

On the detection of effective marketing instruments and causality in VAR models

Csilla Horváth and Pieter W. Otter

University of Groningen
Groningen, The Netherlands

SOM-theme F Interactions between consumers and firms

Abstract

Dynamic multivariate models become more and more popular in analyzing the behavior of competitive marketing environments. Takada and Bass (1998), Dekimpe, Hanssens and Silva-Rosso (1999), and Dekimpe and Hanssens (1999) recommend to use Vector Autoregressive (VAR) models because they provide full-scale linear approximations of dynamic competitive marketing environments, including all structural relationships: sales response effects, competitive reactions, feedback effects, and purchase reinforcement effects. The drawback of VAR models is the large number of parameters to be estimated. This requires preliminary analysis concerning the selection of variables to be included in the model. We propose to use canonical correlation for this purpose. Canonical correlation, furthermore, provides the tool of testing the existence of the structural relationships between (lagged) consumer response and (lagged) marketing instruments. The canonical correlation and causality testing procedures are applied to data consisting of market shares and marketing instruments in a market of nondurable goods.

Key words: Canonical correlation, data reduction, causality testing

1 Introduction

Vector Autoregressive (VAR) systems have become increasingly popular and have been adopted by economists all over the world since the seminal work of Sims (1980). Recently, with an increasing interest in identifying competitive structures and measuring relationships between marketing variables, VAR models have become recognized as an effective modeling technique in marketing. Parsons and Shultz (1976) identify four key elements that need to be incorporated in a parsimonious model for competitive markets: simultaneous relationships, interactions, carryover, and competitive effects. Dekimpe and Hanssens (1995), Takada and Bass (1998), Dekimpe, Hanssens and Silva-Rosso (1999), and Dekimpe and Hanssens (1999) recommend to use Vector Autoregressive (VAR) models because these models incorporate all the structural relationships¹. Takada and Bass (1998) propose to use Vector Autoregressive Moving Average (VARMA) models to analyze competitive marketing behavior because they show that these models are capable of capturing the dynamic competitive structure of the market and find that VARMA models outperform univariate time series models in goodness-of-fit measures as well as in forecasting performance.

The drawback of VAR models is the large number of parameters, which can cause some degrees of freedom problems. If k endogenous variables are included in a VAR of order p , the number of parameters to be estimated is $k^2 \cdot p$, so the number of variables that can be included in the model is rather limited. This requires preliminary analysis concerning the selection of the most influential variables to be included into the model.

We propose to use canonical correlation for this purpose. Canonical correlation is a member of the family of multivariate statistical techniques. A number of authors consider canonical correlation as a logical extension of simple and multiple correlation. Multiple regression analysis is a more special case in which one of the sets of data contains only one variable, while product moment correlation is the most special in that both sets of data contain only one variable. Thompson (1984) states that “Given that canonical correlation analysis can be as complex as reality in which most causes have multiple effects and most effects are multiply caused, an “advanced organizer” regarding some of the research questions that can be addressed using canonical analysis may be helpful.”

¹Hanssens (1980) mention sales response effects, competitive reactions, and feedback effects, which are also captured in VAR models, as the main types of market relationships and urges the development of models that contain all of them simultaneously.

To analyze the interrelationships among the set of multiple criterion variables (market shares) and the set of multiple predictors (marketing instruments) and to assess the extent to which variables intercorrelate to define the same general patterns, we apply canonical correlation analysis. Canonical correlation, furthermore, provides the tool for testing the existence of structural relationships on particular markets, and therefore the appropriateness of the model. We use it for testing Wiener-Granger causality from the instruments to the set of market shares and vice versa.

We first introduce the some theory in section 2 and in section 3 the procedures are applied to a dataset for a market of nondurable goods. Section 4 concludes the article.

2 Wiener-Granger causality and canonical correlation

In this section a brief explanation of Wiener-Granger (WG) causality and canonical correlation is given. The interested reader is referred to Geweke (1982), Geweke, Meese and Deut (1983), and Otter (1990, 1991).

Consider two stationary, zero-mean Gaussian processes $\{y_t\}$ with $\dim\{y_t\} = m_1$ and $\{x_t\}$ with $\dim\{x_t\} = m_2$ where $m_1 \leq m_2$. Define $Y_t = \{y_{t-s}; s \geq 0\}$ and $X_t = \{x_{t-s}; s \geq 0\}$. Let $z_t = (y_t', x_t)'$ and $Z_t = \{z_{t-s}; s \geq 0\}$. Denote the one-step ahead error covariance matrix using Z_t to predict y_{t+1} optimally by $\Sigma_{(y_{t+1}|Z_t)}$.

Definition 1 : X is said to cause Y in the Wiener-Granger sense if, for some t , $\Sigma_{(y_{t+1}|Z_t)} \neq \Sigma_{(y_{t+1}|Z_t \setminus X_t)}$ where \setminus denotes set theoretic subtraction.

Definition 2 : X is said to cause Y instantaneously in the Wiener-Granger sense if, for some t , $\Sigma_{(y_{t+1}|Z_t \cup x'_{t+1})} \neq \Sigma_{(y_{t+1}|Z_t)}$.

Consider the finite vector $Z_{t-1} = (z'_{t-1}, z'_{t-2}, \dots, z'_{t-p})'$, the history of z from p periods beforehand. Then

$$\begin{pmatrix} z_t \\ Z_{t-1} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \right) \quad (1)$$

and the conditional expectation (the best one-step ahead prediction) of z_t given Z_{t-1} is

$$E\{z_t|Z_{t-1}\} \equiv z_t^p = V_{12} \cdot V_{22}^{-1} \cdot Z_{t-1}. \quad (2)$$

From equation (2) it can be seen that z_t^p is the conditional expectation of a VAR(p) model. For details see Geweke (1982) and Mittnik and Otter (1989). The conditional one-step ahead error covariance matrix is

$$\Sigma_{(z_t|Z_{t-1})} = V_{11} - V_{12} \cdot V_{22}^{-1} \cdot V_{21}. \quad (3)$$

In this paper we are interested whether x_t the vector of (lagged) marketing instruments causes y_t , the lagged market shares, and vice versa in WG-sense. Geweke's measure of (instantaneous) WG-causality is:

$$F = \ln \left(\det \Sigma_{(y_t|A)} / \det \Sigma_{(y_t|B)} \right)$$

where \ln denotes the natural logarithm. In case of WG causality (Definition 1.) $A = Z_{t-1}$ and $B = Z_{t-1} \setminus X_t$ and in case of instantaneous causality (Definition 2.) $A = Z_{t-1} \cup x_t$ and $B = Z_{t-1}$. As shown by Otter (1991), F can be expressed in terms of canonical correlation coefficients:

$$F = \sum_{i=1}^{m_1} \ln(1 - \sigma_{i,I}^2) - \sum_{i=1}^{m_1} \ln(1 - \sigma_{i,II}^2)$$

where $\{\sigma_{i,I}^2\}$ are the canonical correlation coefficients² between the sets y_t and A and $\{\sigma_{i,II}^2\}$ are the canonical correlation coefficients between y_t and B with $\sigma_{1,j}^2 \geq \sigma_{2,j}^2 \geq \dots \geq \sigma_{m_1,j}^2$, $j = I, II$. Under the null-hypothesis of no WG-causality ($F = 0$) the large sample test statistic $N \cdot \widehat{F}$ follows a χ^2 -distribution with the number of degrees of freedom equal to the number of prior restrictions in A to get B , see Geweke (1982). \widehat{F} is based on estimated canonical correlation coefficients, which in turn are based on consistent estimates of (co)variance matrices. N denotes the number of observations.

Further, we are interested in the possible direct relation between y_t , the vector of market shares, and x_t , the vector of marketing instruments. Because of the assumptions

²The process of canonical correlation is as follows. We have two sets of variables, $x = (x_1, x_2, \dots, x_{m_1})$ and $y = (y_1, y_2, \dots, y_{m_2})$, and let $m_1 \leq m_2$. On these sets of variables, linear transformations denoted by $(\eta_1, \eta_2, \dots, \eta_{m_1})$ and $(\zeta_1, \zeta_2, \dots, \zeta_{m_2})$ respectively, are carried out, satisfying the following conditions:

1. (a) the η 's are mutually uncorrelated;
- (b) the ζ 's are mutually uncorrelated;
- (c) for $i \neq j$, η_i and ζ_j are uncorrelated;
- (d) for $i = j$, the canonical correlation coefficient between η_i and ζ_i is denoted by σ_i ;
- (e) the σ_i are maximized;
- (f) $\text{var}(\eta_i) = 1$ and $\text{var}(\zeta_j) = 1$; ($i, j = 1, \dots, m_1$).

we have

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right) \quad (4)$$

also the conditional expectation of y_t given x_t is

$$E\{y_t|x_t\} = \hat{y}_t = \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot x_t$$

with error covariance matrix

$$\Sigma_{(y_t|x_t)} = \Sigma_{11} - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \Sigma_{21}.$$

Applying the canonical correlation procedure we have linear transformations $\eta_t = L_1 \cdot y_t$ and $\zeta_t = L_2 \cdot x_t$ such that;

$$\begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} I & \Lambda \\ \Lambda' & I \end{pmatrix} \right) \quad (5)$$

where Λ is obtained from the following singular value decomposition (SVD):

$$\Sigma_{11}^{-1/2} \cdot \Sigma_{12} \cdot \Sigma_{22}^{-1/2} = H' \cdot \Lambda \cdot Q. \quad (6)$$

and $\Lambda = (\Lambda_1; 0)$ with $\Lambda_1 = \text{diag}(\sigma_1, \dots, \sigma_{m_1})$ and $\sigma_1, \dots, \sigma_{m_1}$ are the canonical coefficients (singular values) such that $\sigma_1 \geq \sigma_2 \dots \geq \sigma_{m_1} \geq 0$.

The linear transformation matrices are

$$L_1 = H \cdot \Sigma_{11}^{-1/2} \text{ and } L_2 = Q \cdot \Sigma_{22}^{-1/2}$$

where H and Q are obtained from the singular value decomposition described in equation (6). From the linear transformations we have that the conditional expectations of η_t given ζ_t is

$$\eta_t = \Lambda \cdot \zeta_t = \Lambda_1 \cdot \zeta_{1,t} \quad (7)$$

where $\zeta_{1,t}$ is the m_1 -dimensional subvector of ζ_t which consists of the first m_1 -components of ζ_t . From equation (7) it follows that

$$y_t = L_1^{-1} \cdot \Lambda_1 \cdot L_{2,1} \cdot x_t$$

where

$$\begin{pmatrix} \zeta_{1,t} \\ \zeta_{2,t} \end{pmatrix} = \begin{pmatrix} L_{2,1} \\ L_{2,2} \end{pmatrix} \cdot x_t$$

The conditional covariance matrix is $\Sigma_{(\eta_t|\zeta_t)} = I - \Lambda \cdot \Lambda' = I - \Lambda_1^2$, from which it can be seen that there is a relation between ζ_t and η_t and hence between y_t and x_t if one or more canonical correlation coefficients differ from zero, which, for large samples, can be tested by the Bartlett test statistic;

$$S = - \left[N - \frac{1}{2} \cdot (m_1 + m_2 + 1) \right] \cdot \sum_{i=k+1}^{m_1} \ln(1 - \widehat{\sigma}_i^2) \quad (8)$$

Under the null hypothesis: $\sigma_{k+1} = \sigma_{k+2} = \dots = \sigma_{m_1} = 0$ S follows asymptotically a χ^2 distribution with $(m_1 - k) \cdot (m_1 - k)$ degrees of freedom. The estimated canonical coefficients ($\widehat{\sigma}_i, i = 1 \dots m_1$) are based on consistent estimates of Σ_{11} , Σ_{12} and Σ_{21} . See for details Otter (1990, 1991) .

3 An Application

The market research company, A.C. Nielsen (The Netherlands) B. V. provided aggregated data of 150 stores on a frequently purchased, nondurable, nonfood, consumer product sold in the Netherlands. Our data set consists of weekly series of 76 data points of market shares (Ms) and five marketing mix variables; price (Pr), featuring (Fe), sampling (Sa), bonus (Bo), and refund (Re), for seven large brands. The seven brands together account for approximately 70 percent of the market. To include all these variables into our model we would have 36 endogenous variables³. Therefore, we apply canonical correlation analysis between the set of performance and marketing mix variables to arrive at quantitative relations between the two groups of variables that take into account all available information on the variables concerned. We use the Augmented Dickey Fuller (ADF) test to test for stationarity. All the series turn out to be stationary. All series are standardized.

³Not all the brands used each marketing instruments. Brand one, for example used only price, featuring, and sampling.

3.1 Canonical Correlation Analysis

The canonical correlation procedure, explained in section two, can be applied to the set of market shares, $ms_t = (Ms_{1,t}, \dots, Ms_{7,t})$ with $dim(ms_t) = 7$ and the set of marketing instruments ($inst_t = (Bo_{2,t}, \dots, Pr_{7,t})$) with $dim(inst_t) = 29$. The loadings of instruments are given in Table 1. and Table 2. presents the seven canonical coefficients. All series are standardized.

Table 1. *Canonical loadings of the instruments ($L'_{2,1}$ matrix)*

<i>Pr1</i>	-0,581	0,240	0,506	0,541	-0,227	0,003	0,320
<i>Pr2</i>	-0,215	-0,060	-0,479	-0,257	-0,870	-0,130	-0,023
<i>Pr3</i>	0,128	0,271	0,270	-0,171	0,721	-0,767	0,161
<i>Pr4</i>	-0,086	-0,359	-0,031	-0,016	0,364	0,060	-0,138
<i>Pr5</i>	0,182	-0,203	0,009	0,479	-0,274	-0,757	-0,286
<i>Pr6</i>	-0,184	-0,140	-0,014	-0,363	0,121	0,579	-0,314
<i>Pr7</i>	-0,112	0,290	0,311	-0,128	0,282	0,697	0,079
<i>Fe1</i>	0,292	0,030	-0,201	0,086	-0,152	0,282	0,137
<i>Fe2</i>	-0,056	-0,097	0,000	0,299	0,187	0,118	-0,178
<i>Fe3</i>	-0,063	-0,095	0,082	-0,095	-0,046	0,164	-0,237
<i>Fe4</i>	0,248	0,351	0,173	0,068	-0,360	0,040	0,203
<i>Fe5</i>	-0,082	0,116	-0,084	-0,143	0,307	-0,045	-0,137
<i>Fe6</i>	0,004	-0,064	-0,139	0,178	-0,049	0,174	-0,008
<i>Fe7</i>	-0,172	-0,079	0,143	-0,051	-0,056	-0,180	-0,127
<i>Sa1</i>	-0,042	0,021	0,155	-0,064	0,066	-0,158	-0,020
<i>Sa2</i>	-0,138	-0,058	-0,163	-0,269	-0,028	-0,469	-0,261
<i>Sa3</i>	0,078	-0,155	0,057	0,098	0,062	-0,191	0,301
<i>Sa4</i>	-0,063	0,083	0,221	-1,007	-0,254	0,198	-0,391
<i>Sa5</i>	0,032	-0,058	0,139	0,012	-0,027	0,194	0,112
<i>Sa6</i>	0,106	0,329	-0,886	0,611	-0,247	-0,072	-1,998
<i>Sa7</i>	-0,022	-0,684	-0,473	-0,606	-0,054	-0,280	-0,513
<i>Bo2</i>	-0,270	-0,930	0,047	-0,040	0,676	-0,334	0,627
<i>Bo4</i>	0,314	0,646	0,252	-0,956	-0,582	0,480	-1,768
<i>Bo6</i>	-0,182	0,768	0,254	0,114	0,051	-0,670	-1,337
<i>Bo7</i>	-0,001	-0,281	-0,052	0,505	-0,259	0,114	-1,995
<i>Re4</i>	0,182	0,341	0,047	-0,543	-0,641	0,168	-0,985
<i>Re5</i>	-0,252	0,079	0,137	-0,410	0,239	-0,330	-0,200
<i>Re6</i>	-0,162	-0,160	0,054	-0,485	-0,091	0,659	0,860
<i>Re7</i>	-0,131	0,285	0,077	0,334	-0,005	-0,222	-0,149

Table 2. *Canonical coefficients:*

$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$	$\hat{\sigma}_4$	$\hat{\sigma}_5$	$\hat{\sigma}_6$	$\hat{\sigma}_7$
0,948	0,926	0,915	0,874	0,767	0,757	0,533

We calculated the Bartlett's test statistic, based on equation (8), for the two sets of variables and tested the significance of the Bartlett values. The Bartlett values are presented in Table 3.

Table 3. *Bartlett values:*

$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
541,184	411,349	301,222	198,646	117,056	66,998	18,905

Only $H_0 \sigma_7 = 0$ (case $k = 6$) cannot be rejected. So the first six canonical correlation coefficients contain significant information indicating the dimensionality of the system is six. This makes sense because as the sum of market shares is approximately constant over time, so one of market-share is a linear combination of the other market shares. On the other hand it indicates high complexity of the system as the movements cannot be explained by lower dimensions. From Table 1. it can be seen that all marketing instrument play significant role on the market.

3.2 Instantaneous Wiener- Granger Causality of Market Shares on the Instruments

A main argument for using a VAR model is that it incorporates two important structural relationships: feedback and reinforcement. The purchase decisions of consumers are determined by the previous choices of the customers. This phenomenon is called the purchase reinforcement effect and is captured by the effect of lagged market shares on the current market shares. Feedback effect means the impact of (changes of) sales or market shares on the managerial decisions. This is typically the case when firms use a percent of sales method for determining their advertising budget. Furthermore, as marketing managers fight for better performance of their brand, they try to realign any recent or past objectionable deviation of it. This means that marketing managers adjust their spending depending on their own past and present performance and allow marketing expenditures to change with changing market conditions.

To examine the existence of feedback effects we use Wiener-Granger causality. According to the definition 1 of Wiener-Granger causality described we test for instantaneous Wiener-Granger causality. We define the following conditional expectations (and the conditional covariance matrices): $E\{y_t|A\}$ and $E\{y_t|B\}$, where $y_t = inst_t \in \mathfrak{R}^{29}$, $A = (inst_{t-1}, ms_{t-1}) \in \mathfrak{R}^{35}$, and $B = (inst_{t-1}, ms_{t-1}, ms_t) \in \mathfrak{R}^{42}$.

The estimated Geweke's F measure is $N \cdot \hat{F}_{ms \rightarrow inst} = 525.13$ indicating significant Wiener Granger causality from the market shares towards the marketing instruments (Geweke's F measure follows χ^2 distribution with 29×7 degrees of freedom in all case). It means that the level of market shares of own and competing brands plays an important role in the marketing decisions of the managers. The information in the process of market shares is substantial in addition to all other information in the market. It also means short reaction time in the managerial reaction function.

3.3 Instantaneous Wiener- Granger Causality of the Instruments on Market Shares

Using the same method we also investigated the pertinence of the consumers' reaction. Similarly to the procedure above we define the following conditional expectations (and the conditional covariance matrices): $E\{y_t|A\}$ and $E\{y_t|B\}$, where $y_t = ms_t \in \mathfrak{R}^7$, $A = (inst_{t-1}, ms_{t-1}) \in \mathfrak{R}^{36}$, and $B = (inst_{t-1}, ms_{t-1}, inst_t) \in \mathfrak{R}^{65}$.

Geweke's measure for Wiener-Granger causality is $N \cdot \hat{F}_{inst \rightarrow ms} = 1269.45$. It indicates significant Wiener-Granger causality from the instruments towards the market shares. It means that market share can be predicted more efficiently if the information in the process of instruments is taken into account in addition to all other information. It also means short reaction time in the consumer reaction function.

3.4 Wiener- Granger Causality of Market Shares on the Instruments

According to the definition 2 of Wiener-Granger causality described we test for instantaneous Wiener-Granger causality. We define the following conditional expectations (and the conditional covariance matrices): $E\{y_t|A\}$ and $E\{y_t|B\}$, where $y_t = inst_t \in \mathfrak{R}^{29}$, $A = (inst_{t-1}) \in \mathfrak{R}^{29}$, and $B = (inst_{t-1}, ms_{t-1}) \in \mathfrak{R}^{36}$. The estimated Geweke's F measure is $N \cdot \hat{F}_{ms \rightarrow inst} = 356.62$ indicating significant Wiener Granger causality from the market shares towards the marketing instruments.

3.5 Wiener- Granger Causality of the Instruments on Market Shares

Using the same method we also investigated the pertinence of the consumers' reaction. Similarly to the procedure above we define the following conditional expectations (and the conditional covariance matrices): $E\{y_t|A\}$ and $E\{y_t|B\}$, where $y_t = ms_t \in \mathfrak{R}^7$, $A = (ms_{t-1}) \in \mathfrak{R}^{36}$, and $B = (inst_{t-1}, ms_{t-1}) \in \mathfrak{R}^{65}$. Geweke's measure for Wiener-Granger causality is $N \cdot \hat{F}_{inst \rightarrow ms} = 376.72$. It indicates significant Wiener-Granger causality from the instruments towards the market shares. It means that market share can be predicted more efficiently if the information in the process of

instruments is taken into account in addition to all other information. It also means short reaction time in the consumer reaction function.

4 Conclusion

The canonical correlation procedure and its associated Wiener-Granger causality testing based on the canonical correlation coefficients appeared to be an effective tool to detect the most efficient marketing instruments in a VAR model and to test for the existence of the structural relations in the model. In the application, all the marketing instruments appeared to play an important role on the highly complex market. Also, we found both causality from the set of market shares towards the set of marketing instruments and significant causality from the marketing instruments towards the market shares. Therefore we have bidirectional WG-causality or in other words a feedback system that inspires the application of VAR models in modeling and analyzing the dynamic behavior in competitive marketing systems.

References

- Dekimpe, M. G. and Hanssens, D. M. (1995), 'Empirical generalization about market evolution and stationarity', *Marketing Science* **15**, G109–G121.
- Dekimpe, M. G. and Hanssens, D. M. (1999), 'Sustained spending and persistent response: A new look at long-term marketing profitability', *Journal of Marketing Research* **36**, 397–412.
- Dekimpe, M. G., Hanssens, D. M. and Silva-Rosso, J. M. (1999), 'Long-run effects of price promotions in scanner markets', *Journal of Econometrics* **89**, 269–91.
- Geweke, J. (1982), 'Measurement of linear dependence and feedback between multiple time series', *Journal of the American Statistical Association* **77**, 323–324.
- Geweke, J., Meese, R. and Deut, W. (1983), 'Comparing alternative tests of causality in temporal systems', *Journal of Econometrics* **21**, 161–194.
- Hanssens, D. M. (1980), 'Market response, competitive behavior, and time series analysis', *Journal of Marketing Research* **27**, 470–85.
- Mitnik, S. and Otter, P. W. (1989), Determination of linear feedback between multiple time series, in 'Dynamic Modelling and Control of National Economics, Preprints IFAC Symposium', Vol. 1, Pergamon Press, New York.
- Otter, P. W. (1990), 'Canonical correlation analysis in multivariate time series analysis with an application to one-year-ahead and multiyear-ahead macroeconomic forecasting', *Journal of Business and Economic Statistics* **8**, 453–457.
- Otter, P. W. (1991), 'On Wiener-Granger causality, information and canonical correlation', *Economics Letters* **35**, 187–191.
- Parsons, L. J. and Shultz, R. L. (1976), *Marketing Models and Econometric Research*, North Holland Publishing Company, Amsterdam.
- Sims (1980), 'Macroeconomics and reality', *Econometrica* **48**, 1–48.
- Takada, H. and Bass, F. M. (1998), 'Multiple time series analysis of competitive marketing behavior', *Journal of Business Research* **43**, 97–107.
- Thompson, B. (1984), *Canonical Correlation Analysis: Uses and Interpretation*, Beverly Hills [etc.] : Sage Publications.