Empirically Testing for Dynamic Causality between Promotions and Sales Beer Promotions and Sales in England

Ray Huffaker[1] and Andrew Fearne[2]
[1]Agricultural and Biological Engineering, University of Florida, USA
[2]Kent Business School (Medway Campus), UK
rhuffaker@ufl.edu; a.fearne@kent.ac.uk

Abstract

We devise a decision tool to help economic researchers select a causal detection method compatible with real-world dynamics of an economic system under investigation. We apply it to test for dynamic causality between promotions and sales of a case-study beer brand in England. We find evidence that promotions and sales data for the brand are generated by a nonlinear deterministic dynamic system. Under these circumstances, conventional Granger Causality detection methods impose unreasonable restrictions on real-world market dynamics, and thus must give way to recently formulated Cross Convergent Mapping (CCM) methods. Our application of CCM methods provides strong evidence that promotions and sales for the brand are bi-causal. Promotions have a long-term impact on sales, and sales have a long-term impact on promotion decisions.

Keywords: *promotions, sales, causality, nonlinear dynamics*

1   Introduction

Firms want to know whether promotional expenditures have an impact on sales over time, and economists have been keen to find answers. Empirical causal analysis that is incompatible with real-world marketing dynamics is inaccurate and unreliable for planning. There is critical need to select the causal detection mechanism that matches the dynamic marketing structure of sales and promotions.

Past work investigated dynamic consumer response to promotions in stable markets with infinite distributed lag specifications whose stable behavior restricted post-promotion sales to return to pre-promotion levels [1]. Later work studied dynamic consumer response in stochastically-trending markets. Sales were modeled as linear stochastic processes responding to promotions cast as single-shot random shocks. Stable market behavior allows for post-shock sales to evolve toward higher levels so that promotions can have long-run 'persistent' impacts [2]. Work that followed recast promotions as a random variable joining sales in a bivariate linear stochastic process, and applied cointegration analysis to test for Granger causality (GC) between the two series [3]. If the series are found to share a common stochastic trend, then they co-evolve toward a long-run equilibrium, and GC exists in at least one direction [4].

This approach imposed untested, and possibly unrealistic, restrictions on real-world market dynamics. First, the models required a stable equilibrium to determine causality between promotions and sales. Do real-world promotions and sales data really show evidence of stable steady-state dynamics? Second, the models cast sales and promotions as random variables, or promotions as unexpected random shocks. Do real-world promotions and sales data really exhibit random dynamic behavior, or do they exhibit structured dynamics that economists might expect from strategic market behavior? Finally, Granger Causality (GC) tests conducted to find dynamic causality between promotions and sales required 'separability'—an intrinsic characteristic of purely stochastic and linear models. This means that all the information about a variable can be removed by eliminating it from the model. The removed variable X is said to 'Granger cause' Y if predictability of Y declines when X is removed from the set of possible causal factors [5].
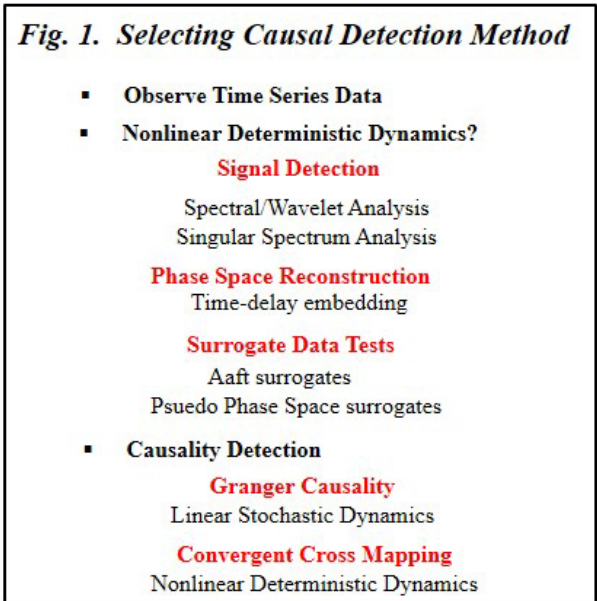
Sugihara *et al.* (2013) demonstrated that GC gives ambiguous results for nonlinear deterministic dynamic systems, which are intrinsically non-separable[6]. These are dynamic systems in which 'everything depends on everything else', or as explained by the naturalist John Muir (1911), *"[w]hen we try to pick something up by itself, we find it hitched to everything else in the universe."* Nonlinear feedback relationships encode information about X into Y, and this information is not lost by removing X from the system. Sugihara *et al.* (2013) developed the Convergent Cross Mapping (CCM) technique—based on nonlinear state space reconstruction—to test for causality in nonlinear deterministic dynamic systems.

A key implication of Sugihara *et al.* (2013) is that researchers can increase the reliability of empirical causality results by first testing whether real-world system dynamics are nonlinear, and then applying the compatible causal detection method. We devise a decision tool to help economic researchers select a causal detection method compatible with the real-world dynamics of the economic system under investigation. We apply this approach to test for dynamic causality between promotions and sales of beer in England. Our data cover 104 weeks (February 2009 to January 2011) of promotions (Dunnhumby data) and sales (Tesco loyalty-card data) of 335 brands of beer in 14 categories. We first test the null hypothesis that promotions and sales data are generated by a nonlinear deterministic dynamic system, acceptance of which calls for application of the CCM causal-determination approach. We find compelling evidence of nonlinear deterministic dynamic structure in the promotions and sales data for several brands. Future research will apply this framework to re-test a number of causal hypotheses from the literature.

2   A Decision Tool for Selecting a Causal Detection Method in Economic Research

Fig. 1 is a broad schematic summarizing how to diagnose real-world dynamics and use this information to select an appropriate causal detection method. Initial steps test whether data are likely generated by nonlinear deterministic dynamic systems. If nonlinear deterministic dynamics are found, CCM is used for causal detection. Alternatively, if linear stochastic dynamics are more likely, then GC is used for causal detection.

Diagnosing System Dynamics. We first apply 'signal detection' methods to each time series to separate structural content from random noise. Spectral analysis identifies periodic patterns (such as daily and seasonal cycles) in the time series, and Singular Spectrum Analysis (SSA) decomposes the time series into the sum of structural (trend and oscillations) and an unstructured residual. R-packages 'spectrum' and 'Rssa' are available for spectral analysis and SSA, respectively.

Fig. 1. Selecting Causal Detection Method

- **Observe Time Series Data**
- **Nonlinear Deterministic Dynamics?**
    Signal Detection
    Spectral/Wavelet Analysis
    Singular Spectrum Analysis

    Phase Space Reconstruction
    Time-delay embedding

    Surrogate Data Tests
    Aaft surrogates
    Psuedo Phase Space surrogates
- **Causality Detection**
    Granger Causality
    Linear Stochastic Dynamics

    Convergent Cross Mapping
    Nonlinear Deterministic Dynamics

Following Vautard (1999), we next use the structural component of the SSA decomposition for each time series to recover the behavioral properties of the real-world dynamic system that generated it ('phase space reconstruction') [7]. This is possible because, if the dynamic system generating the time series is nonlinear, then each time series records the history of its interaction with all other interrelated system variables. We employ the 'time-delay' embedding method of phase space reconstruction which accounts for the multidimensionality of real-world dynamic systems by segmenting a time series *X(t)* into a sequence of delay coordinate vectors: *X(t-d)*, *X(t-2d)*,..,*X(t-(m-1)d)* where '*d*' is the time delay and '*m*' is the number of delayed coordinate vectors (the 'embedding dimension'). The sequence of delay coordinate vectors is collected as columns in an 'embedded data matrix', and the reconstructed phase space is a scatterplot of the multidimensional points constituting the rows of this matrix. The method is firmly rooted in mathematical topology theory thanks to Takens (1980) who showed that an embedding dimension of *m = 2n+1* (where '*n*' is the dimension of the real-world attractor) suffices to reconstruct a

phase space with the same dynamical properties as the phase space from the real-world system ('topological' invariance) [8]. R-package 'tseriesChaos' is available to compute optimal embedding dimension and time delay.

We take three conventional measures to discriminate low-dimensional deterministic structure in reconstructed state space. These are the correlation dimension (measuring the extent to which points on a reconstructed phase space attractor are spatially organized), the Lyapunov exponent (measuring sensitivity to initial conditions and resultant spreading of state-space trajectories over time), and relative forecast error (measuring how well a reconstructed attractor forecasts the future) [9].

Surrogate Data Tests. We conduct surrogate data analysis to test whether apparent nonlinear deterministic structure detected in phase space reconstruction is statistically significant or simply the figment of a mimicking stochastic process [10]. A set of surrogate data vectors is randomly generated destroying intertemporal patterns in observed data while preserving various statistical properties. Phase space is reconstructed for each surrogate vector, and the above three discriminating measures taken. The mean from the distribution of each measure for the set of surrogate vectors is tested for statistical significance from the corresponding value taken from the SSA-filtered data. Statistically insignificant differences indicate that detected structure is better attributed to stochastic behavior.

We generate surrogate data for two conventionally-tested stochastic processes: (1) Aaft (amplitude-adjusted Fourier transform) surrogates generated as static monotonic nonlinear transformations of linearly filtered noise, which preserve both the probability distribution and power spectrum of the SSA-filtered data [10]; and (2) PPS (pseudo phase space) surrogates testing for the presence of a noisy limit cycle by preserving periodic trends in the SSA-filtered data while destroying chaotic structures (Small and Tse, 2002). We followed methods outlined by Kaplan and Glass (1995) to write R-code to generate Aaft surrogate data [9], and methods outlined by Small and Tse (2002) to write R-code generating PPS surrogates [11].

Causal Detection. If surrogate data analysis supports low-dimensional nonlinear dynamics, we apply CCM to detect causality between economic variables of interest, for example, between promotions and sales in a marketing system. In general, CCM tests whether there is correspondence between reconstructed phase spaces for two observed time series variables. The underlying logic is that causally related variables reconstruct the same real-world phase space dynamic. For example, if Y drives X, then phase space reconstructed from X can be used to estimate ('cross map') values of Y, but not vice versa. If Y and X have a bi-causal relationship, then each can be cross mapped from the reconstructed phase space of the other. We followed methods outlined in Supplementary Materials to Sugihara *et al*. (2013) to write R-code for CCM [6].

Alternatively, if surrogate data analysis supports the assumption of separable stochastic dynamics, we apply co-integration time series methods to test for GC. If co-integration analysis finds that time series' share a common stochastic trend, then they co-evolve toward a long-run equilibrium, and GC exists in at least one direction [4].

3   Results

We applied the decision tool outlined in Fig. 1 to test for causality relationships between promotions and sales for the Leffe Blonde brand. The promotions and sales data are reported as market shares of category totals and graphed in Fig. 2A. Promotions (red line) and sales (blue line) market shares appear to oscillate at similar frequencies. This is validated by the SSA reconstructions (Fig. 2B) in which promotions and sales market shares oscillate at 7 and 10 week periods explaining 80% of the combined variance in the observed data.
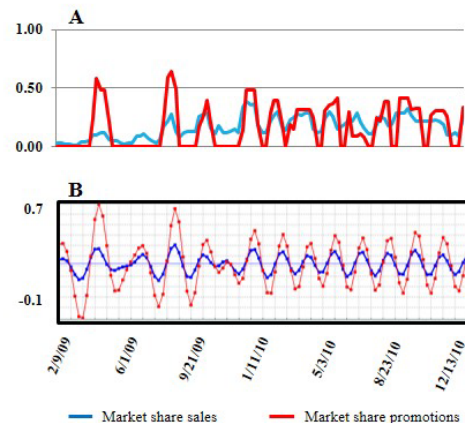

Fig. 2. (A) Graphs of promotions and sales market Shares for Leffe Blonde brand. (B) SSA reconstruction

Nonlinear dynamic methods succeeded in reconstructing low-dimensional and nonlinear attractors[1] from behavioral patterns in SSA reconstructed promotions and sales market shares (Fig. 3). The reconstructed attractors are characterized by irregular 7 and 10 week oscillations. Surrogate data tests reject the hypothesis that the attractors are the figment of a mimicking linear stochastic process. Consequently, convergent cross mapping (CCM) is the proper causal detection method.

CCM results are displayed in Fig. 4. The graphs plot correlation coefficients (ρ) between CCM predictions and the associated observed values for an increasing portion of the data. Correlation coefficients that converge to one are evidence of strong causal relationships. This holds for both the cross mapping of sales market shares with the promotions attractor and vice versa. This indicates that a bi-causal relationship exists between promotions and sales dynamics for the Leffe Blonde brand.

Fig. 4. Convergent Cross Mapping. (A) Cross mapping of sales with promotions phase space. (B) Cross mapping of promotions with sales phase space.

4 Conclusion

Contrary to past work investigating the effectiveness of promotions on sales with stochastic models, our case study uses a data centric approach providing strong evidence that promotions and sales dynamics are characterized by deterministic nonlinear temporal patterns. Under these circumstances, conventional Granger Causality detection methods impose unreasonable restrictions on real-world market dynamics, and thus must give way to Cross Convergent Mapping (CCM) methods. Our application of CCM methods provided strong evidence that promotions and sales are bi-causal. Promotions have a long-term impact on sales, and sales have a long-term impact on promotion decisions.

A limitation of a nonlinear dynamics approach is that attractor reconstruction in market studies is not a given, and can fail for a couple of major reasons [12]. For example, market dynamics in a particular application may not be governed by a low-dimensional attractor, or noisy or limited data may prevent an existing low-dimensional attractor from being detected. When nonlinear dynamic techniques fail to detect market patterns, conventional stochastic approaches remain a viable alternative. However, we propose that researchers initially test for systematic patterns in the data before presuming stochastic structures potentially falling short of real-world complexity.

References

[1] M. Dekimpe, D. Hanssens, Time-series models in marketing: Past, present and future, International Journal of Research in Marketing, 17 (2000) 183-193.

[2] M. Dekimpe, D. Hanssens, V. Nijs, E. Steenkamp, Measuring short- and long-run promotional effectiveness on scanner data using persistence modeling, Applied Stochastic Models in Business and Industry, 21 (2005) 409-416.

[3] V. Nijs, M. Dekimpe, E. Steenkamp, D. Hanssens, The category-demand effects of price promotions, Marketing Science, 20 (2001) 1-22.

[4] C. Granger, Some recent developments in a concept of causality, Journal of Econometrics, 39 (1988) 199-211.

---

[1] In general, an attractor exists when system dynamics pull initial conditions toward a spatially-organized structure upon which system variables either remain constant (a 'point' attractor), orbit periodically (a 'limit cycle' attractor), or orbit irregularly (a 'strange' attractor) as time approaches infinity.
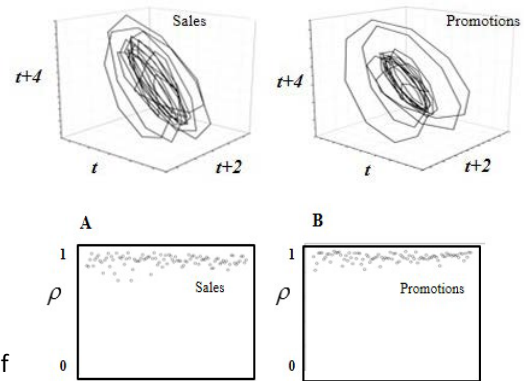
[5] C. Granger, Investigating causal relations by econometric models and cross spectral methods, Econometrica, 37 (1969) 424-438.

[6] G. Sugihara, R. May, Y. Hao, H. Chih-hao, E. Deyle, M. Fogarty, S. Munch, Detecting causality in complex ecosystems, Science, 338 (2012) 496-500.

[7] R. Vautard, Patterns in time: SSA and MSSA in analysis of climate, in: H. von Storch, A. Navarra (Eds.) Analysis of Climate Variability,Springer, 1999.

[8] F. Takens, Detecting strange attractors in turbulence, in: D. Rand, Young, L. (Ed.) Dynamical Systems and Turbulence Springer, New York, 1980, pp. 366-381.

[9] D. Kaplan, L. Glass, Understanding Nonlinear Dynamics, Springer, New York, 1995.

[10] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, Farmer, J., Testing for nonlinearity in time series: The method of surrogate data, Physica D, 58 (1992) 77-94.

[11] M. Small, C. Tse, Applying the method of surrogate data to cyclic time series, Physica D, 164 (2002) 187-201.

[12] G. Williams, Chaos Theory Tamed, John Henry Press, Washington D.C., 1997.