LUND UNIVERSITY
Campus Helsingborg

Generalized Maximum Entropy estimation method for studying the effect of the Management's Factors on the Enterprise Performances

Enrico Ciavolino, Researcher of Statistics, University of Salento,
Department of Philosophy & Social Science, Palazzo Parlangeli,
Via Stampacchia n.1, 73100, Lecce, Italy, enrico.ciavolino@ateneo.unile.it

Jens J. Dahlgaard, Division of Quality Technology,
Linkoping University, Linkoping, Sweden, jenda@ikp.liu.se

Keywords: Generalized Maximum Entropy, Leadership, Human Resources, Strategic Planning, Performance

Category: Research Paper

1. Introduction

The aim of this paper is to show the results of a survey that the University Consortium in Engineering for Quality and Innovation has led. A sample of Italian manufacturing companies was selected in order to verify the abilities to manage the human resources effectively, the spreading level of an effective and aware Leadership and the ability of strategic planning according to of a correct identification of the objectives. Moreover an alternative estimation approach, based on the maximization of the entropy function, is presented for estimating the parameters of the model presented. The data collected was analyzed with a multidimensional statistical method based on a Generalized Maximum Entropy (GME) estimation approach. This method is widely used in linear modelling and is presented here by considering a new algorithm for the estimation of the parameters of the model. Moreover, the data are analyzed by the Partial Least Squares regression (PLS), for a comparative study with GME.

2. The Data Collected and the Model for Evaluating the Management's Factor

Data were collected from a sample of Italian manufacturing companies and represents the answers of 120 enterprises. The sample is based on an analysis of the Italian manufacturing structure, where the sector is characterized by mostly small and medium enterprises (93%) and the remaining part by enterprises with more than 100 employees. From an economic point of view the big enterprises cover 50% of the production of the whole sector and employ 40% of the workforce.
In order to respect the structure of the Italian manufacturing sector the sample was selected as a weighted random sample where each company were assigned sampling weights according to company size. Based on 120 selected companies, the percentage of the enterprises selected follows the following proportions:

- 42%, with less than 50 employees;
- 14%, between 50 and 100 employees;
- 44%, with more than 100 employees.

To measure the level of the impact of management factors on the enterprises' performances, a questionnaire subdivided into four evaluation areas was used (see Table 1). Questionnaire data related to 35 statements were gathered through telephone interviews with the leaders of the selected companies. Respondents were asked to evaluate each statement on an ordinal scale with variation from 1 (disagree) to 2 (neither disagree nor agree) to 3 (agree). Table 1 shows all the variables used formulated as positive statements and LV means Latent Variables.

LV	MANIFEST VARIABLES	LV	MANIFEST VARIABLES
1. Performance	1. The market dimension has been improved during the last three years. 2. The revenue has increased during the last three years. 3. The percentage of Profit has increased 4. The percentage of Return Of Investment has increased during the last three years. 5 The total trend of the company's performance has improved.	*3. Human Resource*	1. The personal's careers are based on specific plans. 2. The company has invested in Research & Development 3. Job satisfaction is being evaluated. 4. Personal skill is being evaluated. 5. The merits are recognised and rewarded. 6. Work groups are used for specific themes 7. The employees have decision autonomy. 8. The employees identify themselves with the companies 9. Middle management has decision autonomy
2. Leadership	1. The reference values are well defined. 2. The leadership styles of governance depend on employee characteristics. 3. Management is open when communicating with employees 4. Management participates in formative events. 5. Management is involved in setting employee rewards 6. Management evaluates its leadership style compared with other company managers. 7. Management listen to considerations from employees 8. Management promotes programs for improving the Society and the Environment 9. Management involves the employees in setting objectives. 10. Management has negotiation capacity in critical situations.	*4. Strategic Planning*	1. Systematic analyses are made for customer expectations and market potentiality 2. Performance indexes are used for medium and long term plans. 3. The strategies consider competitor analysis. 4. Medium and long term plans are used for resource allocation. 5. The strategies are periodically re-evaluated. 6. A structured process defines the objectives and their diffusion. 7.The various operative groups are conformed to the main objectives 8. Each employee knows his objectives and the results 9. The employees are involved in the definition of objective resources. 10. New planning documents are developed for new projects. 11. Documents for the annual operative planning are developed

Table 1 – The Variables used in the Management Factor Model

A short description of the latent variables may help to make the discussion easy. For ***Leadership*** the statements (the manifest variables) selected are highlighting the figure of the leader, managerial capability in long term planning, orientation towards innovation, increasing the value of own collaborators and having a good relationship with the stakeholders. In the ***Human Resource*** area the necessity to have suitable employees that are skilful to bring the enterprise towards excellence has been the focus. The statements are about promotions, re-conversions, careers, training, and recognition of improvements. ***Strategic Planning*** may be the back bone for excellence. The focus is to measure and analyse if managers jointly work with all members of the enterprise through planning, doing and follow up activities driven by management, using systematic methodologies to support and evaluate decisions taken. ***Performance*** is the objective measure of the company health where statements about market dimension and economic indexes such as Return of Investment or profit level measure the efficiency of the company.

The Management Factor Model is shown as a path diagram in figure 1. The latent variables are estimated by using the average of the manifest variables belonging to the respective latent variable. The arrows in the path diagram show relationships between the variables, and these relationships are estimated by a series of entropy regression analyses. The statistical method used to estimate the measure of the relationship is known as *Generalized Maximum Entropy* (GME).

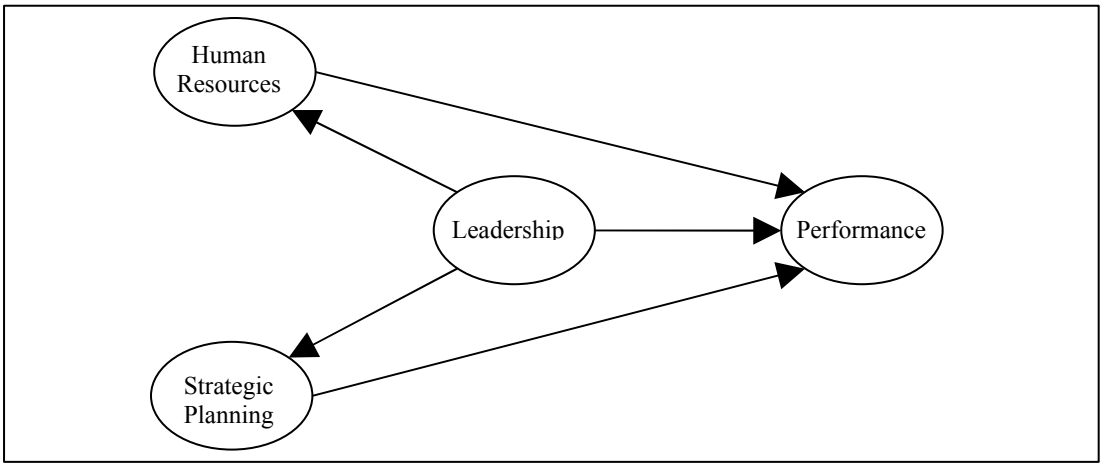

Figure 1 – The Management Factor Model

The model considers as dependent variables (endogenous variables) "*Performance*", "*Human Resources*", and "*Strategic Planning*" and as independent variable (exogenous variable) "*Leadership*". As we can see we have more then one dependent variable so for estimating the GME model we need to specify one equation for each dependent variable, as reported in the following equations system where the *beta coefficients* represents the linear impact of each independent variable.

$$\begin{cases} \text{Performance} = \beta_{1,1} \cdot \text{Leadership} + \beta_{1,2} \cdot \text{Human\_Resoruce} + \beta_{1,3} \cdot \text{Strategic\_Planning} \\ \text{Human\_Resoruce} = \beta_{2,1} \cdot \text{Leadership} \\ \text{Strategic\_Planning} = \beta_{3,1} \cdot \text{Leadership} \end{cases} \quad (1)$$

The equation system is an analytical representation of the path model.

3. The Entropy Theory

The maximum entropy presented here is identified by the *Shannon's entropy function* (1948), which is an axiomatic method to measure the uncertainty (state of knowledge) of a collection of events. Golan *et al* (1996) developed a GME procedure for general linear econometric models in order to permit estimation by the entropy principle when the underling model is incompletely known and the data are limited, partial or incomplete (ill posed problems).

3.1 Shannon's Entropy

Letting X be a random variable with possible outcome x_i, i=1..n, with probability p_i such that $\Sigma p_i=1$, a *global uncertainty measure (global state of knowledge)* of a collections of events is defined by *Shannon's Entropy function* (1948) considering an axiomatic method, based on a unique function:

$$H(X) = -k\sum_{i=1}^{n} p_i \cdot \log(p_i) \quad (2)$$

Where: k is a constant usually equal to 1, $0 \cdot \ln(0)=0$, and $\{\Sigma_i\, p_i=1\}$.
The quantity $\{-\log(p_i)\}$ is called *self-information* of the event x_i. The average of the self-information is defined as the entropy, so the unique function H(P) is called entropy, Shannon's Entropy, or Entropy of Information.

3.2 The Generalized Maximum Entropy Estimation Method
Golan *et al.* proposed an alternative method for parameter estimation of regression models, in case of ill-posed problems, as an extension of the entropy measure introduced by Shannon and generalization of the Maximum Entropy Principle developed by Jaynes (1957, 1968).
The method, which is called Generalized Maximum Entropy (GME), is based on the re-parameterization and re-formulation of a *general linear model* $\mathbf{y}=\mathbf{X\beta}+\boldsymbol{\varepsilon}$ in order to estimate the parameters inside the framework of the MEP, by the following equation:

$$\mathbf{y}_{n,1} = \mathbf{X}_{n,m}\boldsymbol{\beta}_{m,1} + \boldsymbol{\varepsilon}_{n,1} = \mathbf{Z}_{n,m}\cdot \mathbf{p}_{m,M}\cdot \mathbf{V}_{m,1} + \mathbf{w}_{n,N}\cdot \mathbf{V}_{N,1} \quad (3)$$

The matrices **Z** and **V** are diagonal and the generic element is represented by the vectors $\mathbf{z}_k' = [-c \quad -c/2 \quad 0 \quad c/2 \quad c]$ and $\mathbf{v}_k' = [-c \quad -c/2 \quad 0 \quad c/2 \quad c]$.
These vectors define the support variables called fixed points, usually of five elements (M=N=5), equally distributed around zero.
The super vectors **p** and **w** associated are probabilities and have to be estimated by maximization of the Shannon entropy function $H(P,W) = -\mathbf{p}'\cdot\ln\mathbf{p} - \mathbf{w}'\cdot\ln\mathbf{w}$, subjected to some normalization and consistency constraints. The steps for the GME algorithm are reported in the following table 2.

---

1. **Re-parameterize** the unknown parameters and the disturbance terms as a convex combination *of expected value of a discrete random variable*;
2. **Re-formulate the model** with the new re-parameterization as the data constraint;
3. **Define the GME problem as non-linear programming** problem in the following form:

   Objective Function = Shannon's Entropy Function
   1. The consistency constraints, which represents the new formulation of the model;
   2. The normalization Constraints.
4. **Solve the non-linear programming** by using numerical method

---

Table 2 – The Generalized Maximum Entropy Algorithm

4. Partial Least Squares Regression

PLS generalizes and combines features from principal component analysis (PCA) and multiple regression. Considering a *linear regression model* $\mathbf{y}=\mathbf{X\beta}+\boldsymbol{\varepsilon}$, using *X* original variables, PLS extracts independent latent variables **T**, from the original set of *m* variables. The regression vector is calculated from these latent variables, hence

overcoming difficulties in ordinary least squares such as multicollinearity. The PLS can be formalized by the following equation:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_{PLS} + \mathbf{E} \qquad (4)$$

Where $\boldsymbol{\beta}_{PLS}$ are the PLS regression coefficients equal to $\mathbf{w}^{*}\mathbf{C}^{T}$, respectively the covariance in term of the original variables, and the regression coefficients of the **y** on the **T** latent variables.

5. Analysis of Results

Both GME and PLS methods are used for estimating the relationships (path coefficients or impact scores) and significance, reported by the *Path Diagrams* in figures 2 and 3. Figure 2 shows the GME estimated parameters, and figure 3 shows the PLS parameters.

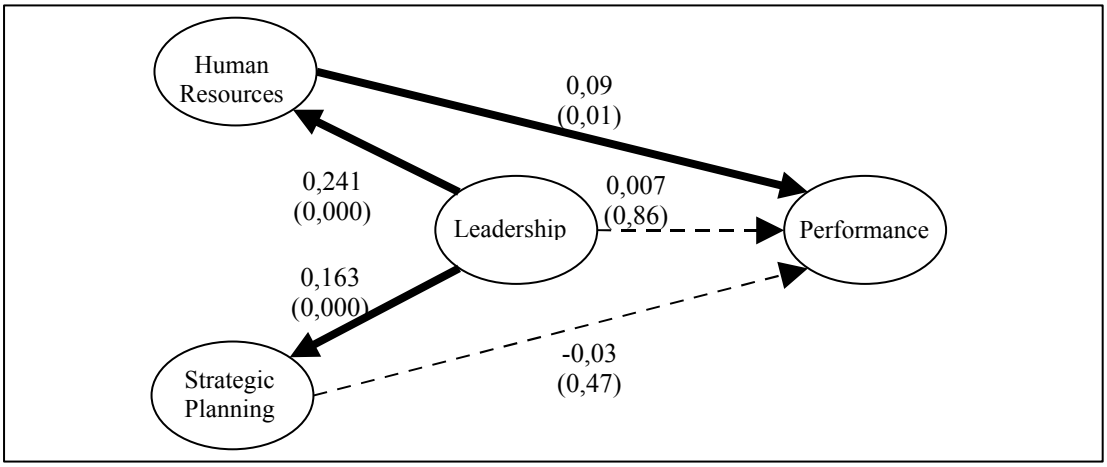

Figure 2 – Results of the GME estimation

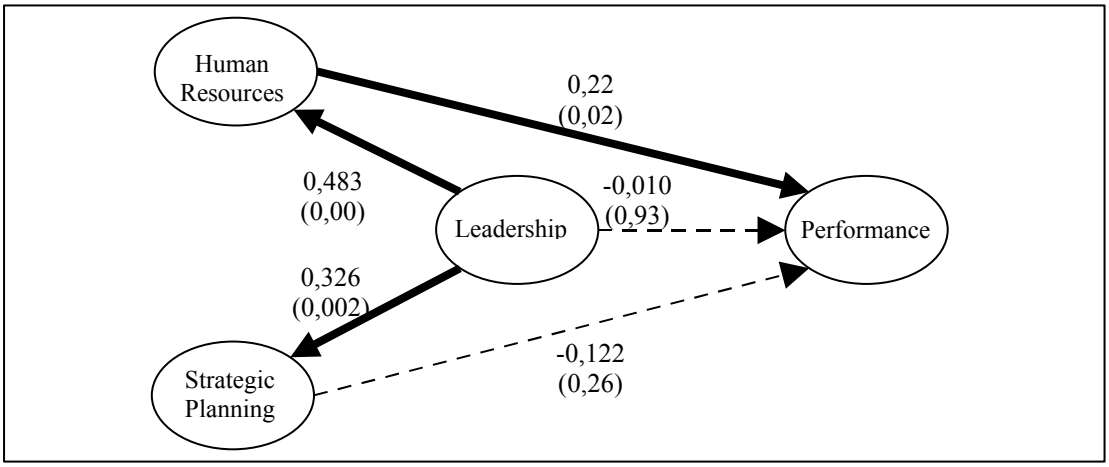

Figure 3 – Results of the PLS estimation

The estimated coefficients are reported near the line of the relationships, and significant relationships have been highlighted by a bold line, and non significant relationships are shown by non continuous lines. The numbers in the brackets are the p-values. The significance of the variable is calculated via bootstrap re-sampling, considering 100 of dimension 120.

It is seen that there are differences in the size of the impact scores but the two methods show the same results with respect to what are the significant relationships. The statistical analyses showed a little improvement in terms of Mean Squared Error (MSE) for the GME with respect the PLS. The results are reported in the table 3, but a deep comparative study has been conducted (Ciavolino *et. al.*, 2006, Ciavolino, 2007), showing an over-performing of the GME on the PLS in case of strongly correlated variables and more general in cases of the ill-posed problems.

	Mean Squared Error (GME)	Mean Squared Error (PLS)
Performance	0,00847	0,00882
Human Resource	0,00705	0,00656
Strategic Planning	0,00692	0,00766

Table 3 – The MSE comparison between GME and PLS.

The estimated coefficients suggest that *Human Resource* is the only variable having a significant and positive direct impact on *Performance*. The variables *Leadership* and *Strategic Planning* had no significant direct impact on *Performance*, but the impact of Leadership on *Human Resources* and *Strategic Planning* were significant. Table 4 reports the estimated values of the latent variables, obtained as the average value of the manifest variables.

L.V.	Performance	Leadership	Human Resource	St. Planning
Values	2,275	2,796	2,583	2,633

Table 4 – Estimated Values of the Latent Variables.

The exogenous latent variables were categorized into two groups of agreement level, where the variables in the first group have a high level of agreement with the positive statements, and the variables in the second group have a low level of agreement. The level of agreement is in this case recognised as high when the agreement value is 2.60 or higher and low when the level of agreement is below 2.60.
By using this categorization it became possible to construct an interventions matrix, by combining the information of the path coefficients and the average agreement reached, in order to group the variables according to importance and agreement, as reported in the following table 5.

		Agreement	
		Low	**High**
Importance	**Low**		Strategic Planning
	High	Human Resource	Leadership

Table 5 – Interventions Matrix

The "message" of the interventions matrix is that improvements should be prioritised to variables where importance is high and agreement low. Hence the general message of table 5 is that Italian industrial companies are relatively weak on *Human Resources* and they should first of all improve the *Human Resource* aspects. This message is based on the fact that *Human Resources* were the only variable with a significant impact on *Performance*, and the level of agreement reached was smaller than for the other management factors of the model.
The following table 6 shows the average scores for each manifest variable. Regarding *Human Resource* aspects the variables with the lowest scores were "job satisfaction is

being evaluated" and "the personals' careers are based on specific plans". For *Leadership* the variables with the lowest scores were "management promotes programs for improving the Society and the Environment" and "management evaluates its leadership style compared with other company managers". Under *Strategic Planning* we have the lowest scores related to "a structured process defines the objectives and their diffusion" and "new planning documents are developed for new projects". Regarding *Performance* we can see that most or the average scores were relatively low indicating that there has not been a positive trend in Italian industrial companies' performance during the last 3 years.

Performance	**Av**	**Human Resource**	**Av**
1. The market dimension has been improved during the last three years.	2,169	1. The personal's careers are based on specific plans.	2,349
2. The revenue has increased during the last three years.	2,119	2. The company has invested in Research & Development	2,649
3. The percentage of Profit has increased	2,777	3. Job satisfaction is being evaluated.	2,033
4. The percentage of Return Of Investment has increased during the last three years.	2,284	4. Personal skill is being evaluated.	2,674
5 The total trend of the company's performance has improved.	2,027	5. The merits are recognised and rewarded.	2,583
		6. Work groups are used for specific themes	2,358
		7. The employees have decision autonomy.	2,983
		8. The employees identify themselves with the companies	2,758
		9. Middle management has decision autonomy	2,866
Leadership	**Av**	**Strategic Planning**	**Av**
1. The reference values are well defined.	2,976	1. Systematic analyses are made for customer expectations and market potentiality	2,678
2. The leadership styles of governance depend on employee characteristics.	2,872	2. Performance indexes are used for medium and long term plans.	2,512
3. Management is open when communicating with employees	2,990	3. The strategies consider competitor analysis.	2,845
4. Management participates in formative events.	2,580	4. Medium and long term plans are used for resource allocation.	2,612
5. Management is involved in setting employee rewards	2,672	5. The strategies are periodically re-evaluated.	2,837
6. Management evaluates its leadership style compared with other company managers.	2,588	6. A structured process defines the objectives and their diffusion.	2,370
7. Management listen to considerations from employees	2,863	7.The various operative groups are conformed to the main objectives	2,978
8. Management promotes programs for improving the Society and the Environment	2,547	8. Each employee knows his objectives and the results	2,695
9. Management involves the employees in setting objectives.	2,780	9. The employees are involved in the definition of objective resources.	2,545
10. Management has negotiation capacity in critical situations.	2,995	10. New planning documents are developed for new projects.	2,373
		11. Documents for the annual operative planning are developed	2,520

Table 6 – Average Scores of the Manifest Variables

5. Concluding Remarks

This paper has focused on a study of the relationships between ***Leadership***, ***Human Resources*** and ***Strategic Planning***, and the impact of these latent variables on ***Performance.*** The data analysed were collected by telephone interviews with leaders from 120 Italian industrial companies.

The analysis of the model has been conducted by considering also the PLS estimation method. It has been shown how the two methods, in this case, showed the same overall results, but also showed a little improvement of the GME on the PLS in terms of MSE. Simulation studies have been conducted showing that in case of ill-posed problems, and especially for small samples, the GME outperforms the PLS estimation method in term of MSE.

The analysis of the survey data showed interesting and unexpected results regarding the non significant relationship between Leadership and Performance. What does that mean? Doesn't good Leadership influence enterprises' performance?

The answer to this question is that even if there is no direct relationship between the two variables, the effect of Leadership is obtained by an indirect relationship through Human Resources. The *combination* of Leadership and Human Resources has hence been identified as the variables which have the highest impact on the performance of Italian industrial companies. This result is totally in accordance with the findings and suggestions by Dahlgaard & Dahlgaard (2003, 2006) in their "4P" model for business Excellence. The message from this model is that a general strategy for improving performance is to improve "the 4P" – People, Partnerships, Processes, and Products – in this order. And because the foundation of "the 4P" is Leadership improvements always starts with Leadership. Without Leadership no sustainable improvements, and improvements of "the 4P" go through Leadership. The statistical analyses shown in this paper support this strategy.

Another interesting and unexpected result was that there was no significant impact of Strategic Planning on Performance. It seems that the leaders of Italian industrial companies have not understood that good strategic planning is a necessary condition for achieving excellence. It seems they have not understood what excellent companies have learned during the last decades that good strategic planning with effective policy deployment is *the backbone* of Total Quality Management and Business Excellence.

So another improvement area, which was not highlighted by the interventions matrix, is in fact Strategic Planning. This area should have the highest priority and responsibility of any top management team and the focus should include how to establish a strong relationship between strategic planning and performance. If statistical data analyses, as shown in this report, show now correlations between strategic planning and performance, then we have a strong indication that something is wrong. It is not enough that Leadership is doing Strategic Planning – Leadership is also about studying and follow up on results in order to assure impacts on performance. This link seems to be missing in Italian industrial companies (as indicated in figures 2 and 3).

REFERENCES

Al-Nasser A. D. (2003). Customer Satisfaction Measurement Models: Generalized Maximum Entropy Approach. *Pak Journal of Statistics*, 19(2), 213–226;

Ciavolino E. (2007). The Entropy Theory for evaluating the Job Satisfaction, *GFKL 2007*, Freiburg, march 2007;

Ciavolino E., Dahlgaard J.J., ECSI - Customer Satisfaction Modelling and Analysis: a case study, in press on Total Quality Management & Business Excellence, Article ID: 223946 (TQM & BE 108/05);

Ciavolino E., Al Nasser A.D., D'Ambra A. (2006). The Generalized Maximum Entropy Estimation method for the Structural Equation Models, *GFKL 2006*, Berlino, marzo 2006;

Dahlgaard, J. J. & Park-Dahlgaard, S. M. (2003), "The "4P" Quality Strategy for Breakthrough and Sustainable Development", *the 7th QMOD Conference*, Monterrey, Mexico

Dahlgaard, J. J., Park-Dahlgaard, S. M. & Martensen, Anne (2006), Measuring and Diagnosing Innovation Excellence – Simple contra advanced approaches - A Danish Study, Proceedings of the 9th QMOD Conference, Liverpool, UK, August 2006

Park-Dahlgaard, S.M. & Dahlgaard, J.J. (2006), In Search of Excellence – Past, Present and Future, in: H. Schnauber (ed.), Kreativ und Konsequent – Herbert Masing, Hanser Verlag, München, Wien

Jaynes E.T. (1957). Information Theory and Statistical Mechanics, *The Physical Review* 106 (4), 620-630, May 15, 1957;

Jaynes, E. T. (1968). Prior Probabilities, *IEEE Transactions On Systems Science and Cybernetics*, vol. sec-4, no. 3, 227-241;

Golan A., Judge G. & Karp L. (1996). A maximum entropy approach to estimation and inference in dynamic models or counting fish in the sea using maximum entropy, *Journal of Economic Dynamics and Control*, 20, 559-582;

Golan, A, G George & D Miller (1996). Maximum Entropy Econometrics, Wiley, London;

Paris, Q. (2001): Multicollinearity and maximum entropy estimators. *Economics Bulletin*, 3(11), 1–9.

Shannon C. E. (1948). A mathematical Theory of Communications, *Bell System Technical Journal*, 27, 379–423.

Wold S., Sjostrom M., Eriksson L., (2001), PLS-regression: a basic tool of chemometrics, *Chemometrics and Intelligent Laboratory Systems*, vol. 58, 109-130;