

HOW THE RESULTS OF MICRO-MARKETING CAMPAIGNS OF A NEWSPAPER COULD BE LINKED TO THE SOCIO-ECONOMICAL LOCAL REALITIES

Maria Silvia Farina, Alís

1. Introduction

General market condition of Italian newspapers

The origin of most Italian newspapers is typically linked to a region or a geographical area where the publication initially took place. Historically, during their life time, many newspapers have tried expand beyond their geographical limit, increasing the involvement in general themes, skipping from a local dimension to a more national perspective and including international content. At the end of the above mentioned process, the need to gain new markets forced the national newspapers to face up to the problem of circulation inside the local areas different from the original one where they developed their market, with the contrasting strength and historical power of local publishers.

The local publishers have high penetration results in term of readership, thanks to a specialised service focused on a rich contribution of local information together with an overview of general information referring to national themes. To improve their performances in local areas, the national newspapers, traditionally providers of general information service at a national level (news, politics), developed specialised sections for regions and provinces inside the regions, by increasing the readers interest taking into account readers' reduced budget and their unwillingness to renounce local information.

The case of Sole 24 Ore

Inside the national newspapers market, the situation of the "Sole 24 ore", the Italian market leader title for economic news, is roughly the same in terms of the need to increase local penetration. But, in contrast to other national newspapers, the case of "Sole 24 ore" presents the following characteristics

1. the difficulty in adapting the editorial content by adding local information
2. facility to be sold as an addition to local newspaper service, because of its specialised contribution.

In fact the information supplied by the title comprehends the traditional information (national news, politics) and, particularly, specialised sections devoted to economic resolutions, laws, prescribed times of administrative executions. Thanks to its specialised content Sole 24 ore doesn't represent a real competitor for local newspapers, because it plays the role of a title giving complementary information.

The above considerations convinced the marketing staff of "Sole 24 ore" to launch a campaign to improve the local circulation of the title, with copies of the title sold, through a special offering called "sandwich", jointly with a local newspaper.

The study

Because of the general successful trend of the title, measured both in terms of circulation and readership, the publisher was interested in knowing the probable reasons of variations read in small geographical areas, where the "sandwich" campaign took place.

Considering the type of information supplied by the title "Sole 24 ore", to prepare the future local campaigns, the marketing management decided to promote a study at micro geographical level to find out:

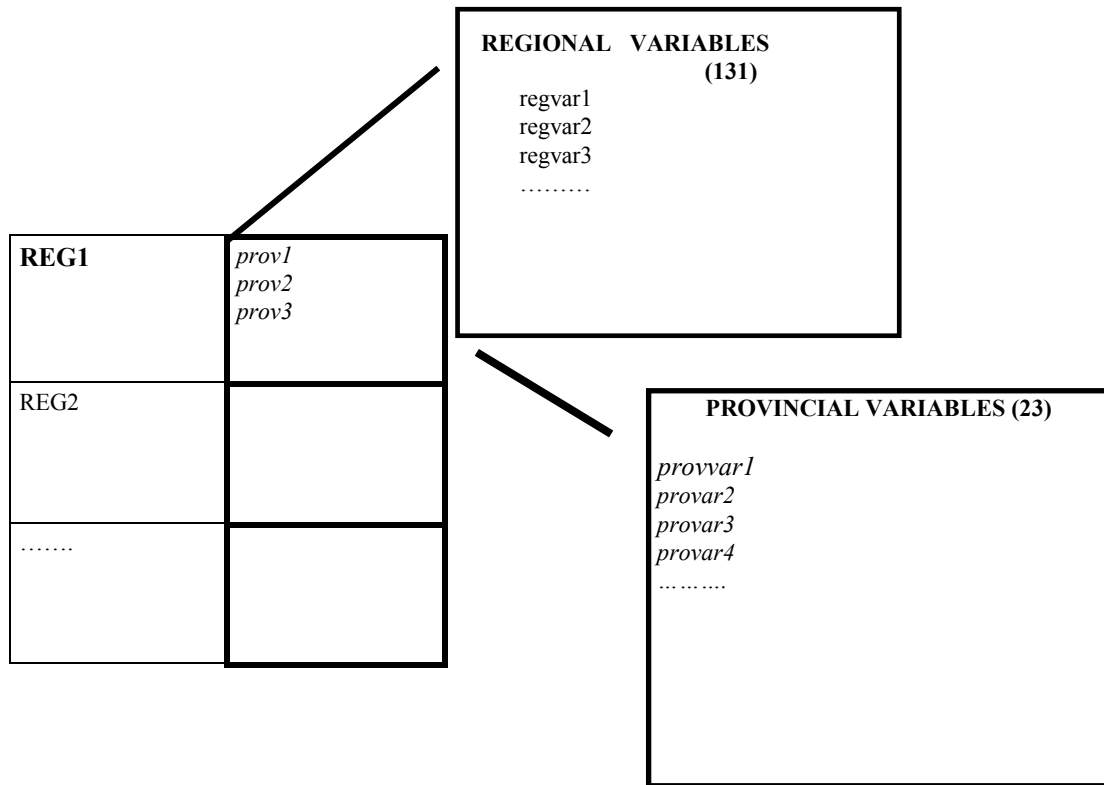
- a. what were the external variables conditioning the "core" component of the title penetration
- b. what were the external variables more related with the "movement" component of the title penetration.

For external variables we mean variables regarding demographic, social and economical environment.

Because of the obvious limits of the information available, it was necessary to review the client’s request, adapting the objectives of the study in the following way:

- a. the “CORE” component of the penetration was studied using the widely available information for regions
- b. the “MOVEMENT” component of the penetration was studied at provincial level, considering the small geographical realities where the changes are more consistent and accepting limited information contribution.

The structure of the study can be represented in the following diagram .



The study is conceived on two complementary dimensions:

- A. the “STRUCTURAL/STATIC” dimension observed in a fixed time frame (observations referred to a specific year, the most recent available) at regional level
- B. the “MOVEMENT/DYNAMIC” dimension observed in a series of time frames (repeated observations referred to a couple of years) at provincial level, considering provinces as different situations inside the homogeneous structure of their region.

It’s clear that the goal of the study was to find a global description of phenomena influencing the title penetration, by evaluating the relative weight of static and dynamic components.

Results

The study shows that

- 1. the title penetration is related to some macro-economical variables, measuring both the traditional economical phenomena of regions and the recent development of new activities
- 2. the title’s penetration is also linked to the presence of some socio-cultural variables
- 3. finally the variations in the title’s penetration are closely linked to the socio-cultural variables.

The study indicates a basic relation with the presence of the traditional needs of economic information (industry, companies, services) and a new relation with socio-cultural variables measuring the individual interest of a typically private market, characterised by curiosity, new needs and flexibility.

2. Methodology

Information basis

The variables, considered as dependent phenomena, are circulation and readership of "Sole 24 ore".

The variables measuring the external phenomena are

- 131 variables for regions
- 23 variables for provinces.

To avoid the unwanted effects arising from the use of absolute values, we decided to consider dimension-less values, created using common general reference standards, represented by the national figures.

Applying this procedure, all the values, including absolute values, are transformed into indices

penetration variables

- firstly we calculated the percentage, as a ratio between the data and total population at national/ regional /provincial level
- secondly we calculated an index, as a ratio between the regional/provincial percentage and the national percentage obtained in the previous step

quantitative variables on population

- firstly we calculated the mean value at national / regional / provincial level, dividing values by the reference population
- secondly we calculated an index, as ratio between the regional/provincial mean and the national mean obtained in the previous step.

All the data are used on the scale 0-100, leaving out the real quantitative contribution to the results .

The chosen normalisation of data supply a good premise for developing a strict data analysis procedure.

Models

The study was carried out developing sequential steps designed according to the following criteria:

- the progressive reduction of the available variables
- the identification of relations using the original variables or linear combination of them.

Regional mode

In the case of the regional variables we have to manage such a large number of variables that we decided to reduce them by mainly using a variables clustering procedure. Thanks to this approach, we can ensure the control of the well known effect of strong correlation between economic variables . The procedure adopted is used like a filtering process finalised to recognise the best representative variables inside homogeneous sets of information.

The total set of 131 variables measure demographic, social and economical phenomena. Our objective is to reduce the variables to an ideal maximum of ten.

The results of the clustering procedure are the following

- identify homogeneous groups of (correlated) variables
- attribution of each variable to a group (only one group)
- describing quantitatively the groups in terms of variance caught
- measuring variables contribution in terms of relative weight inside its groups.

Step 1

The means to chose the best subset is achieved using a multivariate procedure consisting of clustering variables.

We used the procedure as a variable-reduction method. In fact, a large set of variables can be often replaced by the set of the variables that more contribute to the clusters' principal components. Furthermore if cluster components do not generally explain as much variance as principal components on the full set of variables, nevertheless the cluster components are usually easier to interpret than the principal components.

According to the above beliefs, we wanted to build hard disjoint clusters, where the original variables are grouped inside different clusters. The clustering procedure is based on the correlation matrix and the clusters are chosen to maximise the variation accounted by the first principal component of each cluster.

The procedure attempts to divide variables into non overlapping clusters in such a way that each cluster can be interpreted as essentially unidimensional. The first principal component is a weighted average of variables that explains as much variance as possible inside the cluster.

$$y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p \quad \text{i.e.} \quad y_1 = a'_1x$$

The second principal component, is the linear combination

$$y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p \quad \text{i.e.} \quad y_2 = a'_2x$$

As previously stated there is no constraint that components are uncorrelated. The procedure tries to maximise the sum across clusters of the variance of the original variables that is explained by the clusters components.

The procedure find 8 groups and it isolates only one principal component for each group, satisfying the condition to have the corresponding eigenvalue greater than one.

The disjoint groups built from the initial set of 131 variables are here described below.

Table 1

GROUPS	DESCRIPTION	NO.OF ORIGINAL VARIABLES	STEP 1 CHOSEN VARIABLES
TOTAL	-	131	18
NO.1	DEMOGRAPHICS	16	2
NO.2	FAMILY INCOME	12	2
NO.3	INDUSTRY STRUCTURE	22	1
NO.4	LABOUR FORCES GEN. ACTIVITY SECTION	18	3
NO.5	LABOUR FORCES PROF.	7	2
NO.6	CONSUMPTION	24	3
NO.7	CULTURAL INTEREST	12	2
NO.8	READING BEHAVIOUR/ INTEREST	20	3

Step 2

The chosen variables are included in a linear regression model where they play the role of independent variables. The dependent variables were the circulation and readership of “Sole 24 ore” .

The most important variables, chosen following the criteria of the portion explained of the independent variable, came from the set of 18 variables chosen at step 1.

The final result is summarised in the following Table 2.

Table 2

GROUPS	DESCRIPTION	NO.OF ORIGINAL VARIABLES	STEP 1 CHOSEN VARIABLES	STEP 2 CHOSEN VARIABLES
TOTAL	-	131	18	12
NO.1	DEMOGRAPHICS	16	2	1
NO.2	FAMILY INCOME	12	2	2
NO.3	INDUSTRY STRUCTURE	22	1	1
NO.4	LABOUR FORCES GEN. ACTIVITY SECTION	18	3	2
NO.5	LABOUR FORCES PROF.	7	2	1
NO.6	CONSUMPTION	24	3	-
NO.7	CULTURAL INTEREST	12	2	2
NO.8	READING BEHAVIOUR/ INTEREST	20	3	3

Provincial model

The objective of this section of the study was to find demo-social-economic variables correlated with the title variations measured on the time dimension, excluding the contribution of other components such as marketing activities and title's distribution. In fact the work hypothesis is to assume that the environmental conditions, depending on general newspaper market and title selling initiatives, remain exactly the same or change in a uniform way for provinces, in the course of the time period analysed.

The study attempted to correlate the penetration variations inside provinces, in terms of readership and distribution, with a reduced number of variables related to the same information groups identified for the regional analysis.

Step 1

The goal of this section is to reduce the variables to an ideal maximum number strictly less than ten. We applied the same method used for regional analysis to the provincial analysis. The way to choose the best subset was carried out using a multivariate procedure consisting of clustering variables. We used the procedure as a variable-reduction method, applying exactly the same technique described in Step 1 of regional analysis. The same considerations proposed for regions can be applied to the provincial part of the study.

The disjoint groups built from the initial set of 23 variables are described below in Table 3.

Table 3

GROUPS	DESCRIPTION	NO.OF ORIGINAL VARIABLES	STEP 1 CHOSEN VARIABLES
TOTAL	-	23	5
NO.1	DEMOGRAPHICS	4	-
NO.2	FAMILY STANDARD OF LIVING	5	3
NO.3	BUSINESS AND MICRO ECONOMY	4	-
NO.4	PUBLIC SERVICES EFFICIENCY	4	-
NO.5	TIME SPENT FOR PERSONAL INTEREST (entertainment,sport, culture)	6	2

Step 2

The chosen variables are included in a linear regression model where they play the role of independent variables. The model used is based on the methodology developed for analysis of panel data. The presence of cross sections of provinces observed over time, allowed us to construct and test more realistic behavioural models that could not be identified using only a cross section or a single time series data set. *****

The provinces series examined refers to a 2-years period in which the most important local campaigns took place.

The provinces data represent a longitudinal, or panel, data set that follows the provinces over time, and thus provides multiple observations on each province over time.

Besides the advantage that panel data allow us to construct and test more complicated behavioural models than purely cross-sectional or time-series data, the use of panel data also provides a means of resolving or reducing the magnitude of a key problem that often arises in empirical studies, namely, the often-heard assertion that the real reason one finds (or does not find) certain effects is because of omitted (wrongly measured, not observed) variables that are correlated with explanatory variables . By using information on both the intertemporal dynamics and the individuality of the entities being investigated, one is better able to control,in a more natural way, the effects of missing or unobserved variables.

For example, we can consider a simple model postulated as

$$y_{it} = \alpha_i^* + \beta_i x_{it} + u_{it}$$

$$i = 1, \dots, N$$

$$t = 1, \dots, N$$

where x is a scalar exogenous variable and u_{it} is the error term with mean zero and constant variance σ_u^2 . The parameters α_i^* e β_i may be different for different cross-sectional units, although they stay constant over time.

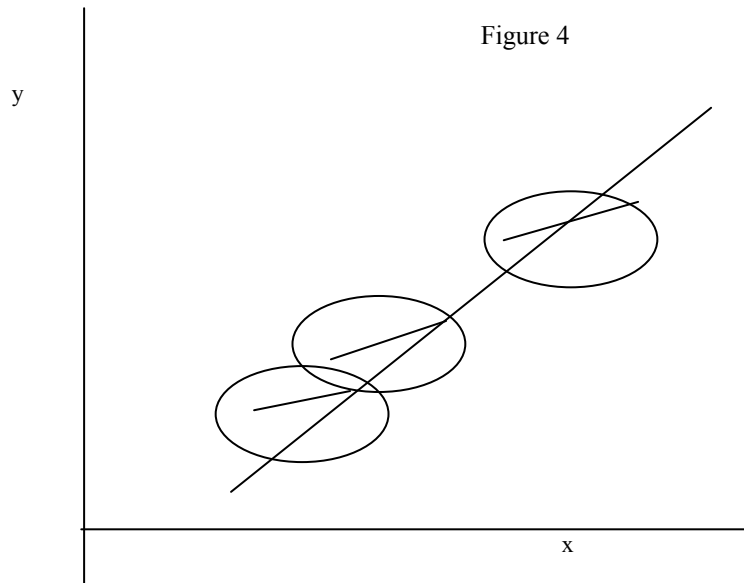
Following this assumption, a variety of sampling distributions may occur. Such sampling distributions can seriously mislead the least squares regression of y_{it} on x_{it} when all NT observations are used to estimate the model

$$y_{it} = \alpha_i^* + \beta x_{it} + u_{it}$$

$$i = 1, \dots, N$$

$$t = 1, \dots, N$$

For instance the data can present heterogeneous intercept ($\alpha_i^* = \alpha_j^*$) and homogeneous slope ($\beta_i = \beta_j$) . A possible circumstance is showed in the Figure 4.



The model chosen postulates homogeneous slope and heterogeneous intercepts depending on regions where the province is found .

References

- Box,G.E.P., and Tiao G.C.(1973), *Bayesian inference in Statistical Analysis*, Wiley Classic Library edition published 1992, New York:John Wiley & Sons,Inc.
- Biswas S. (1991), *Topics in Statistical Methodology*, New York: John Wiley & Sons, Inc.
- Carrol,R.J. and Ruppert,D.(1988),*Transformation and Weighting in Regression*,London: Chapman and Hall.
- Diggle,P.J., Liang,K.Y., and Zeger,S.L. (1994), *Analysis of Longitudinal Data*, Oxford: Clarendon Press
- Everitt B.S. and Dunn G. (1991), *Applied Multivariate Analysis*,Edward Arnold, London
- Fuller,W.a. and Battese, *Introduction to Statistical Time series*, New York: John Wiley & Sons, Inc.
- Hamburg M(1970)., *Statistical Analysis for Decision Making*, New York: Harcourt, Brace & World, Inc.
- Hsiao C., *Analysis of panel data*, Econometric Society Monographs
- SAS/STAT[®] Software, Changes and Enhancements*, Sas Institute Inc.
- SAS/STAT[®] Software, User's Guide*, Sas Institute Inc

