**Supplementary Information File**

**Title:** Estimatingthe lagged effect of price discounting: a time-series study using transaction data of sugar sweetened beverages.

**Authors:**

Hiroshi Mamiya, Alexandra M. Schmidt, Erica EM. Moodie, David L. Buckeridge

Contents

[Appendix S1. Detailed Description of Calculation of Category-Level Discounting 2](#_Toc83553755)

[Supplementary Figure S1 3](#_Toc83553756)

[Appendix S3. Shapes of promotional lag 5](#_Toc83553757)

[Supplementary Figure S2: 5](#_Toc83553758)

[Appendix S3. Statistical analysis: time-series regression to incorporate the lagged effect of discounting, covariates, and intercept 6](#_Toc83553759)

[Appendix S4 9](#_Toc83553760)

[Supplementary Figure S3. 9](#_Toc83553761)

[Supplementary Figure S4 10](#_Toc83553762)

[Supplementary Figure S5 11](#_Toc83553763)

[Supplementary Figure S6 12](#_Toc83553764)

[Supplementary Figure S7 13](#_Toc83553765)

[Supplementary Table 2 14](#_Toc83553766)

[References for Supplementary Information File 14](#_Toc83553767)

# Appendix S1. Detailed Description of Calculation of Category-Level Discounting

As in previous studies, we used a conventional marketing approach to calculate the category-level discounting of Sugar Sweetened Beverages (SSB) for each store-week from store-level scanner data (Raju, 1992).

The terms used in the analysis are defined as:

1) *Item* refers to a unique beverage product as identified by Universal Product Code in transaction data. A large supermarket often carries more than 1000 distinct SSB items.

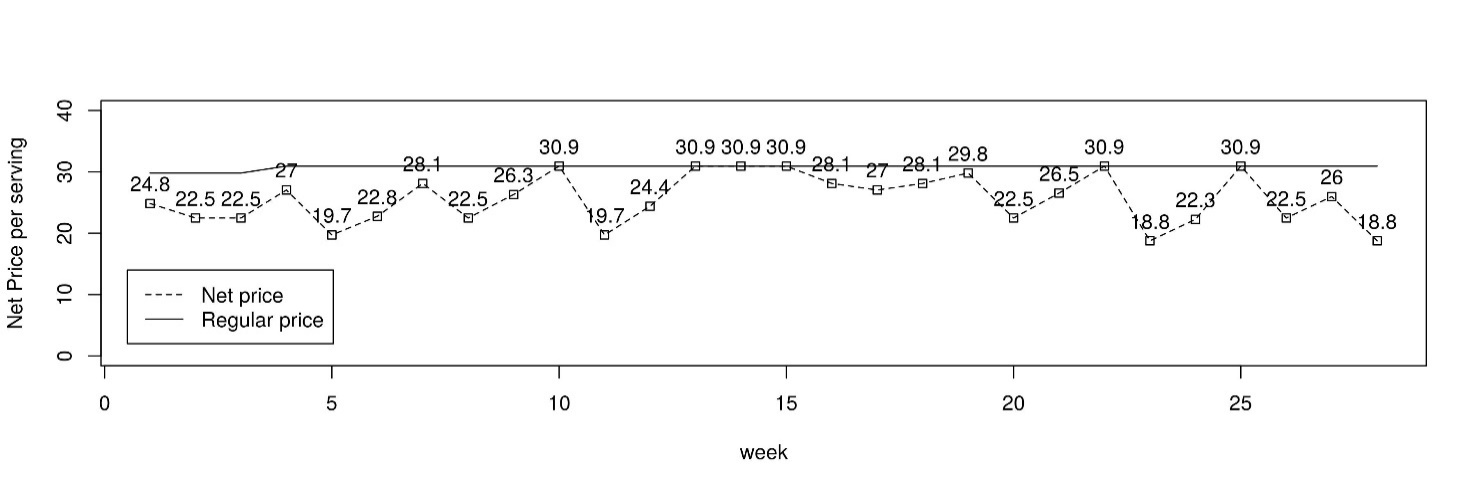
2) *Net price* indicates the price of a SSB item after discounting, standardized to a single serving size (240ml).

3) *Regular price* is the baseline (reference) price of a single serving without any discounting, and

4) *Item* *discounting* indicates the difference between the net and regular price, thus the extent or depth of temporary price reduction per serving. Note that net and regular price are in Canadian Cents.

Because the data provide only the weekly net price after discounting for each item but not the regular price nor discounting directly, we calculated the regular price for each item, and then we computed discounting as the difference between the calculated regular price and the observed net price. We then computed the category-level discounting as the weighted mean of item discounting across items in the same store-week, with the weights representing the market share of each item.

For simplicity, we focus on the calculation of discounting for the soda category as an example. However, the calculations below were applied to each of the five SSB categories. Let represent an item of soda sold in store at week . We denote the net price of item in store at week as . We examined the price history of each item. The regular price of soda item is identified as the highest price in a 3-month window, e.g., a moving maximum net price of item in store in a 3 months widow. Supplementary Figure S4-1 below illustrates a series of weekly net price of one of popular beverage items and its regular price.



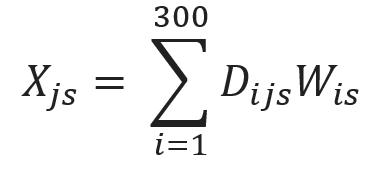
Supplementary Figure S1. Weekly Net Price and Regular Price of a Single Soda Item in Cents.

Item discounting, , is defined as a percent price decrease from regular price as shown below;

= – /

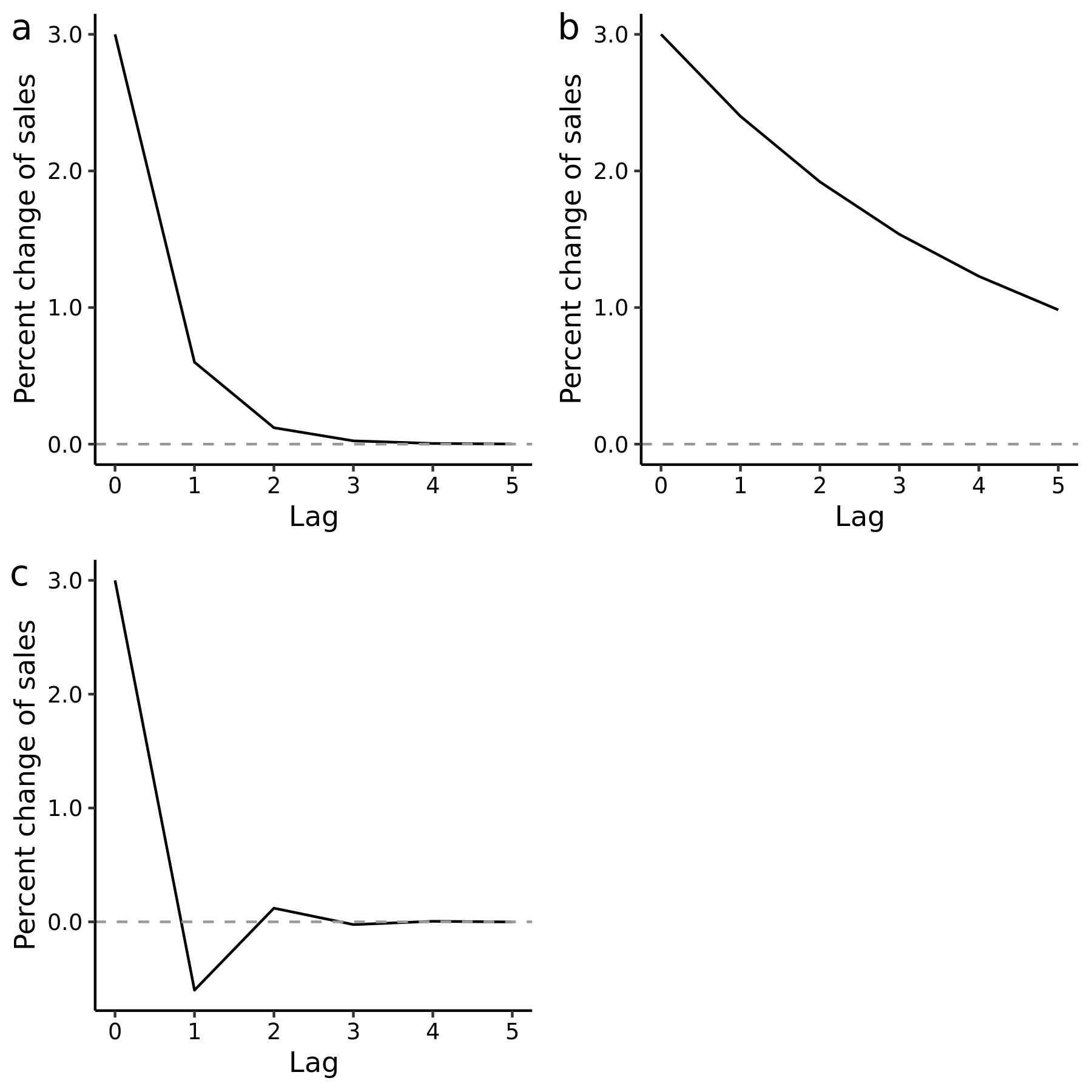
Supplementary Figure S1 indicates that there is no item discounting (zero percent item discounting) in weeks 10, 13-15, 23, and 25, so the net price is the same as the regular price for these weeks. In other weeks, there is discounting since the net price is less than the regular price. In week 11, for example, the discounting is computed as: (30.9 – 19.7) / 30.9 x 100 = 36.25 percent.

Finally, we aggregated the individual item discounting into an overall measure for the discounting in the each of five SSB categories. Suppose there are 300 soda items sold in store in week . Our main exposure of interest, the category-specific store-week discounting of all soda items, ,is defined as

,

where represents a normalized (sum to one across all items in store ) weight calculated from the annual market share of each soda items within the soda category at store . Therefore, the value of the weights is large and stable over time for soda items of major brands.

# Appendix S3. Shapes of promotional lag



Supplementary Figure S2: Hypothetical shapes of the lagged effect of price discounting as a) a monotonic decay returning to the original (baseline, pre-promotion) level of sales with a weak estimated lag effect (, and b) a more persistent lag ( such that sales have not decayed to pre-discounting levels after 5 weeks.

The functions seen in plot a) and b) are generated when the value of is constrained to be between 0 and 1 by an analyst.

Discounting in this plot represents a ‘shock’ exposure, that is, its value is increased by one unit (i.e., one percent discounting) at the week where *x* = 0 (promotion period), then set to zero, or returning to the baseline, pre-discounting, price in the following weeks.

The y-axis represents the effect of one percent discounting on sales at each week. The value at *x* = 0 represents the immediate effect, which is the change of sales during the week of discounting, as represented by the value of .

# Appendix S3. Statistical analysis: time-series regression to incorporate the lagged effect of discounting, covariates, and intercept

*Transfer function*

First, we specify the koyck lag specification discussed in the main text. The monotonic decay of discounting at time *t* is summarized by a structural variable as:

,

where represents price discounting at time *t*, is the coefficient for the immediate effect of discounting, is the decay parameter as described in the main document, and *h* indicates lag length. The equation above allows the effect of discounting to geometrically decay over lag. Using transfer function, the same equation can be recursively summarized as:

, which indicates that the structural variable captures immediate effect of and its cumulative effect up to time *t* through lagged variable (1). We have provided further details about transfer functions (2).

*Time-series model*

We added the structural variable to a dynamic liner model, a time-series linear regression model that naturally incorporates time-varying parameters, crucially above, to smoothly change over time under structural formulation (1,3). In our analysis, the model is specified as:

,

where is the structural variable capturing time-varying influence of discounting combining the immediate and lagged effects up to time *t* as descried in the main text. The outcome is the natural log-transformed sales outcome of each SSB category denoted as .

The intercept is a ‘dynamic’, or time-varying intercept, which captures long-term trends and temporal fluctuations of sales. This is a series of s (one for each week, and thus denoted as ), interpreted as baseline sales at time *t* and is determined by its previous value, and random noise, hence this process is called random walk. Further details in the “local level model” specification are available in references (1,3).

The term represents a vector of covariates described in the main document, the corresponding vector of time-fixed coefficients is denoted as . The term consists of nonlinear regressors as defined by sine and cosine functions fitted to the harmonic wave, with its regression coefficients controlling amplitude of the seasonal wave. Its periodicity is , where as the indicator of week, as 52.2 representing the cycle (number of weeks in a year) and = as the constant pi. With the regression coefficients, and defining the estimated amplitude (vertical extent).

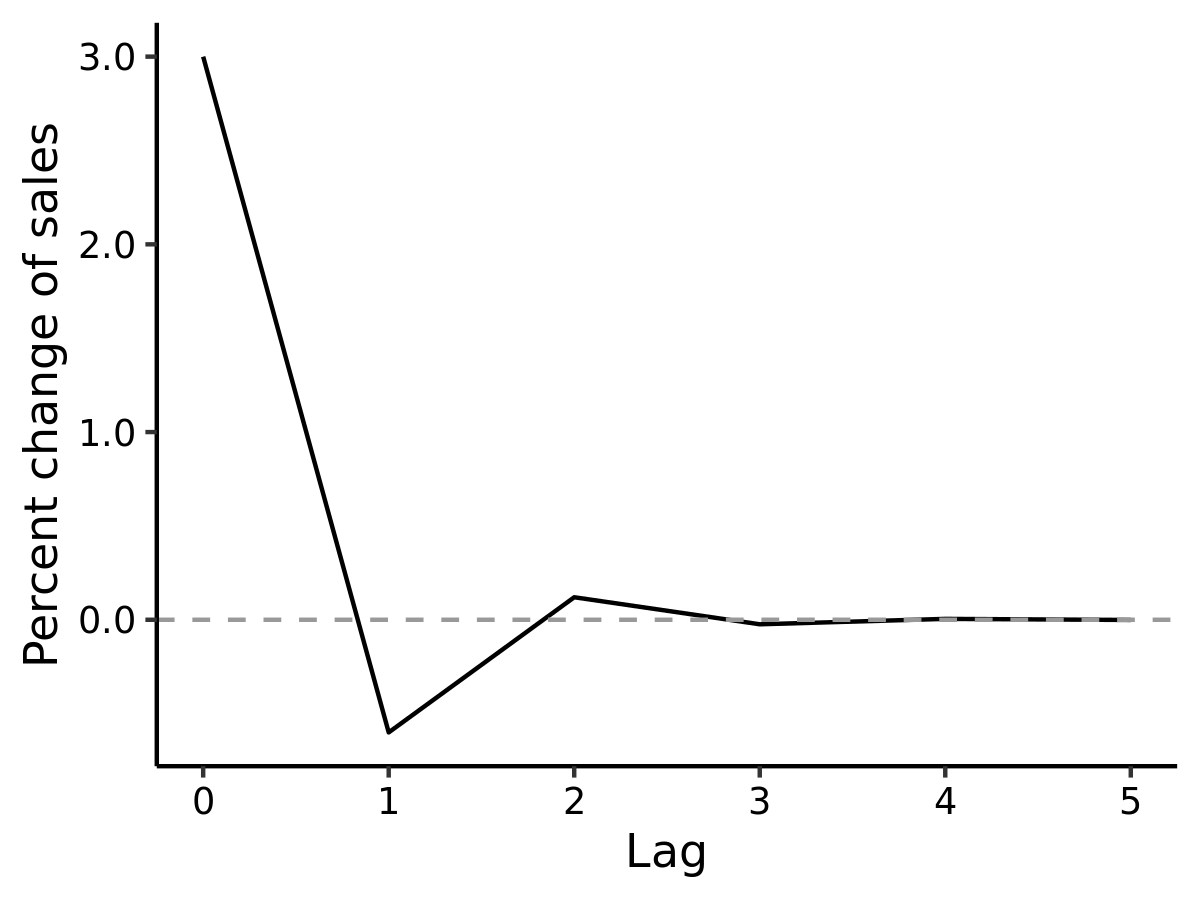
We used Stan statistical software to fit the models (4). The value of the discounting variable as well as all covariates except dummy indicators for holiday were mean centered to improve the mixing of the Markov chain for successful mixing. We generated 3,000 Markov chains as burn-in sample, followed by 30,000 chains as the posterior distribution for inferencing. Stan codes to the model are available in:

https://github.com/hiroshimamiya/promotionLag/blob/main/discountLag\_KoyckTransfer.stan.

*Prior probabilities*

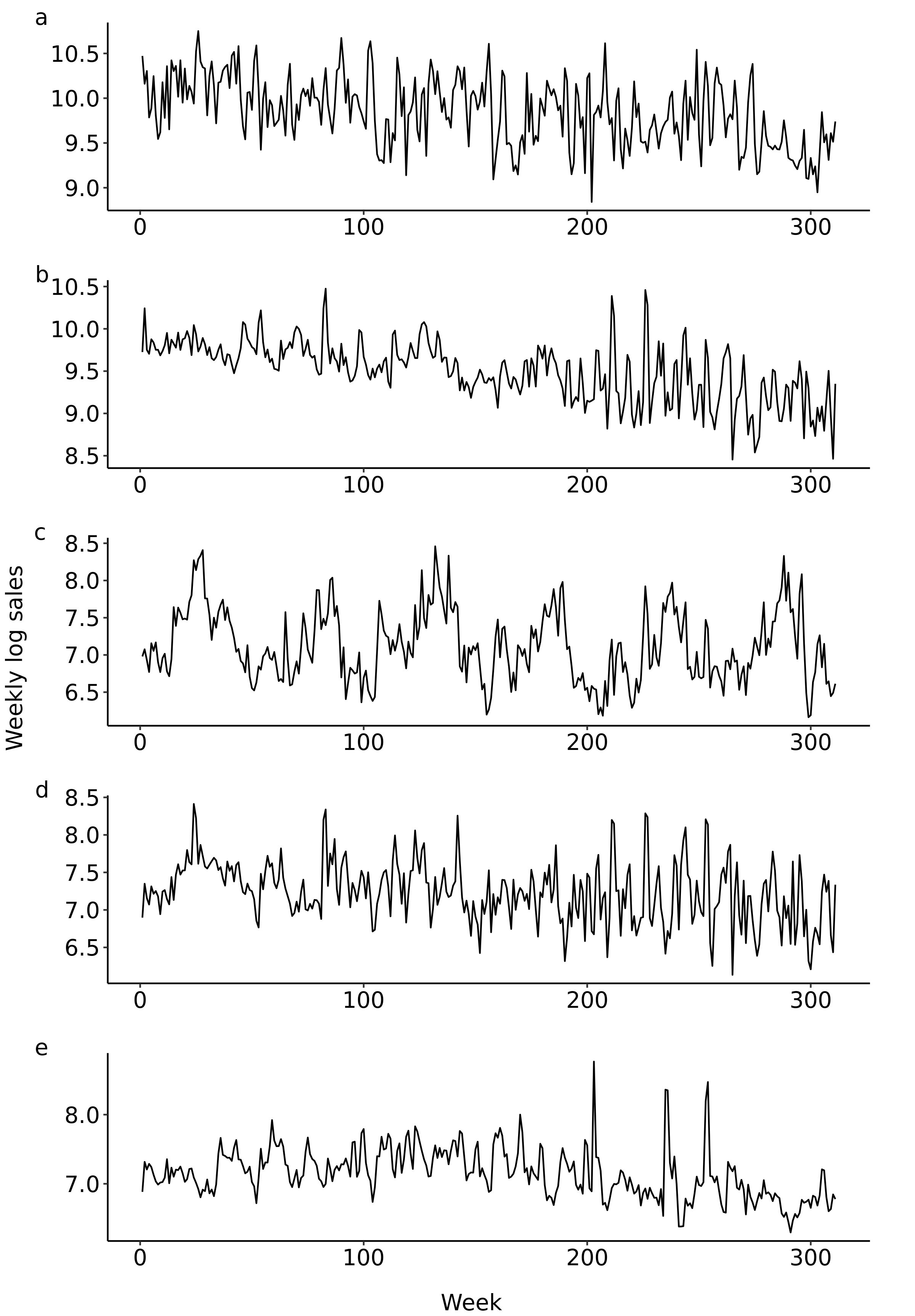
The natural log-transformed sales outcome follows a normal distribution, with its error terms following a zero-mean non-informative (i.e., diffuse) Cauchy distribution , which is a commonly suggested distribution for a variance parameter (5). The superscript + indicates that the distribution is constrained to take a positive value (i.e., constrained to the positive half of the Cauchy distribution). We also used a positive-constrained normal distribution to examine posterior sensitivity with the same values of the scaling parameter as above, which led to nearly identical results. The prior probability of the regression coefficients, including the season coefficients was specified as independent non-informative normal distributions . For the time-varying intercept, , the prior probability of the initial value at was , which was subsequently allowed to evolve as was with a random noise .

# Appendix S4

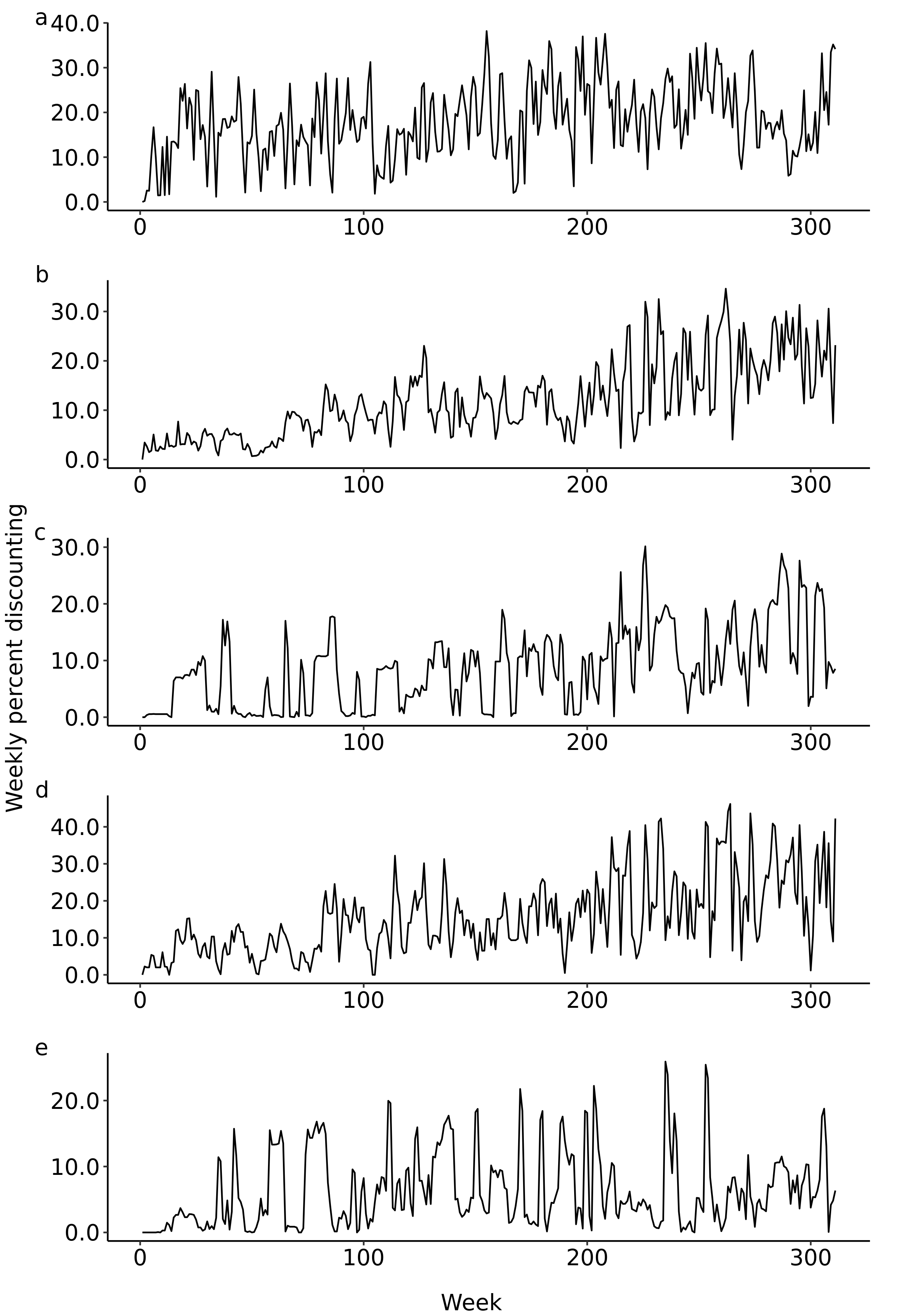


Supplementary Figure S3.Post-promotion dip of sales at one week after the discounting period (*x* = 0).

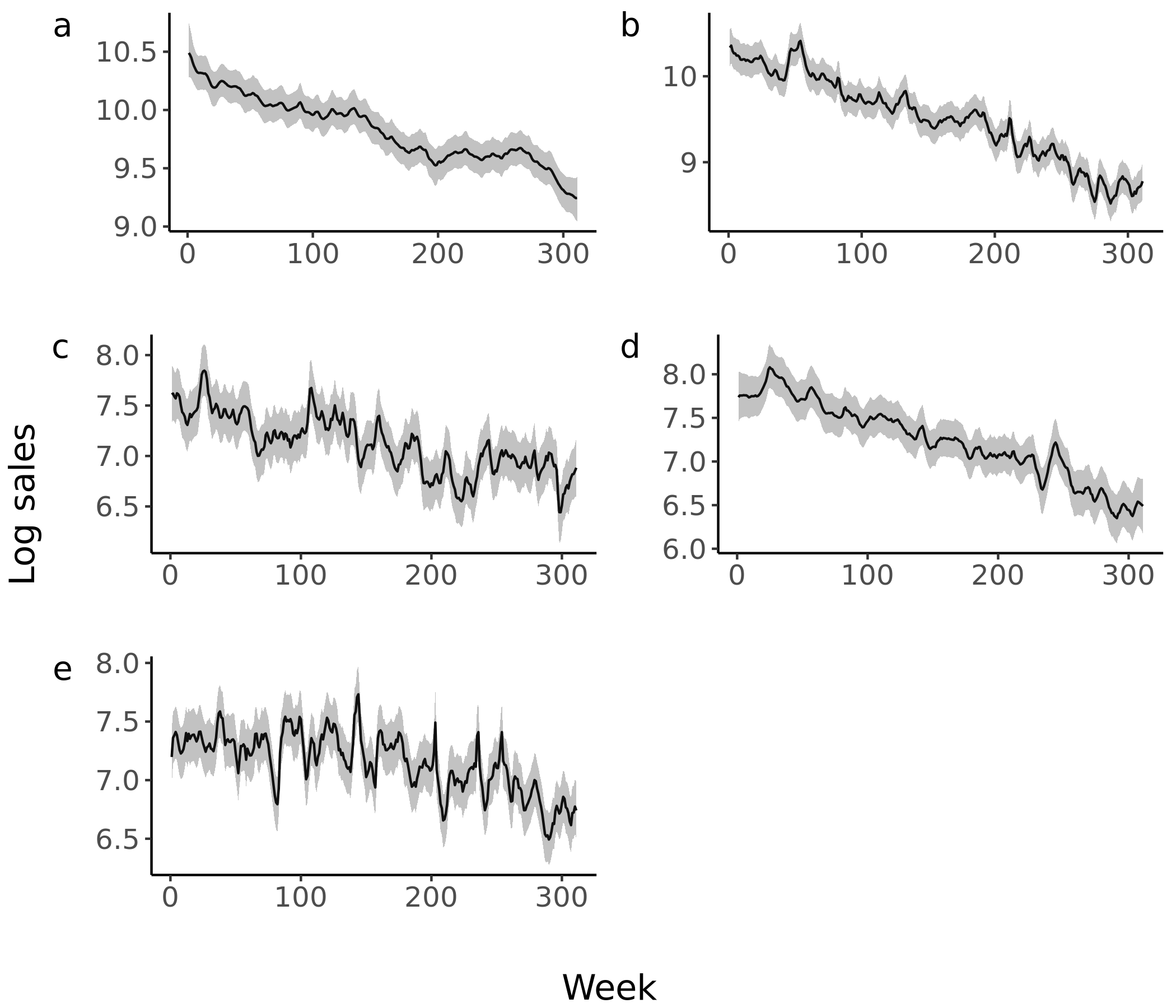
The y-axis represents the percent change of sales relative to the baseline (pre-discounting) sales due to one percent discounting. Unlike the monotonic decay seen in Supplementary Figure S2 a and b, the oscillating pattern described here is captured by constraining the value of as: . The decay function in this plot is generated by The dip is attributed to reduced purchasing activities if households stockpile the promoted items during the discounting period. However, the dip may not be observed when stockpiled items are consumed rapidly, leading to an immediate re-purchasing after discounting (6–8).



Supplementary Figure S4: Weekly series of category-level percent sales for a) soda, b) fruit drinks, c) sports and energy drinks, d) coffees and teas, and e) drinkable yogurt.



Supplementary Figure S5: Weekly series of category-level percent discounting for a) soda, b) fruit drinks, c) sports and energy drinks, d) coffees and teas, and e) drinkable yogurt.



Supplementary Figure S6. Posterior mean (solid line) and 95% credible interval (grey band) of weekly-varying intercept, , of log sales for a) soda, b) fruit drinks, c) sports and energy drinks, d) coffees and teas, and e) drinkable yogurts. Note that the value of the discounting variable was mean centered; therefore, the estimated level of the intercept represents the sales at the mean level of discounting.

A screenshot of a computer

Description automatically generated with medium confidence

Supplementary Figure S7**.** Autocorrelation function of residuals for : a) soda, b) fruit drinks, c) sports and energy drinks, d) coffees and teas, and e) drinkable yogurt category. The grey band around solid line indicates 95% credible interval. It should be noted that the fitted sales in these plots do not represent forecasting distribution.

Supplementary Table 2. Non-log serving-standardized sales quantity of SSB categories attributed to lagged effect along (left pane) and both lagged and immediate effect (right pane).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Sales due to lagged effect | | |  | Sales due to lagged and immediate effect | | |
| SSB category | Mean | Median | 95%CI |  | Mean | Median | 95%CI |
| Soda | 306,483 | 238,042 | (9,943- 977,152) |  | 4,308,369 | 4,283,077 | (3,418,117 - 5,343,486) |
|  |  |  |  |  |  |  |  |
| Fruits drinks | 195,659 | 174,491 | (10,659 - 508,513) |  | 2,317,649 | 2,306,233 | (1,876,257 - 2,824,259) |
|  |  |  |  |  |  |  |  |
| Energy and sports drinks | 28,459 | 26,345 | (2,661 - 67,253) |  | 131,606 | 130,446 | (96,155 - 173,625) |
|  |  |  |  |  |  |  |  |
| Sweetened coffees and teas | 20,079 | 16,627 | (802 - 58,650 |  | 281,883 | 279,539 | (216,303 - 360,958) |
|  |  |  |  |  |  |  |  |
| Sweetened drinkable yogurts | 6,007 | 4,870 | (213- 18,060 |  | 133,117 | 132,531 | (108,460 - 161,073) |

# References for Supplementary Information File

1. West M, Harrison J. Bayesian forecasting and dynamic models. Second edition. New York: Springer; 1997. (Springer series in statistics).

2. Ravines RR, Schmidt AM, Migon HS. Revisiting distributed lag models through a Bayesian perspective. Applied Stochastic Models in Business and Industry. 2006;22(2):193–210.

3. Petris G, Petrone S, Campagnoli P. Dynamic linear models with R. New York: Springer-Verlag; 2009. (Use R!).

4. Carpenter B, Gelman A, Hoffman MD, Lee D, Ben Goodrich, Betancourt M, et al. Stan: a probabilistic programming language. Journal of Statistical Software. 2017 Jan 11;76(1):1–32.

5. Gelman A. Prior distributions for variance parameters in hierarchical models. Bayesian Analysis. 2006;1(3):515–33.

6. Hendel I, Nevo A. The Post-Promotion Dip Puzzle: What do the Data Have to Say? Quantitative Marketing and Economics. 2003 Dec 1;1(4):409–24.

7. van Heerde HJ, Neslin SA. Sales promotion models. In: Wierenga B, van der Lans R, editors. Handbook of marketing decision models. Springer International Publishing; 2017. p. 13–77. (International Series in Operations Research & Management Science).

8. Neslin SA, Schneider Stone LG. Consumer inventory sensitivity and the postpromotion dip. Market Lett. 1996 Jan 1;7(1):77–94.