UNIVERSITÉ DE GENÈVE

Archive ouverte UNIGE
https://archive-ouverte.unige.ch

Rapport de recherche 2002

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Using Partial Least Regression in Marketing Research

Graber, Stéphane; Czellar, Sandor; Denis, Jean-Emile

How to cite

GRABER, Stéphane, CZELLAR, Sandor, DENIS, Jean-Emile. Using Partial Least Regression in Marketing Research. 2002

This publication URL: https://archive-ouverte.unige.ch/unige:5822

© This document is protected by copyright. Please refer to copyright holder(s) for terms of use.

Using Partial Least Squares Regression in Marketing Research

Stéphane Graber

Sandor Czellar

Jean-Emile Denis

University of Geneva

December 2002

Using Partial Least Squares Regression in Marketing Research

Abstract

Small datasets, missing values and the presence of multicollinearity often plague samples used in marketing research. Partial-least squares regression (PLS) is a statistical tool specifically designed to cope with such problems. The paper proceeds first with a concise review of the PLS method. It then highlights particular conditions when applying a standard OLS regression leads to unstable statistical results. The argument is supported with an empirical application on the market orientation/performance link. In conclusion, it is called for an increased reliance on the use of PLS in research.

Keywords: marketing research; multivariate analysis; PLS regression; market orientation: business performance

Partial Least-squares Regression

Marketing researchers are often faced with the following problems: a survey is administered but the number of respondents is lower than expected; those who did reply did not answer to all the items contained in the survey; to top it all, responses to different parts of the survey are strongly correlated. How, then, may these data be reliably used for regression-type analyses? Under these conditions, ordinary least squares (OLS) regression yields unstable results due to a small sample size and missing values. Moreover, multicollinearity between predictors in an OLS regression increases the standard error of their estimated coefficients (Field, 2000). In other words, a high level of multicollinearity increases the risk of a theoretically sound predictor to be rejected from the regression model as a non-significant variable.

Partial least squares (PLS) regression analysis was developed in the late seventies by Herman O. A. Wold (Wold, 1975, Wold & al. 1984). It is a statistical tool that has been specifically designed to deal with multiple regression problems where the number of observations is limited, missing data are numerous and the correlations between the predictor variables are high. These characteristics of PLS regression have been demonstrated both with real data and in simulations (Garthwaite, 1994; Tenenhaus, 1998). PLS regression has known a great success in scientific fields where the problem of linking a great number of correlated variables

through a limited number of observations is particularly acute, such as chemistry (Frank & Friedman, 1993; Wold, 1993). Its application by researchers in management and marketing science - who may well face the same kind of difficulty - has remained surprisingly rare, despite calls from several scholars (Ryan, Rayner & Morrison, 1999; Cassel, Hackl & Westlund, 2000).

In technical terms, PLS regression aims at producing a model that transforms a set of correlated explanatory variables into a new set of uncorrelated variables (Tenenhaus, 1998; Umetri AB, 1998). The parameter coefficients in a PLS regression are derived from the direct correlations between the predictor variables and the criterion variable. The SIMCA 8.0 software package has specifically been designed for PLS analyses (Umetri AB, 1998) and is used in this study. Popular statistical packages also incorporate PLS extension modules, such as the PLS procedure in the SAS System (SAS Institute, 1996).

Methodology

We shall illustrate the usefulness of the PLS method by choosing as a case in point empirical research on the market orientation – performance relationship. Theory holds that a focus on marketing orientation leads to a better business performance (Jaworski & Kohli, 1990). Most studies test the main and interaction effects of market orientation, external variables (especially market growth, distributor power, competition intensity, entry barriers) and internal variables (mainly firm size and cost) on performance (e. g. Slater & Narver, 1994; Jaworski & Kohli, 1993; Greenley, 1995). In this paper, we will only investigate the main effect of market orientation and control variables on business performance. We chose this topic as an illustrative case because past studies investigating the impact of market orientation on a firm's performance have exclusively relied on standard ordinary least squares (OLS) regression models. Through a systematic literature review we identified twelve studies carried out in various countries investigating the main effect of market orientation on performance, which include internal and external control variables (Jaworski & Kohli, 1993; Narver & Slater, 1990; Slater & Narver, 1994; Greenley, 1995; Appiah-Adu, 1998; Appiah-Adu & Ranchhod, 1998; Kumar & Subramanian, 1998; Pulendran, Speed & Widing, 2000; Dawes, 2000; Subramanian & Gopalakrishnan, 2001; Harris, 2001; Rose & Shoham, 2002). The overall result of these studies is rather inconclusive: market orientation is shown to exert a significant positive impact on performance in only half of the studies, and surprisingly most

control variables appear to have no direct effect on performance. However, none of these studies addressed rigorously the issue of multicollinearity. This paper shows that in this type of research multicollinearity may well be present and lead to unstable OLS regressions.

The authors of this paper undertook an empirical study similar to those identified above. It was carried on Swiss firms from four business sectors: food, beverages, hygiene products and cosmetics. Companies were selected using three databases, in a progressive order. First, a list of consumer goods companies belonging to the top 2000 Swiss firms was produced (TOP 2000, 1996). This list was then expanded with two other Swiss commercial business databases: Sysmar (provided by a local marketing research firm) and the Kompass Suisse Repertoire (Kompass Switzerland and Liechtenstein, 1997/98). The names of the key marketing managers were then secured for each firm. A questionnaire was administered through CATI to the sample of respondents. It consisted of questions on market orientation, external and internal control variables/moderators and performance measures according to the self-reported format adopted in previous studies (e.g. Slater & Narver, 1994; Greenley, 1995). As in past research, all items were measured on a 1-5 Likert scale. 372 people were contacted, and 229 answered the questionnaire resulting in an effective response rate of 55%. The number of usable questionnaires was 175. Both the response rate and the number of usable questionnaires is line with those of past research. In the past twelve studies previously identified, the response rates typically vary between 30 and 60 % and the number of usable questionnaires between 60 and 250.

Results

We first ran a classical OLS analysis investigating the impact of market orientation on firm's efficiency, in the spirit of Narver and Slater (1990) and Greenley (1995). Three indicators of potential multicollinearity are commonly used and therefore investigated here: (1) the matrix of bivariate correlations, (2) the variance inflation factors, and (3) the un-centred cross-products matrix (Field, 2000).

Bivariate correlations between the predictor variables are presented in Table 1. Seventeen out of the 28 correlation coefficients between the predictor variables are significant at the 95% level, some of them with values higher than 0.6. It is not unlikely to observe high correlations between some market environment variables. Most of them were a priori expected on the

basis of mainstream business strategy theory (Porter, 1998; Jarillo, 1993). Note that most of the past studies did not report the correlation matrix at all. Those who did report it obtained correlation scores and structures comparable to ours (e.g. Subramanian & Gopakrishnan, 2001). They did not investigate however the impact of this condition on the stability of the OLS regression parameters.

Table 1 - Correlations between independent variables
(Pearson correlation coefficients and number of cases per cell)

	Size	Cost	Market growth	Distr. Power	Tech. turbulence	Comp. Intensity	Entry barriers	Market orientation
Size	1 155							
Cost	.003 142	1 153						
Market growth	.225* 155	-.001 149	1 171					
Distr. Power	.202* 120	-.114 116	.018 131	1 133				
Tech. turbulence	.105 154	-.094 148	.103 166	.131 131	1 169			
Comp. Intensity	.211* 159	-.196* 153	.045 171	.590* 133	.290* 169	1 175		
Entry barriers	.287* 159	-.255* 152	.054 169	.666* 131	.322* 167	.767* 173	1 173	
Market orientation	.310* 159	-.188* 153	.087 171	.465* 133	.260* 169	.590* 175	.605* 173	1 175

The variance inflation factor (VIF) is another way to assess the magnitude of multicollinearity in regression models. Although there is no valid statistical rule on which to judge a VIF as critical, Field (2000) suggests that an average VIF value over 1 should raise concerns. Past studies either did not examine VIF factors or if they did (like Dawes, 2000), they applied less stringent criteria, such as the cut-point of 10 (Greenley & Foxall, 1998, Hair & al, 1995). Table 2 presents the VIF for each main variable in the model.

Table 2 - Impact of market orientation on efficiency: variance inflation factors (VIF)

Independent variable	VIF
Market orientation	1.853
Relative size	1.152
Relative cost	1.133
Market growth	1.034
Distributor's power	2.051
Technological turbulence	1.132
Competition intensity	2.976
Entry barriers	3.699

As the data in Table 2 show, several VIF factors are much greater than 1 meaning that significant collinearity could be present in the data. The VIF of market orientation and

external variables such as distributor power, competition intensity and entry barriers are particularly high, in accord with results obtained by Dawes (2000).

A third collinearity indicator is the eigenvalues of the scaled, un-centred cross-products matrix. They are not reported here because of space constraints but are available upon request from the authors. If it contains an eigenvalue that is much greater than the others, the regression is unstable and might easily be affected by small variations in the variables. This is effectively the case since our first eigenvalue was 8.243 whereas the other eight ranged between 0.014 and 0.303. None of the past twelve studies referred to earlier had used this information to assess the stability of their OLS regression.

On the evidence of these three indicators, we conclude that the data does present significant multicollinearity problems. Therefore, the reliance on an OLS regression model is likely to yield biased parameter estimates, and potentially lead to the deletion of some predictor variables as statistically non-significant even when they are actually strongly linked to the criterion variable. For this reason, both OLS and PLS regressions were performed on the data. We used standard OLS regression as it was used in the past twelve studies under comparison. Table 3 presents the direct effects of market orientation and control variables on efficiency measured by OLS and PLS regressions as well as the Pearson correlation coefficients.

Table 3 - Impact of market orientation on efficiency: Comparison of PLS and OLS regression results

Independent variable	Pearson correlation with the dependent variable	PLS regression parameter estimate (n=148)[1]	OLS regression parameter estimate (n=92)
Intercept		.486**	.222**
Market orientation	.218**	.140**	.068
Relative size	.434**	.201**	.388*
Relative cost	.080	.040	.149
Market growth	.124	.063	.022
Distributor's power	.231*	.079*	.105
Technological turbulence	.127	.063	-.040
Competition intensity	.220**	.076**	.130
Entry barriers	.233**	.073**	.015
Adjusted R^2		.155	.132

* p<.05, ** p <.01

[1] In PLS, *n* is greater than in OLS, due to the fact that in PLS regression missing values for one independent variable do not influence the parameter estimation of the other variables; whereas in standard OLS, a missing value for one variable leads to the deletion of the whole observation.

Several anomalies are observed in the OLS analysis. Theoretically, in absence of multicollinearity, the Pearson coefficients and OLS parameter estimates should show a similar significance pattern. However, only one of the eight predictor variables has a significant coefficient in the OLS, whereas one would expect several more to be significant on the basis of the strength of the Pearson correlation coefficients listed in the second column in Table 3. Moreover, marketing theory and past empirical evidence support the significance of PLS parameter estimates obtained for most external variables. Alarmingly, the market orientation parameter is insignificant in the OLS whereas it is significant in the PLS model. Table 3 illustrates the risk of obtaining misleading results in studies such as this one since multicollinearity leads to the rejection of key variables when relying on the standard OLS regression model. The problem is of considerable concern since in our study OLS regression lead to the rejection of the key variable of interest namely market orientation.

For the sake of comparison, let us have a look at Greenley's (1995) study, which is quite similar to ours. It is based on 240 usable questionnaires filled out by managers selected from a Dun & Bradstreet listing of the largest UK firms. Greenley's OLS results are very similar to our own OLS results in terms of R^2 and the sign/significance of parameter estimates. It should be noted in particular that both his and our own OLS analysis led to the conclusion that there was no significant relation between market orientation and efficiency. However, this conclusion is reversed in our study when performing a PLS regression. Based on these results, we argue that a re-examination of past studies on the market orientation – performance link with the PLS regression technique is likely to lead to more convergent results across different studies. This, in turn, could lead to a greater generalisability of empirical evidence on the market orientation - performance relationship.

Future Research

Survey data are sensitive to multicollinearity, which warrants the use of appropriate statistical tools. This paper illustrates that using standard statistical tools under these conditions leads to erroneous conclusions. Variables are rejected from a theoretical model when in fact they should not be. Future empirical research on topics involving potentially correlated constructs should consider using partial least squares and other statistical techniques, which do cope with multicollinearity rather than standard OLS, which does not. Our study shows that high multicollinearity may be expected in firm-based surveys where several simultaneous measures of external market characteristics are taken (e.g. technological turbulence, market

growth, distributor power, competition intensity, entry barriers etc). Other research domains that may benefit from PLS-type analysis include studies covering many related independent variables such as research on consumer loyalty (Ryan, Rayner & Morrison, 1999); investigations of specific industries characterized by very few observations such as market niches or surveys of sensitive subjects (e.g. proprietary business information, sensitive consumption aspects), which are likely to lead to incomplete and correlated answers from respondents.

References

Appiah-Adu, K. (1998). Market orientation and performance: empirical test in a transition economy. *Journal of Strategic Marketing*, 6 (1), 25-45.

Appiah-Adu, K., & Ranchhod, A. (1998). Market Orientation and Performance in the Biotechnology Industry: An Exploratory Empirical Analysis. *Technology & Strategic Management*, 10 (2), 197-210.

Cassel, C. M., Hackl, P., & Westlund, A. H. (2000). On Measurement of Intangible Assets: A Study of Robustness of Partial Least Squares. *Total Quality Management, 11*(7), 897-908.

Dawes, J. (2000). Market Orientation and Company Profitability: Further Evidence Incorporating Longitudinal Data. *Australian Journal of Management, 25*(2), 173-199.

Field, A. (2000). *Discovering Statistics Using SPSS for Windows*. London: Sage
Bartlett, J. Gani, Ed. London: Academic Press.

Garthwaite, P. H. (1994). An Interpretation of Partial Least Squares. *Journal of the American Statistical Association*, 89 (425), 122-27.

Greenley, G. E. (1995). Market Orientation and Company performance: empirical evidence from UK Companies. *British journal of Management*, 6, 1-13.

Greenley, G. E., & Foxall, G. R. (1998), "External moderation of associations among stakeholder orientations and company performance," *International Journal of Research in Marketing*, 15, 51-69.

Harris C., L. (2001). Market orientation and performance: objective and subjective empirical evidence from UK companies. *Journal of Management Studies*, 38 (1), 17-43.

Jarillo, J. C. (1993). *Strategic Networks: Creating the Borderless Organisation*. Oxford: Butterworth-Heinemann.

Jaworski, B. J., Kohli, A. K. (1993), "Market orientation: Antecedents and Consequences," *Journal of Marketing*, 57 (3) : 53-70.
Kompass: Répertoire Général de l'Économie de la Suisse et du Liechtenstein, 47° éd. 1997/98 (1997) Zurich, Editions Kompass Suisse S.A.

Kumar, K., & Subramanian, S. (2000). Navigating the External Environment Through a Market Orientation. *S.A.M. Advanced Management Journal*, 65 (1), 16-30.

Narver, J. C., & Slater, S. F. (1990). The Effect of a Market Orientation on Business Profitability. *Journal of Marketing*, 54 (October), 20-35.

Porter, M. E. (1998). *Competitive Strategy*. New York: Free Press.

Pulendran, S., Speed, R., & Widing, R. E. (2000). The Antecedents and Consequences of Market Orientation in Australia. *Australian Journal of Management, 25*(2), 119-143.

Rose, G. M., Shoham, A. (2002). Export performance and market orientation: Establishing an empirical link. *Journal of Business Research*, 55, 217-25.

Ryan, M. J., Rayner, R., & Morrison, A. (1999). Diagnosing Customer Loyalty Drivers. *Marketing Research, 11*(Summer), 19-26.

SAS Institute, Inc. (1996), *SAS/STAT Software, Version 6.11, Proc PLS.* Cary NC: SAS Institute Inc.

Subramanian, R., & Gopalakrishna, P. (2001). The market orientation- performance relationship in the context of a developing economy. An empirical analysis. *Journal of Business research* (53), 1-13.

Tenenhaus, M. (1998). La régression PLS. Paris : Publi-Union.
TOP 2000, Les plus grandes entreprises en Suisse, 1996. 1996. Genève et Zurich: PME Magazine et Handelszeitung.

Umetri AB (1998), *User's Guide to SIMCA Version 7.0.* Umetri AB, Umea: Sweden.

Wold, S., A. Ruhe, H. Wold, and W. J. III Dunn (1984), "The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses," *SIAM Journal on Scientific and Statistical Computing*, 5 (3), 735-43.

Wold, H. (1975). Soft Modeling by Latent Variables: the Non-linear Iterative Partial Least Squares Approach, in *Perspectives in Probability and Statistics, Papers in Honour of M. S.*