),W(w),M(m),t}$. The empirical auto-correlation function of the Pearson residuals from this model does not suggest including higher-order lags into the model. Fig. 3 shows the empirical auto-correlation function of the final model (left panel), and for illustrative purposes, also shows the series of Pearson residuals (right panel).

The final model fits 223 parameters to a time series with 953 observations that exclude the promotion period. That this parameterization does not include parameters that model the effect of the promotion is worth stressing, because the purpose of the model is to predict the number of visits to the store in each one-hour time-band, had the

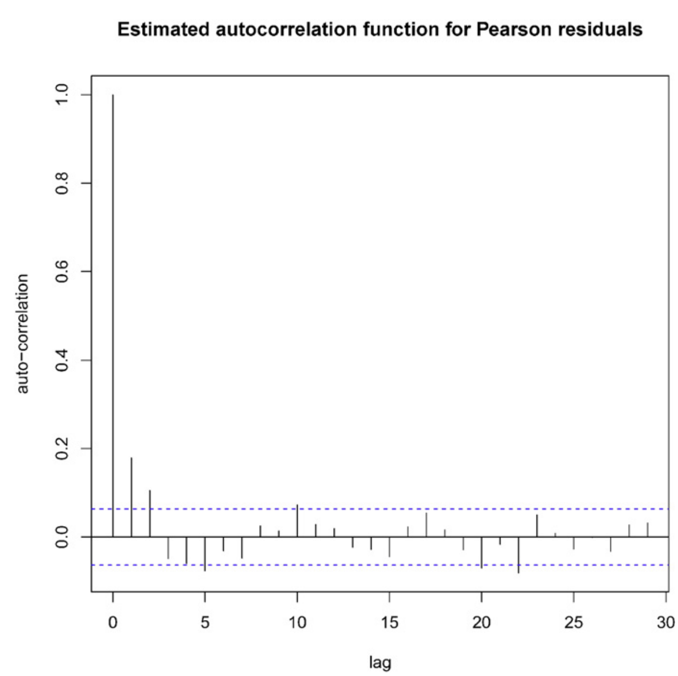

Fig. 2. Estimated autocorrelation function of Pearson residuals.

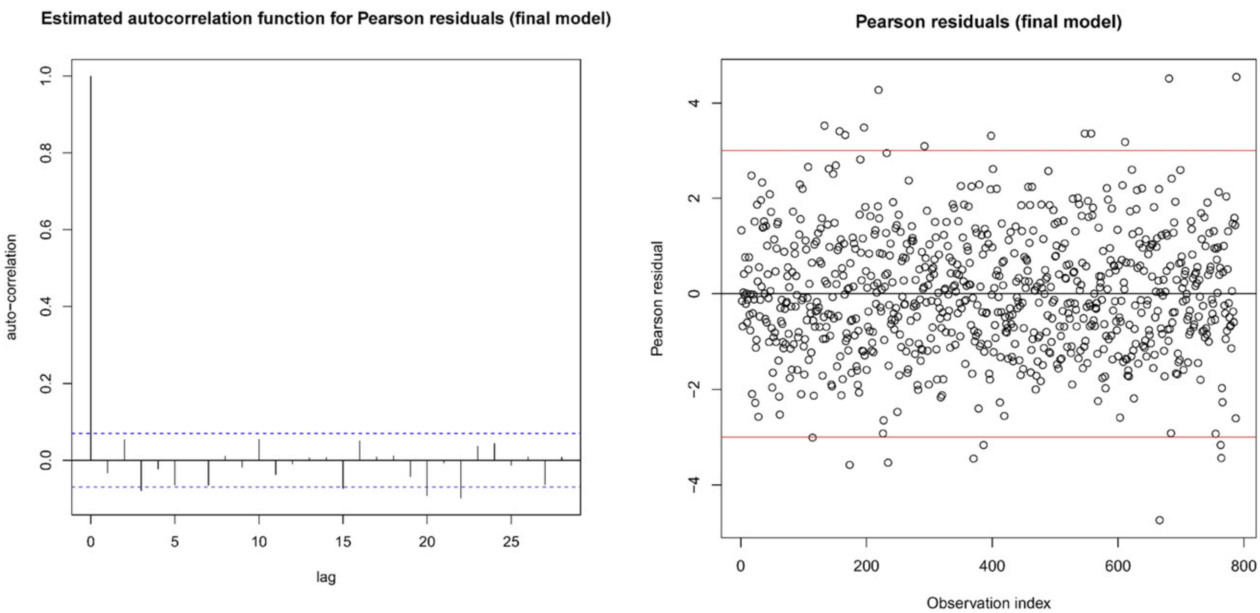

Fig. 3. Left panel: Estimated autocorrelation function for the Pearson residuals (final model). Right panel: Pearson residuals (final model).

promotion not taken place. One may attempt to interpret the final model, namely to examine the estimates of the parameters in the model. However, these interpretations would distract from the illustration of the method, which focuses on the principle that evaluations of MPC should use predicted-effects "observation minus prediction." The approach does not use models that require parameterizing the intervention effect. This is an important advantage of the approach, in part because observations during interventions often show irregular patterns and modeling them could be a challenge difficult to overcome, especially if the intervention takes place during a short time interval, in which case one may be able to collect too few observations during this interval to estimate promotion-specific parameters. This limitation compounds with the complexity of a model that requires additional parameters to incorporate changes in variance and correlations during the promotion. Certainly, estimating parameters that represent intervention effects may be, and have been, the interest of other studies.

3. Analysis and results

Fig. 4 shows in red the predicted number of visits during the promotion. These predictions represent one's expectations about the number of visits in each one-hour time-band had the promotion not taken place. These predicted counts, therefore, provide a baseline to evaluate the effect of the promotion.

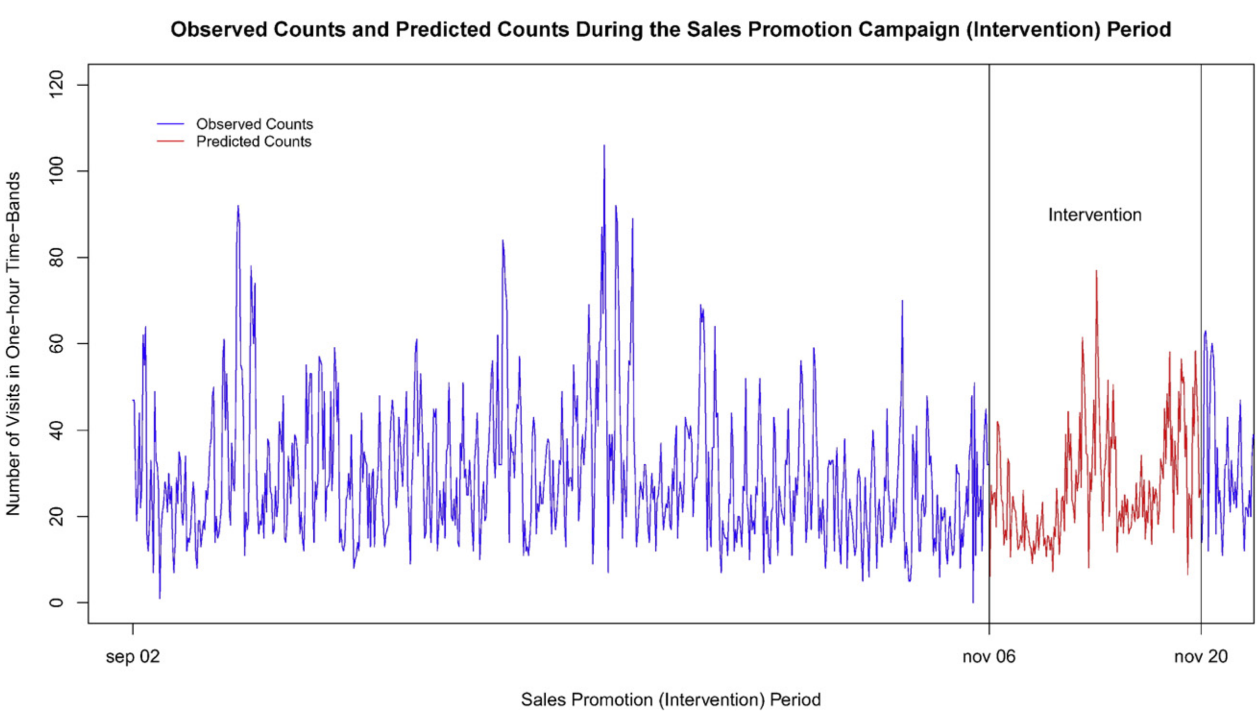

Fig. 4. Observed and predicted counts during the promotion (intervention) period.

During the promotion and for a given time-band, if the number of visits to the store is larger than the baseline count, then one might attribute the excess visits to the promotion. This study calls this difference "time-band specific predicted effect" or "predicted residual." With symbols,

$$E_{B(b),D(d),W(w),M(m),t} = Y_{B(b),D(d),W(w),M(m),t} - \hat{Y}_{B(b),D(d),W(w),M(m),t}.$$

At the end of the study, namely after November 30, 2011, with the data and the predictions available, one may compute the time-band specific observed predicted effect,

$$\hat{e}_{B(b),D(d),W(w),M(m),t} = y_{B(b),D(d),W(w),M(m),t} - \hat{y}_{B(b),D(d),W(w),M(m),t}.$$

Fig. 5 shows the time-band specific counts and the corresponding predicted-effects during the promotion period. Blue dots are the observed counts, red dots are the predicted counts (baseline), green segments represent the positive values of $e_{B(b),D(d),W(w),M(m),t}$, and black segments represent the negative values of $e_{B(b),D(d),W(w),M(m),t}$. A positive value of $e_{B(b),D(d),W(w),M(m),t}$ corresponds to a time-band where the count of visits is larger than the expected count. A parallel interpretation goes with a negative value of $e_{B(b),D(d),W(w),M(m),t}$.

Fig. 6 shows the time-band specific predicted-effects during the promotion. The figure reveals that positive effects are much more frequent than negative effects, suggesting that at the end of the promotion the number of visits is larger than they would have been if the promotion did not occur. That is, the promotion associates with store traffic increase.

One should raise the question as to whether the time-band specific predicted-effects, either positive or negative are the result of the variability of the data and not a result of the promotion. To answer this question, Fig. 7 shows with black dots the time-band specific observed predicted-effects during the promotion and uses red segments to represent the corresponding point-wise 95%-level prediction intervals. Fig. 6 represents these same predicted effects with segments.

Observe that each interval combines the variability of the conditional Poisson counts and the variability of the estimated model. For a specific time-band, one may interpret the size of an observed predicted-effect as follows: if the observed predicted-effect is above the corresponding prediction interval, then such magnitude may not be attributable to the variability of the data. Thus, one could infer that the additional arrivals are attributable to the promotion. In a few cases, approximately six, the observed predicted-effects are below their corresponding prediction interval. With 95% prediction intervals, one would expect approximately 2.5% of the effects to have values below the lower bounds of their corresponding intervals. Thus, during the period November 6–November 20, 15 days with $15 \times 11 = 165$ time-bands, one would expect approximately $165 \times 0.025 \approx 4$ effects to have values below the lower bounds of their respective intervals, similar to the observed number (six) of such effects. Hence, these six predicted effects could be attributable to the variability of the data. On the other hand, one would expect four observed predicted effects to lie above their corresponding intervals. However, more than 40 observed predicted effects are above their intervals. This larger than expected number of effects above their intervals cannot be attributable only to the variability of the data and strongly suggest that the promotion is associated with an increased probability of observing larger than expected arrival counts. A more detailed analysis could use simultaneous prediction bands, but their interpretation would be somewhat more distant from managerial practice.

When a manager plans a sales promotion campaign, she hopes to attain a specified goal. Assessing the attainment of a promotion goal requires effect measures that aggregate time-band effects throughout the campaign. When counting the number of visits to a store, the aggregate assessment is the number of additional visits by the end of the campaign that one may attribute to it. Fig. 8 precisely reflects the evolution of the cumulative effect. Indeed, this figure shows the plot of the cumulative number of additional visits from the start of the campaign to a given time-band, say b_0 in a given campaign day t_0. By the campaign's end, the cumulative effect was approximately 661 additional visits, or a mean 661/15, approximately 44, additional visits per day.

Fig. 8 allows the manager and the analyst to see features that are more difficult to see in previous plots. For instance, the number of additional visits grows at an approximately constant rate after the start of the campaign, Nov. 8–Nov. 11, next follows a period where the number of visits does not seem to grow, November 12–November15. After

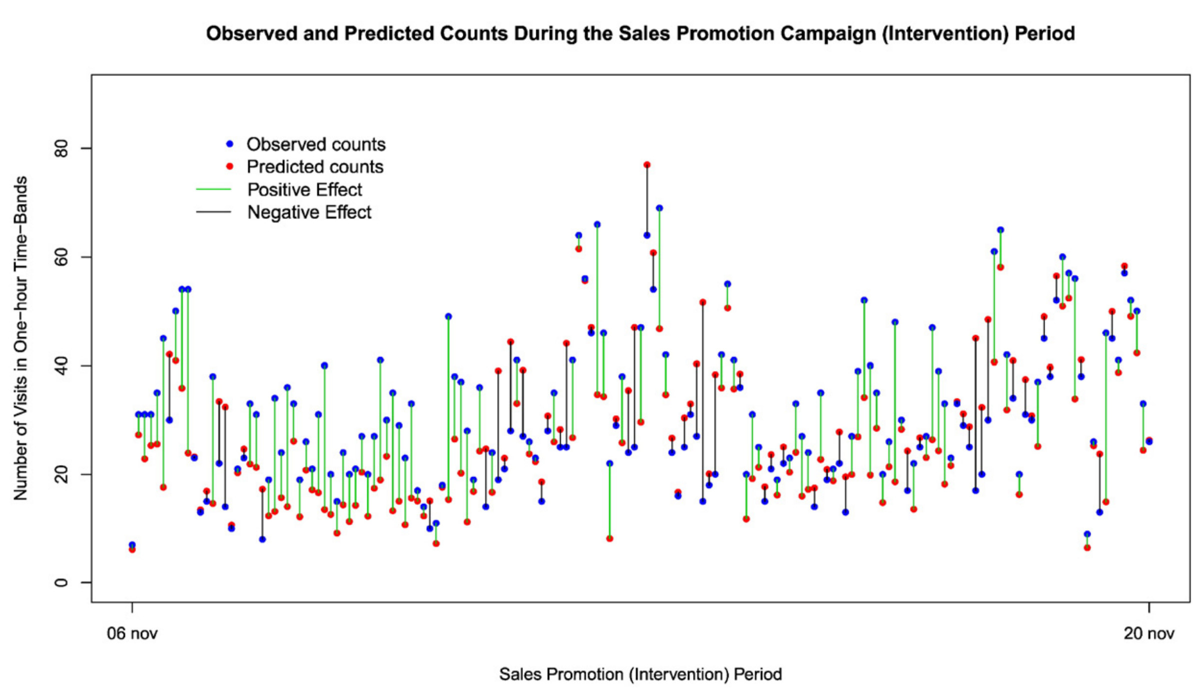

Fig. 5. Observed and predicted counts during the promotion (intervention) period.

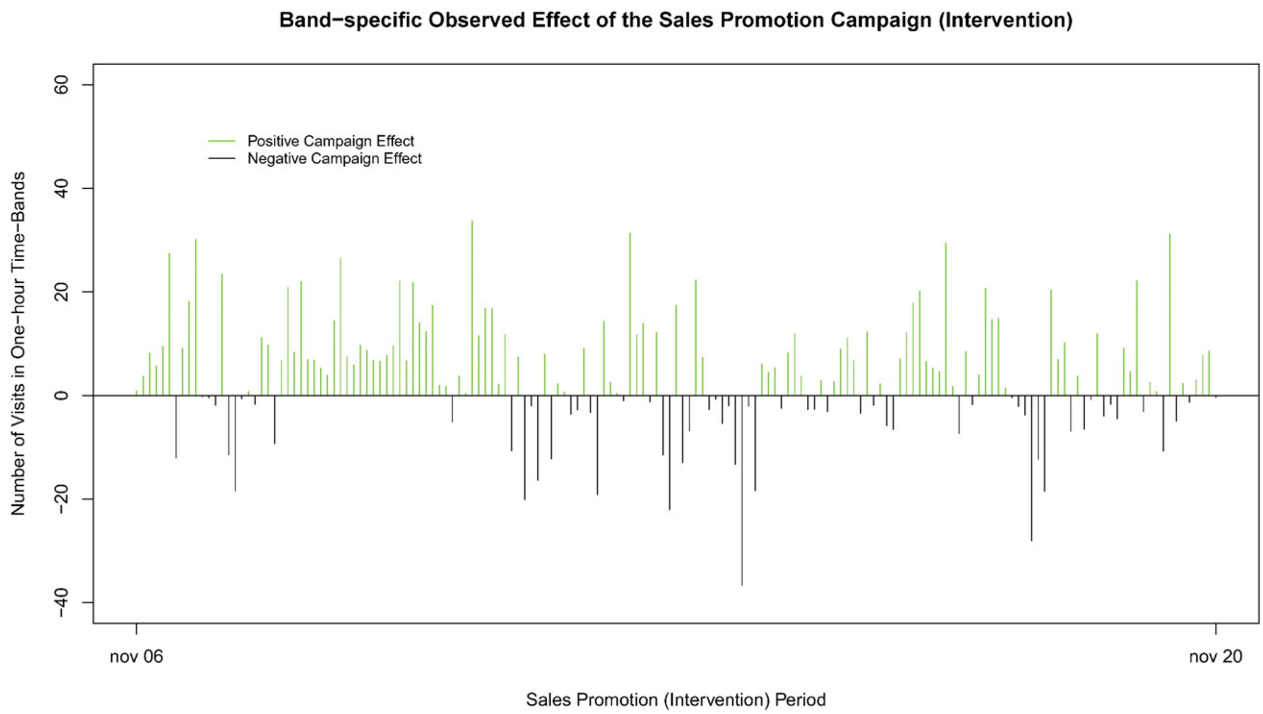

Fig. 6. Band-specific observed effects during the promotion (intervention).

November 15 the growth of visits restarts, but at a rate lower than immediately after the start of the campaign. The leveling off of the effect after the initial growth is a phenomenon worth investigating. Fig. 8, of great value to marketing managers, shows both the evolution of the cumulative effect and the accumulated effect at the end of the campaign. Features of the distribution of the total effect the promotion accumulates deserve closer examination, such as its variability. Indeed, a cumulative effect with a large variability corresponds with a high probability that the cumulative effect is smaller than the estimated 661 additional visits. Assessing the variability of the cumulative effect is a challenge that the approach addresses with the bootstrap method. This method estimates the prediction distribution of the cumulative effect. This prediction distribution provides a much more informative inference than a single measure of variability. A description of the parametric bootstrap scheme appears below (Efron & Tibshirani, 1993).

In the application the bootstrap generates many replicates of time-series of counts with a distribution that approximates the distribution of the actual data. Each one of these series is a "bootstrap replicate". With

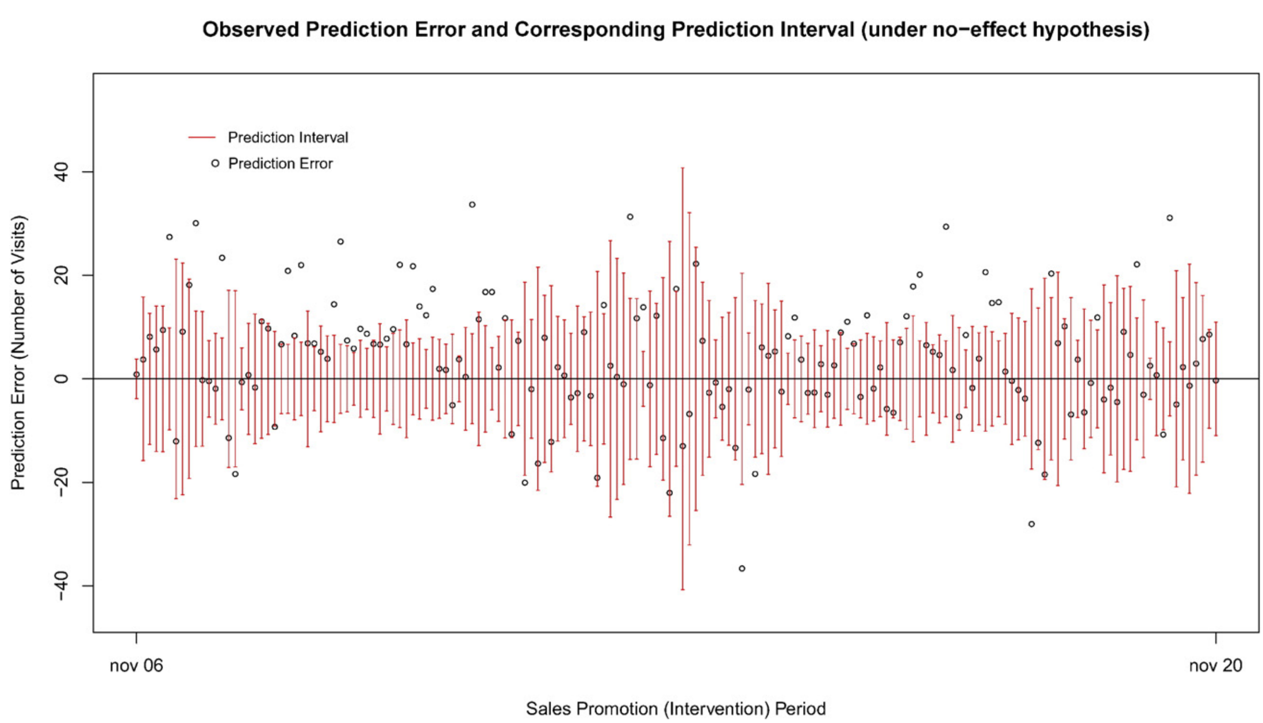

Fig. 7. Observed prediction error and corresponding prediction interval (under no-effect hypothesis).

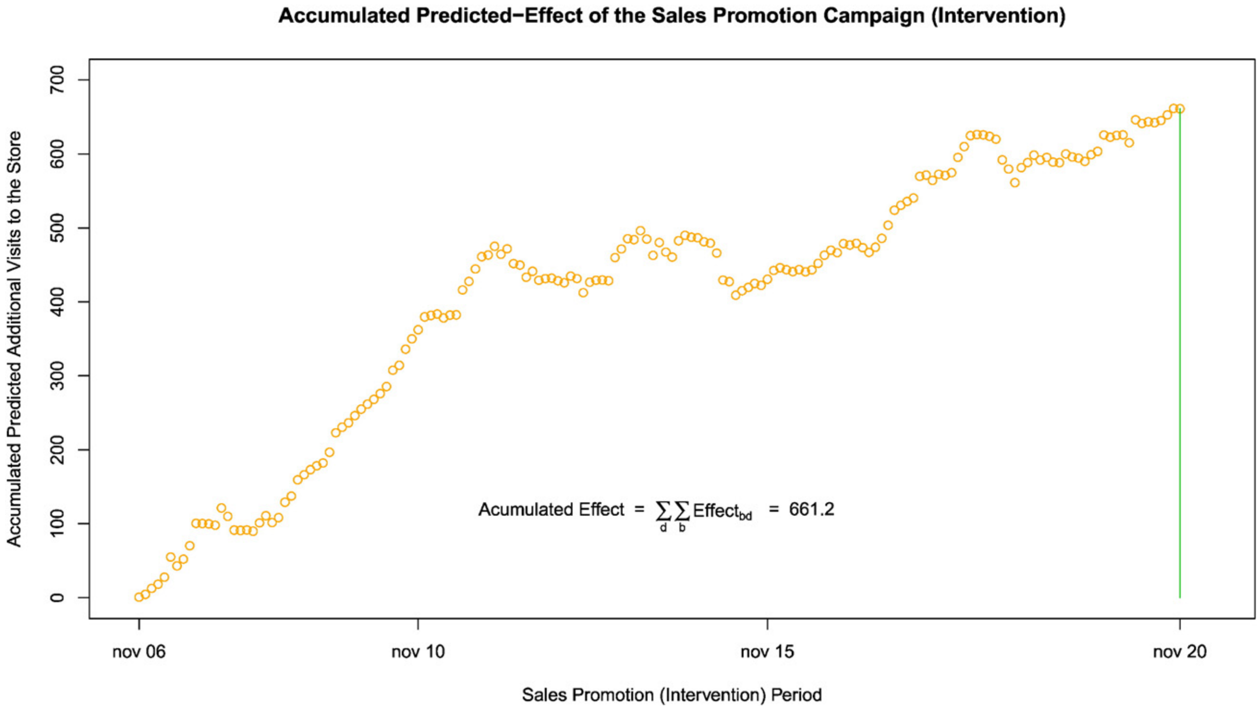

Fig. 8. Accumulated predicted-effect of the promotion (intervention).

each one of these replicates one builds a series of predictions of counts of visits for the promotion period. Each one of these series provides a corresponding baseline with variability consistent with the data outside the promotion period. One uses each one of the baseline series to compute a series of predicted-effects. The collection of predicted-effect series provides estimates of their distribution.

More specifically, the bootstrap scheme is as follows: (1) fit the model to the data, excluding the promotion period; (2) use this model to generate $r_b = 1000$ time series of counts (bootstrap replicates); (3) fit the model to each one of the r_b bootstrap replicates; (4) with each one of the r_b model-fits from step (3) compute a corresponding series of baseline predictions for the number of visits during the promotion period. With each one of the r_b series of baseline predictions, compute the corresponding r_b series of prediction-effects; and finally (5) with the r_b sets of prediction effects, compute the corresponding series of accumulated effects.

Fig. 9, left panel, shows $r_b = 1000$ trajectories, each one of them represents the accumulated effect of a bootstrap replicate of prediction-effects. The purpose of the set of trajectories is to represent the variability of the prediction-effects. The variability in this plot compounds two

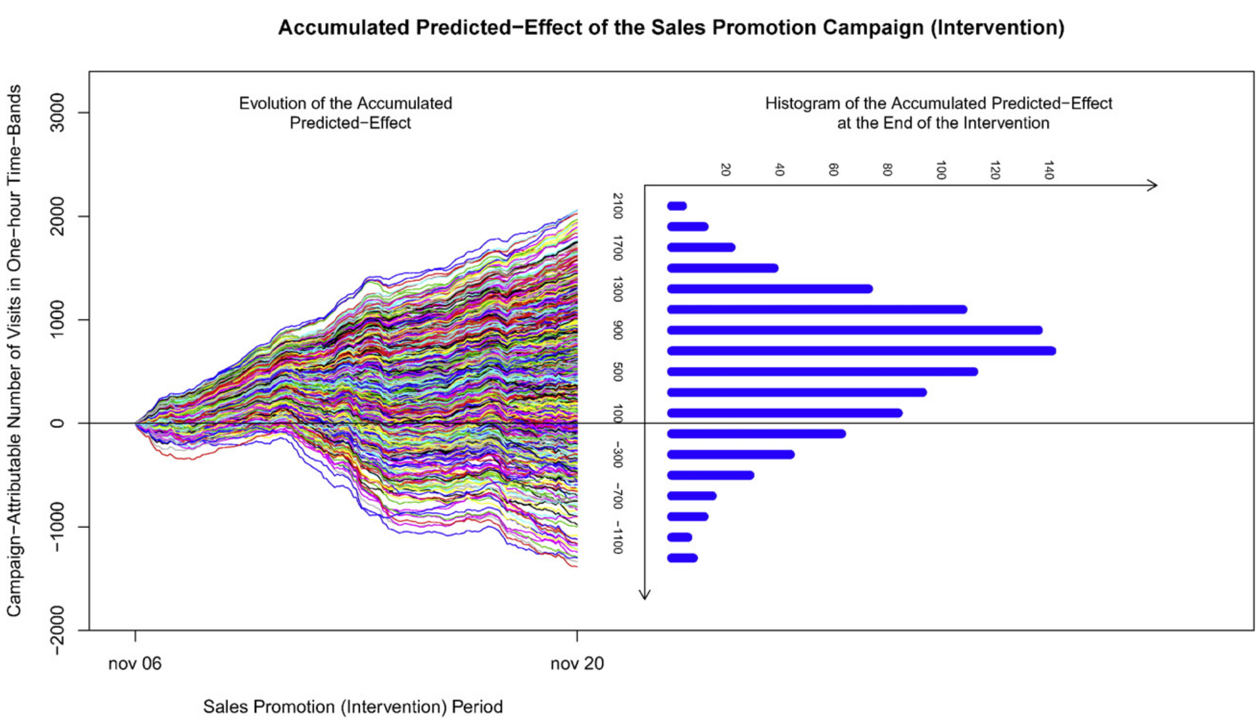

Fig. 9. Accumulated predicted-effect of the promotion (intervention).

sources of variability: the variability of the data (sampling variability) and the uncertainty that corresponds to that prediction. The right panel in Fig. 9 is a histogram of the cumulative predicted-effect at the end of the promotion period.

The prediction distribution, right panel of Fig. 9, communicates the uncertainty about the cumulative outcome of the promotion and complements the plot in Fig. 8, where one sees that the cumulative predicted-effect is 661 visits, a point-wise effect prediction. In Fig. 9, the distribution of the predicted-effect has a mode near 661. Additional analyses illustrate the use of the predictive distribution. One sees that the approximate range of values for the predicted-effect, or additional visits, is the interval $(-1200;2100)$. For a given interval $(a;b)$ in the (vertical) scale of additional visits, the fraction of bootstrap predicted-effects in $(a;b)$ estimates the probability that the predicted-effect takes a value in this interval. Call C_E the cumulative predicted-effect. Thus, $P(C_E<0) = 177/1000 = 0.177$ estimates the probability of a negative predicted-effect. Therefore, under conditions identical to this study, with probability 0.177 the promotion could have a negative effect on the total number of visits. The value $P(550 \leq C_E \leq 750) = 136/1000 = 0.136$, estimates the probability of the predicted-effect taking a value in the interval $(550;750)$. One can use these same ideas to assess the predicted-effect of other measures of performance to evaluate promotions (e.g., sales). The evolution of the cumulative predicted-effect is of interest to managers to monitor the evolution of the campaign effect. One may view the evolution of the cumulative predicted-effect as a stochastic process that can answer questions such as the following: if by day $t_1 = 10$ of the promotion, the cumulative number of additional visits is 300, what is the probability that at the end the promotion the cumulative number of additional visits will be 600? A promotion campaign may have a cumulative negative effect just by chance. If the evolution of a campaign suggests an adverse trajectory, a manager may consider making adjustments. Thus, at that point the manager may want to answer questions such as: if by day $t_1 = 10$ of the promotion, the cumulative number of additional visits is 50, what is the probability that at the end the promotion the cumulative number of additional visits will be only 100?

4. Discussion

4.1. Overview of contribution

This research presents an approach to make inferences about the effect of a sales promotion on visitors' traffic to a retail store. This approach uses the actual measure of interest rather than a surrogate (e.g. Mulhern & Padgett, 1995, Walters & Rinne, 1986). The method is more accurate and provides inferences of the promotion effects that are more thorough than inferences in prior research (e.g., Abraham and Lodish, 1993; Ando, 2008). This method has the additional advantage that provides good measures of the promotion effect even with very short promotion periods. Thus, this research makes an important methodological contribution of practical relevance.

This approach creates opportunities to develop new and useful tools for managers. This approach shifts the focus from parameters in a model to predicted-effects. This shift serves managers and business better than the statistical significance of a parameter, or a parameter estimate. Indeed, what matters to managers are the effects of promotions and campaigns on measures of performance, such as changes in the number of visits to a store or the returns on the investments in promotions. Translating inferences into business actions is more difficult with the conventional emphasis on parameter values. Indeed, in a complex model with many main effects and interactions, to summarize an overall conclusion that combines the effects of the parameters on a single performance measure is difficult in most cases. Further, expanding this approach will allow managers to explore, on the basis of predictive distributions, whether an intervention is more effective under certain conditions, or combinations of conditions, namely, specific days of the week, or specific weeks of the month, or specific months, or holidays, etc.

4.2. Research extensions

One may expand the approach in various ways that are both relevant for academics and managers alike. Some of these expansions depend on data collection abilities soon to become available. One of these expansions first identifies a person looking at the store window, then determines whether that person enters the store, and then follows her and registers her behavior, including her purchase decisions. The exact time of the purchase allows analysts to match the purchase with cash-register data to determine the amount of the purchase. Certain businesses value predictions of the number of customers that arrive at specific time intervals. This information is often useful to plan the size of the sales force.

Capturing customer-specific data already brings interesting modeling challenges. These data may guide managers to establish product development or service strategies that fit customers. Finally, real-time data collection technologies may allow managers to conduct real-time marketing field experiments at the points of sale. With appropriate statistical control variables one may compare the performance of stores a manager assigns to an intervention with the performance of stores she designates as controls. Statistical planning of these experiments, that blend designed and observational features, may become a fertile ground for development.

Adaptations of the ideas put forward in this article can be useful to analyze the effectiveness of interventions with other performance measures such as return on investment. Real time data collection is just beginning to answer a number of open questions. Some of these questions relate to the persistence of effects of promotions or other actions after the action and for how long. Other questions relate to occasional shift of the timing of purchases associated with the promotions, sometimes delaying the purchases until the promotion starts and sometimes accelerating the purchases into promotion period. Assessing the effect of a promotion is more difficult when the promotion shifts the timing of purchases. The refined data that electronic imaging technologies gather may prove useful to evaluate whether promotion shifts purchases, and if so to what extent.

Additionally, expanding this methodology, retailers will be able to compare of effectiveness of different promotions that take place at the same time in the same store, or that take place at the same time but in different stores. The development of models will be able to take into account store-to-store differences as well. Further, the method allows for the study of interactions among promotions within and among stores. Finally, using the accumulation of data from past actions can contribute to planning and assessing the effectiveness of future promotion campaigns.

Acknowledgments

The authors are grateful to Skillup-Chile for providing the data and for helpful discussions. Milberg gratefully acknowledges partial support from the Research Center for International Competitiveness UAI, SOC 1105 (CONICYT, Chile).

References

Abraham, M. M., & Lodish, L. M. (1993). The Integrated System for Improving Promotion Productivity Using Store Scanner Data. *Marketing Science*, *12*(3), 248–269.

Ando, T. (2008). Measuring the baseline sales and the promotion effect for incense products: A Bayesian state-space modeling approach. *Annals of the Institute of Statistical Mathematics*, *60*, 763–780.

Dobson, A. J. (2008). *An introduction to generalized linear models* (3rd ed.). Chapman & Hall/CRC.

Doyle, P., & Saunders, J. (1986). Measuring the true profitability of sales promotions. *The Journal of the Operational Research Society*, *37*(10), 955–965.

Efron, B., & Tibshirani, R. J. (1993). An introduction to the bootstrap. *Chapman & Hall/CRC monographs on statistics & applied probability*.

Kondo, F. N., & Kitagawa, G. (2000). Time series analysis of daily scanner sales: Extraction of trend, day-of-the-week effect and price promotion effect. *Marketing Intelligence & Planning*, *18*(2), 53–66.

Lam, S. Y., Vandenbosch, M., Hulland, J., & Pearce, M. (2001). Evaluating promotions in shopping environments: Decomposing sales response into attraction, conversion, and spending effects. *Marketing Science*, *20*(2), 194–215.

Leone, R. P. (1983). Modeling sales-advertising relationships: An integrated time series-econometric approach. *Journal of Marketing Research*, *20*(3), 291–295.

Mulhern, F. J., & Padgett, D. T. (1995). The relationship between retail price promotions and regular price purchases. *Journal of Marketing*, *59*, 83–90 (October).

Parsons, A. J. (2003). Assessing the effectiveness of shopping mall promotions: customer analysis. *International Journal of Retail & Distribution Management*, *31*(2), 74–79.

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing (ISBN 3-900051-07-0, http://www.R-project.org.).

Walters, R. G., & Rinne, H. J. (1986). An empirical investigation into the impact of price promotions on retail store performance. *Journal of Retailing*, *62*(3), 237–266.