



Marketing response and temporal aggregation

Philip Hans Franses¹

Revised: 28 September 2020 / Accepted: 31 December 2020
© The Author(s) 2021

Abstract

This paper deals with inferring key parameters on marketing response at a true high frequency while data are partly or fully available only at a lower frequency aggregate levels. The familiar Koyck model turns out to be very useful for this purpose. Assuming this model for the high-frequency data makes it possible to infer the high-frequency parameters from modified Koyck type models when lower frequency data are available. This means that inference using the Koyck model is robust to temporal aggregation.

Keywords Advertising response · Current effect · Carryover effect · Total effect · Temporal aggregation

Introduction

This paper makes the case that the familiar Koyck (1954) or geometric lag model, which is often used to model marketing response¹ to for example advertising, is robust to temporal aggregation. The assumption is that there is an adequate Koyck model for high-frequency data. The focus is on four cases, and these are the case (a) of high frequency, say, sales, and advertising, the case (b) of high-frequency sales and aggregated advertising, the case (c) of aggregated sales and higher frequency advertising and the well-known case² (d) of both low-frequency sales and low-frequency advertising. The two cases (b) and (c) which involve one of the two variables being aggregated have not been studied before.

Temporal aggregation is relevant for the analysis of marketing response, in particular when examining carry-over effects, see also Leone (1995). When the true marketing response process occurs at a high-frequency level, of say hours, and the data are only available at a lower frequency level, say days, the estimation results from models for lower frequency data cannot be one-to-one translated to the true higher frequency process. This point was already made³ in the seminal paper by Clarke (1976), who argued that the model for aggregated data must differ from the model for the high-frequency data.

A key issue for temporal aggregation is that aggregation makes the model to change. That is, if one fits an econometric time series model, like the Koyck model, to sales and advertising data, then the model changes due to aggregation.⁴ Clarke (1976) was the first to recognize that not modifying the model leads to biased results and incorrect advertising duration intervals. In the present paper, it is demonstrated that the Koyck model is very useful even when the data are all or partly available in temporally aggregated format. In fact, it is shown that the Koyck model is robust to such aggregation. The focus is on just two variables for

¹ Nowadays, the model is also used to estimate such effects for various marketing variables like satisfaction, quality, distribution, and online chatter on a variety of dependent variables like sales, market shares, and even earning and stock market returns. Examples of studies using versions of the Koyck model are Berkowitz et al. (2001), Breuer et al. (2011), Chessa and Murre (2007), Dekinder and Kohli (2008), Graham and Frankenberger (2011), Herrington and Dempsey (2005), Kappe et al. (2014), Prabhu et al. (2005), Tellis et al. (2000), Yoo and Mandhachitara (2003), Farace et al. (2019), and Villarroel Ordenes et al. (2019). Recent studies using the Koyck model in other disciplines than marketing are Mulchandani et al. (2019), and Acar and Temiz (2017).

² See Kanetkar et al. (1986), Bass and Leone (1983, 1986), and Weinberg and Weiss (1982), and recently Tellis and Franses (2006).

³ Today still, the impact of temporal aggregation acquires much attention in the marketing literature, see for example, Calli et al. (2012), Kappe et al. (2014), Lambrecht and Tucker (2013), Sethuraman et al. (2011), Sood et al. (2014), Tirunillai and Tellis (2012), and Xi et al. (2014).

⁴ See, for example, exercise 3.3 in Franses, van Dijk and Opschoor (2014, p. 75) which concerns the case where an autoregression of order 1 becomes an autoregressive moving average model of order (1,1). A classic study in this context is Working (1960).

✉ Philip Hans Franses
franses@ese.eur.nl

¹ Econometric Institute, Erasmus School of Economics,
Erasmus University Rotterdam, Rotterdam, The Netherlands



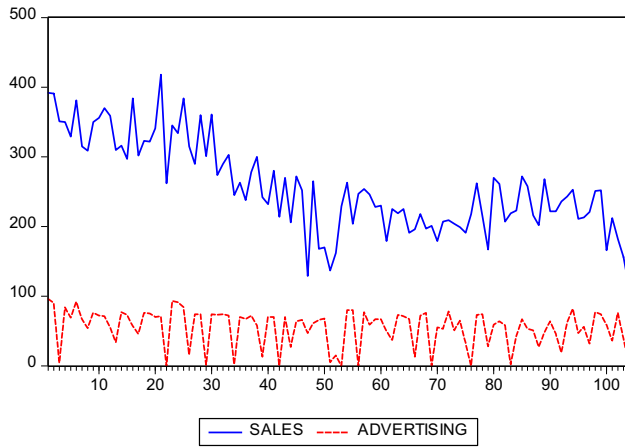


Fig. 1 Weekly sales and advertising data. The data source is Tellis et al. (2000)

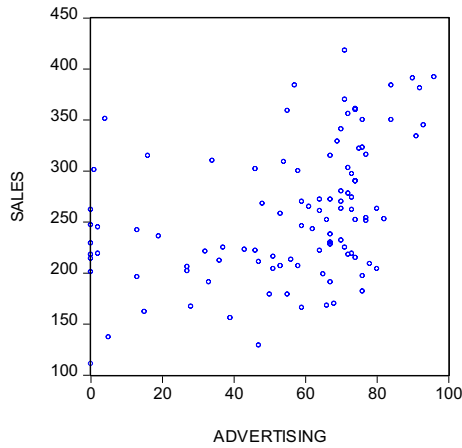


Fig. 2 Scatter plot of weekly sales against weekly advertising

notational convenience, but extensions to more variables are conceptually straightforward.

The outline of the paper is as follows. Section 2 presents the Koyck model. Section 3 discusses four variants of temporal aggregation. Each variant is illustrated with a real-life example. The data source is Tellis et al. (2000). A scatter plot of sales against advertising in Fig. 2 suggests that there is a positive correlation between the two variables. Mathematical derivations are delegated to the technical appendix. Section 4 presents the results of some simulation experiments, and Sect. 5 concludes.

The Koyck model

The Koyck (1954) model, or geometric lag model, yields insights in the key parameters on marketing response. When sales are denoted as y_t , and advertising (or any other marketing-mix variable) as x_t , and L is the familiar lag operator with

$$L^k y_t = y_{t-k}, \quad k = \dots -2, -1, 0, 1, 2, \dots,$$

and deleting the intercept for notational convenience, the original Koyck model reads as follows:

$$y_t = \beta x_t + \beta \lambda x_{t-1} + \beta \lambda^2 x_{t-2} + \dots + \varepsilon_t, \quad (1)$$

where ε_t is an uncorrelated white noise process with mean 0 and variance σ_ε^2 and $|\lambda| < 1$. When using the L operator, it reads as follows:

$$y_t = (\beta + \beta \lambda L + \beta \lambda^2 L^2 + \dots) x_t + \varepsilon_t. \quad (2)$$

Because $|\lambda| < 1$, it holds that.

$$\beta + \beta \lambda L + \beta \lambda^2 L^2 + \dots = \beta (1 + \lambda L + \lambda^2 L^2 + \dots) = \frac{\beta}{1 - \lambda L}. \quad (3)$$

Hence, the infinite regression model in (1) can be written as follows:

$$y_t = \frac{\beta}{1 - \lambda L} x_t + \varepsilon_t. \quad (4)$$

This expression suggests what became known as the ‘‘Koyck transformation,’’ i.e., when both sides of (4) are multiplied with $1 - \lambda L$, one obtains

$$y_t = \lambda y_{t-1} + \beta x_t + \varepsilon_t - \lambda \varepsilon_{t-1}. \quad (5)$$

The Koyck model has an autoregressive term λy_{t-1} , a term involving current advertising βx_t and a so-called moving average term $\varepsilon_t - \lambda \varepsilon_{t-1}$. From the model parameters, one can derive the short-run (or current or direct) effect of advertising, using the partial derivative:

$$\frac{\partial y_t}{\partial x_t} = \beta. \quad (6)$$

The total (or carry-over) effect of advertising follows from

$$\frac{\partial y_t}{\partial x_t} + \frac{\partial y_t}{\partial x_{t-1}} + \frac{\partial y_t}{\partial x_{t-2}} + \dots = \beta + \beta \lambda + \beta \lambda^2 + \dots = \frac{\beta}{1 - \lambda}. \quad (7)$$

As the focus is on the direct effect and the carry-over effect, in practice, one usually considers the unrestricted version of (5), i.e.,



$$y_t = \lambda y_{t-1} + \beta x_t + \varepsilon_t - \theta \varepsilon_{t-1}, \quad (8)$$

where λ and θ are not restricted to be equal from the start.⁵ Suppose that an analyst knows that the advertising response process works at the high-frequency data level, denoted as t , with $t = 1, 2, \dots, N$. For example, t can be associated with weeks within a period of 4 weeks. Suppose further that the sales and advertising data can be available after temporal aggregation at a lower frequency, denoted as T . For example, weekly data could be aggregated to four-weekly data. To introduce some formal notation, consider the polynomial $S(L)$ defined as

$$S(L) = 1 + L + L^2 + \dots + L^{K-1} \quad (9)$$

which amounts to a temporal aggregation of the high-frequency data over K periods. In the case of weeks and hours, K would be equal to 168. Hence, $T = 1, 2, \dots, \frac{N}{K}$. Further, consider the notion of skip sampling at every K th observation at frequency t . This means that, for t equal to $K, 2K, 3K$, and so on, there is an observation at the lower frequency T , with $T = 1, 2, 3, \dots, \frac{N}{K}$. For the hourly case, where the first hour of the week can be 1.00AM on Monday morning, then $K = 168$ concerns 12.00PM on Sunday evening.

Four cases of aggregation

In relation to the frequencies t and T , there are now four cases of potential interest and practical relevance.

High-frequency sales and high-frequency advertising

The first and most simple case is when the analyst has data on sales and advertising both at the high frequency t . A Koyck model as in (8) can be estimated using Maximum Likelihood for the illustrative data, where now also an intercept is included. This results in the following estimates (with estimated standard errors in parentheses) of the two key parameters:

$$\hat{\lambda} = 0.939(0.031)$$

$$\hat{\beta} = 0.279(0.139).$$

The R^2 of this model is 0.682. The short-run effect is 0.279, and the total long-run effect is

$$\frac{0.279}{1 - 0.939} = 4.543.$$

Suppose now that this model for the weekly data corresponds with the true frequency of the sales and advertising relationship. The topic of interest in this paper is that it can happen that one does not have the weekly data, but for example, only four-weekly data. This can occur when commercials are only broadcasted once per four weeks, while sales are measured per week. Or, the other way around that commercials are broadcasted once per week, while sales are only measured at a four-weekly level.

It might be the case that one has (a) weekly data for both sales and advertising as above, (b) weekly data on sales but only four-weekly data for advertising, (c) four-weekly data on sales and weekly data on advertising, or (d) four-weekly data for both sales and advertising. The question is now whether in cases (b), (c), and (d), one can estimate the parameters concerning the true high, weekly, frequency-relating sales with advertising. The key assumption is of course that (a) amounts to the correct frequency, but, note again, this is here for illustration only. Whether it is true for the illustration data is unknown, and therefore, later on a simulation experiment will be carried out. In the high-frequency case, skip sampling will lead to suboptimal inference in terms of efficiency as information will be lost. Consider

$$y_t = \lambda y_{t-1} + \beta x_t + \varepsilon_t - \lambda \varepsilon_{t-1}$$

and K periods later:

$$y_{t+K} = \lambda y_{t+K-1} + \beta x_{t+K} + \varepsilon_{t+K} - \lambda \varepsilon_{t+K-1}.$$

Skip-sampling towards the frequency implied by K would allow the inclusion of y_{t-1} and x_t in the model, but not the moving average term with ε_{t-1} , ε_{t+K-1} , and so on. This, thus, leads to bias in estimating λ . So, when all high-frequency data are available, it is recommended to consider a model for the high-frequency data and not to temporally aggregate the high-frequency data. See also Tellis and Franses (2006) for evidence based on simulations.

High-frequency sales and low-frequency advertising

The second case (b) is where sales are observed at frequency t , while advertising is observed at the lower frequency T after aggregating over K units. In the Appendix, it is derived that the modified Koyck model becomes

⁵ When $\theta = \lambda$, estimation and inference on the parameters in (1) have to incorporate that when $\beta = 0$, the model in (5) collapses to $y_t = \varepsilon_t$ as the term $1 - \lambda L$ cancels on both sides. Put formally, under the null hypothesis of no effect of advertising, the parameter λ is not identified. This so-called Davies (1987) problem makes inference on β non-standard. Franses and van Oest (2007) provide the proper tools for inference, which involves the more complicated method of conditional maximum likelihood. Simulation experiments in Franses and van Oest (2007) show that for large samples, the differences between estimating (8) or (5) are small when it comes to estimating the short-run and carry-over effects.



$$Y_T = \lambda S(L)y_{t+K-1} + \beta X_T + u_T - \theta u_{T-1}. \quad (10)$$

The parameters in (10) can be estimated using the unrestricted maximum likelihood method.⁶ Note that the parameters in (10) are estimated for N/K observations instead of N , and temporal aggregation means loss of efficiency.

For the running example with the data in Figs. 1 and 2, the key estimation results for (10) (with an intercept) for 25 effective observations are

$$\hat{\lambda} = 0.942(0.023)$$

$$\hat{\beta} = 0.363(0.209)$$

The R^2 of this model is 0.974. The short-run effect is 0.363, whereas the total long-run effect at the high frequency is

$$\frac{0.363}{1 - 0.942} = 6.259.$$

We see that this long-run effect is a bit larger than the “true” high-frequency effect of 4.543, while the “true” short-run effects are close to each other.

Low-frequency sales and high-frequency advertising

The third case⁷ is when sales are observed at frequency T , after aggregating over K units, while advertising is observed at the higher frequency t . In the Appendix, it is derived that the modified Koyck model becomes

$$Y_T = \lambda^K Y_{T-1} + \beta(1 + \lambda L + \lambda^2 L^2 + \dots + \lambda^{K-1} L^{K-1}) S(L)x_t + \varepsilon_T - \lambda^K \varepsilon_{T-1}. \quad (11)$$

When the unrestricted version of this model is estimated, that is, when we replace $\varepsilon_T - \lambda^K \varepsilon_{T-1}$ in (12) by $\varepsilon_T - \theta \varepsilon_{T-1}$, then, for the running example, the estimation results obtained using iterative Maximum Likelihood for (11) are

$$\hat{\lambda}^4 = 0.947(0.191),$$

$$\hat{\beta} = -0.133(0.361),$$

which gives $\hat{\lambda} = \sqrt[4]{0.947} = 0.986$. The R^2 of this model is 0.769. The short-run effect is, however, not significant. This may perhaps reflect that the weekly frequency cannot be assumed to be the true frequency.

Low-frequency sales and low-frequency advertising

Finally, the fourth case (d) arises where both sales and advertising are observed only after temporal aggregation at the low-frequency T . Tellis and Franses (2006) conveniently show that when it is assumed⁸ that an advertising impulse occurs only once in each K th period, and at the same time within that K th period, (5) can become

$$Y_T = \lambda^K Y_{T-1} + \beta_1 X_T + \beta_2 X_{T-1} + \varepsilon_T - \lambda^K \varepsilon_{T-1}, \quad (12)$$

where β_1 and β_2 are functions of β and λ , such that

$$\frac{\beta_1 + \beta_2}{1 - \lambda^K} = \frac{\beta}{1 - \lambda}.$$

Tellis and Franses (2006) recommend that if aggregation is necessary, one should collect data such that the key assumption on the advertising process holds.

For the illustrative four-weekly data, the key estimation results for (12), where we replace $\varepsilon_T - \lambda^K \varepsilon_{T-1}$ in (12) by $\varepsilon_T - \theta \varepsilon_{T-1}$, are

$$\hat{\lambda}^4 = 0.902(0.176)$$

$$\hat{\beta}_1 = 1.195(0.967)$$

$$\hat{\beta}_2 = -1.109(0.763)$$

which gives $\hat{\lambda} = \sqrt[4]{0.902} = 0.975$. The R^2 of this model is 0.796. The short-run effect $\beta = \frac{(1-\lambda)(\beta_1+\beta_2)}{1-\lambda^4}$ is 0.022, whereas the total long-run effect for the high-frequency data would be

$$\frac{1.195 - 1.109}{1 - 0.902} = 0.966.$$

We now see that this long-run effect is about one fifth of the “true” high-frequency effect of 4.543. This result may perhaps be driven by the potential fact that the advertising impulse does not occur only once in each four-weekly period, at least for these illustrative data.

⁶ Moreover, in this case, a test of $\beta = 0$ does not suffer from the Davies problem, and hence, standard inference is possible. So, if advertising is only available, say, four-weekly, while sales are recorded weekly, the analysis of the Koyck model follows standard procedures. Note that this also implies that one can purposely aggregate the data in order to avoid the Davies problem.

⁷ In macroeconomics, one often resorts to so-called MIDAS models if one encounters this particular case for non-Koyck models. Relevant references for the MIDAS approach are Andreou et al. (2010), Ghysels et al. (2002, 2007, 2020), Forni et al. (2015).

⁸ This assumption results in much more simpler expressions than those in See Kanetkar et al. (1986), Bass and Leone (1983, 1986), and Weinberg and Weiss (1982).



Table 1 Average estimates of λ , β , θ , and $\frac{\beta}{1-\lambda}$, when averaged over 100 replications, $K = 5$, sample size is 1000

Cases	λ	θ	β $\beta_1 + \beta_2$ in case (d)	$\frac{\beta}{1-\lambda}$
(a)	0.800	0.799	5.001	25.001
(b)	0.804	0.620	4.899	24.994
(c)	0.785	0.946	3.511	16.311
(d)	0.820	0.944	15.443	25.037

In the data-generating process, we set $\lambda = 0.8$ and $\beta = 5$.

Simulation experiments

The empirical results in the previous section for just a single illustrative case in part seem to confirm that the Koyck model is robust to temporal aggregation, at least, after proper modification. Cases (c) and (d) did not work so well in the illustration, although the parameter λ is estimated at a consistent value across the four cases. As this is just a single empirical case with actual data, we now turn to simulation experiments.

The data-generating process (DGP) is

$$y_t = \lambda y_{t-1} + \beta x_t + \varepsilon_t - \lambda \varepsilon_{t-1}$$

with $\varepsilon_t \sim N(0, 1)$, $y_0 = 0$, and x_t is the absolute value of a draw from a $N(0, 1)$ distribution halfway the K -period, and otherwise, it is zero. So, x_t has a positive value after each K periods, where the size of the value can change over time. Tellis and Franses (2006) use the same format for the simulations. Here, we set $K = 5$.

The sample size is set at 1000. The x_t obtains a positive non-zero value at observation 3 within $K = 5$. The short-run effect is set at $\beta = 5$, and we set the decay parameter at $\lambda = 0.8$. Hence, the true carry-over effect is $\frac{5}{1-0.8} = 25$.

Table 1 reports the estimates of λ , β , θ and $\frac{\beta}{1-\lambda}$, when averaged over 100 replications, which is a reasonable amount for a sample size of 1000. Each time, as in the illustration before, we use an unrestricted version of the Koyck model, in terms of the moving average part. The simulation results seem to confirm the theory that the Koyck model is robust to temporal aggregation, for cases (b) and (d), although we observe some bias for case (c). This last bias seems to be caused by (on average) too small an estimate of λ and too small an estimate of β .

Conclusion

This paper has shown that the Koyck (1954) model is a useful model to estimate advertising response at the true high frequency, even when the analyst has temporally

aggregated sales data or temporally aggregated advertising data, or both. Inference using the Koyck model is robust to temporal aggregation. An empirical example, in part, and a simulation exercise, almost fully, emphasized the theoretical claims. Further research should concern more illustrations to see how the Koyck model fares in other empirical settings. Also, more theoretical results can be derived that in case, the Koyck model is extended to more than a single explanatory variable.

The practical implications are that, given a situation of partial or full temporal aggregation of the data, a practitioner can retrieve the proper current and carry-over effects of marketing efforts on marketing response at the high frequency.

Appendix

Case (b): sales are observed at frequency t , while advertising is observed at the lower frequency T after aggregating over K units. To see what the consequences are for the Koyck model, consider applying

$$S(L) = 1 + L + L^2 + \dots + L^{K-1}$$

to both sides of

$$y_t = \lambda y_{t-1} + \beta x_t + \varepsilon_t - \lambda \varepsilon_{t-1}. \text{ This gives}$$

$$S(L)y_t = \lambda S(L)y_{t-1} + \beta S(L)x_t + S(L)\varepsilon_t - \lambda S(L)\varepsilon_{t-1}.$$

Moving ahead K units in time t , this equation reads as

$$S(L)y_{t+K} = \lambda S(L)y_{t+K-1} + \beta S(L)x_{t+K} + S(L)\varepsilon_{t+K} - \lambda S(L)\varepsilon_{t+K-1}.$$

Skip-sampling at every K th observation in t amounts to

$$Y_T = \lambda S(L)y_{t+K-1} + \beta X_T + u_T - \theta u_{T-1}.$$

The Y_T is the temporally aggregated sales variable in a K -period interval, where $S(L)y_{t+K-1}$ can simply be constructed from the available high-frequency sales data at time $t = K, 2K, 3K$, and so on, and where $u_T - \theta u_{T-1}$ is a first-order moving average process with mean zero and where u_T has variance σ_u^2 .

Case (c): sales are observed at frequency T , after aggregating over K units, while advertising is observed at the higher frequency t .

To see how this translates to the Koyck model, one can replace y_{t-1} on the right-hand side of (5) by

$$y_{t-1} = \lambda y_{t-2} + \beta x_{t-1} + \varepsilon_{t-1} - \lambda \varepsilon_{t-2}$$

and repeat this K times to obtain

$$y_t = \lambda^K y_{t-K} + \beta(1 + \lambda L + \lambda^2 L^2 + \dots + \lambda^{K-1} L^{K-1})x_t + \varepsilon_t - \lambda^K \varepsilon_{t-K}$$



Multiplying both sides of this last expression with $S(L)$ gives

$$S(L)y_t = \lambda^K S(L)y_{t-K} + \beta(1 + \lambda L + \lambda^2 L^2 + \dots + \lambda^{K-1} L^{K-1}) S(L)x_t + S(L)\varepsilon_t - \lambda^K S(L)\varepsilon_{t-K}.$$

Skip sampling at each K th observation results in a model for the temporally aggregated data like

$$Y_T = \lambda^K Y_{T-1} + \beta(1 + \lambda L + \lambda^2 L^2 + \dots + \lambda^{K-1} L^{K-1}) S(L)x_t + \varepsilon_T - \lambda^K \varepsilon_{T-1}.$$

With high-frequency data on advertising, the analyst can rely on an iterative Maximum Likelihood method to alternate between estimating λ and creating the relevant observations for $\beta(1 + \lambda L + \lambda^2 L^2 + \dots + \lambda^{K-1} L^{K-1}) S(L)x_t$.

Acknowledgements In 2019, it was 65 years ago that Leendert Koyck defended his PhD thesis at the Econometric Institute of the Netherlands School of Economics, now Erasmus University Rotterdam. His thesis supervisor was Jan Tinbergen. Leendert Koyck died at the age of 44 in 1962. His thesis is still cited today, and it belongs to the heritage of marketing science history. I thank Michael McAleer, Gerard Tellis, and an anonymous reviewer for helpful comments. Thanks are due to Max Welz and Olivier Mulkin for their help with the simulations.

Compliance with ethical standards

Conflict of interest There is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acar, M., and H. Temiz. 2017. Advertising effectiveness on financial performance of banking sector: Turkey case. *International Journal of Bank Marketing* 35 (4): 649–661.
- Andreou, E., E. Ghysels, and A. Kourtellis. 2010. Regression models with mixed sampling frequencies. *Journal of Econometrics* 158 (1): 246–261.
- Bass, F.M., and R.P. Leone. 1983. Temporal aggregation, the data interval bias, and empirical estimation of bimonthly relations from annual data. *Management Science* 29 (January): 1–11.
- Bass, F.M., and R.P. Leone. 1986. Estimating micro relationships from macro data: A comparative study of two approximations of the brand loyal model under temporal aggregation. *Journal of Marketing Research* 23 (August): 291–297.
- Berkowitz, David, Arthur Allaway, and Gilles D'Souza. 2001. Estimating differential lag effects of multiple media across multiple stores. *Journal of Advertising* 30 (4): 59–65.
- Breuer, R., M. Brettel, and A. Engelen. 2011. Incorporating long-term effects in determining the effectiveness of different types of online advertising. *Marketing Letters* 22 (4): 327–340.
- Calli, M.K., M. Weverbergh, and P.H. Franses. 2012. The effectiveness of high-frequency direct-response commercials. *International Journal of Research in Marketing* 29 (1): 98–109.
- Chessa, A.G., and J.M.J. Murre. 2007. A neurocognitive model of advertising content and brand name recall. *Marketing Science* 26 (1): 130–141.
- Clarke, D.G. 1976. Econometric measurement of the duration of advertising effect on sales. *Journal of Marketing Research* 13 (4): 345–357.
- Davies, R.B. 1987. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 64: 247–254.
- Dekinder, J.S., and A.K. Kohli. 2008. Flow signals: How patterns over time affect the acceptance of start-up firms. *Journal of Marketing* 72 (5): 84–97.
- Farace, S., A. Roggeveen, F. VillarroelOrdenes, K. De Ruyter, M. Wetzels, and D. Grewal. 2019. Patterns in motion: How visual patterns in ads affect product evaluations. *Journal of Advertising* 49 (1): 3–17.
- Foroni, C., M. Marcellino, and C. Schumacher. 2015. Unrestricted mixed data sampling (U-MIDAS): MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society, Series A* 178: 57–82.
- Franses, P.H., and R. van Oest. 2007. On the econometrics of the geometric lag model. *Economics Letters* 95: 291–296.
- Ghysels, E., V. Kvedaras, and V. Zemlys-Balevicius. 2020. *Mixed data samples (MIDAS) regression models, Chapter 4 in Handbook of Statistics*, vol. 42, 117–153.
- Ghysels, E., P. Santa-Clara, and R. Valkanov. 2002. The MIDAS touch: mixed data sampling regression models, Working paper, UNC and UCLA.
- Ghysels, E., A. Sinko, and R. Valkanov. 2007. MIDAS regressions: further results and new directions. *Econometric Reviews* 26 (1): 53–90.
- Graham, R.C., and K.D. Frankenberger. 2011. The earnings effects of marketing communication expenditures during recessions. *Journal of Advertising* 40 (2): 5–24.
- Herrington, J.D., and W.A. Dempsey. 2005. Comparing the current effects and carryover of national-, regional-, and local-sponsor advertising. *Journal of Advertising Research* 45 (1): 60–72.
- Kanetkar, V., C.B. Weinberg, and D.L. Weiss. 1986. Recovering micro parameters from aggregate data for the Koyck and brand loyal models. *Journal of Marketing Research* 23 (August): 298–304.
- Kappe, E., A.S. Blank, and W.S. DeSarbo. 2014. A general multiple distributed lag framework for estimating the dynamic effects of promotions. *Management Science* 60 (6): 1489–1510.
- Koyck, L.M. 1954. *Distributed Lags and Investment Analysis*. Amsterdam: North Holland.
- Lambrecht, A., and C. Tucker. 2013. When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research* 50 (5): 561–576.
- Leone, R.P. 1995. Generalizing what is known about temporal aggregation and advertising carryover. *Marketing Science* 14 (3): G141–G150.
- Mulchandani, K., K. Mulchandani, and R. Attr. 2019. An assessment of advertising effectiveness of Indian banks using Koyck model. *Journal of Advances in Management Research* 16 (4): 498–512.
- Prabhu, J.C., R.K. Chandy, and M.E. Ellis. 2005. The impact of acquisitions on innovation: Poison pill, placebo or tonic? *Journal of Marketing* 69 (1): 114–130.



- Sethuraman, R., G.J. Tellis, and R.A. Briesch. 2011. How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *Journal of Marketing Research* 48 (3): 457–471.
- Sood, A., E. Kappe, and S. Stremersch. 2014. The commercial contribution of clinical studies for pharmaceutical drugs. *International Journal of Research in Marketing* 31 (1): 65–77.
- Tellis, G.J., R.K. Chandy, and P. Thaivanich. 2000. Which ad works, when, where, and how often? Modeling the effects of direct television advertising. *Journal of Marketing Research* 37 (1): 32–46.
- Tellis, G.J., and P.H. Franses. 2006. Optimal data interval for estimating advertising response. *Marketing Science* 25 (3): 217–229.
- Tirunillai, S., and G.J. Tellis. 2012. Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Science* 31 (2): 198–215.
- VillarroelOrdenes, F., D. Grewal, S. Ludwig, K. De Ruyter, D. Mahr, and M. Wetzels. 2019. Cutting through content clutter: How speech and image acts drive consumer sharing of social media brand images. *Journal of Consumer Research* 45 (5): 988–1012.
- Weinberg, C.B., and D.L. Weiss. 1982. On the econometric measurement of the duration of advertising effect on sales. *Journal of Marketing Research* 19 (November): 585–591.
- Xi, L., J.A. Duan, and A. Whinston. 2014. Path to purchase, A mutually exciting point process model for online advertising and conversion. *Management Science* 60 (6): 1392–1412.
- Yoo, B., and R. Mandhachitara. 2003. Estimating advertising effects on sales in a competitive setting. *Journal of Advertising Research* 43 (3): 310–321.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

