3. THE RELATION BETWEEN PREDICTABILITY AND COMPLEXITY: DOMESTIC AND PUBLIC CONSUMPTION IN THE ROMANIAN ECONOMY[1]

Cornelia SCUTARU[*2]
Corina SÂMAN[**]
Cristian STĂNICĂ[***]

Abstract

We continue[3] *with the problem of the relation between predictability and complexity in the Romanian economy, analyzing other two components of GDP: domestic consumption and public consumption. The basic idea of this work is that the unpredictability of a system gives a measure of its complexity, so that in order to predict a future state of a complex system one must find the system structure explained by some simpler components that can be predicted. The complexity of the economic system is reflected in the synthetic macroeconomic indicators (GDP and its components). We find the principal components of the macroeconomic variables as a preprocessing step and model them as linear combinations of some simpler non observable predictable variables; we have constructed empirical models for domestic and public consumption; it was shown that these models are sufficiently accurate to predict for two or three periods ahead.*

[1] Prepared under the Programme: "Modelling and assessing the impact of direct investments on national and international labour market and macroeconomic evolutions of Romania", Contract MEC 91-052/ 10 September 2007.

[*]Senior researcher, Ph.D., Institute for Economic Forecasting, Romanian Academy, Bucharest; e-mail: corneliascutaru@yahoo.com.

[2] Member of the Academy of Romanian Scientists. Bucharest;

** Researcher, Ph.D. Student, Institute for Economic Forecasting, Romanian Academy, Bucharest; e-mail: csaman@ines.ro .

*** Researcher, Ph.D., Institute for Economic Forecasting, Romanian Academy, Bucharest; e-mail: stanica_cris@yahoo.com .

[3] We have treated this problem for the case of gross fixed capital formation in “Predictability and Complexity in Macroeconomics. The case of Gross Fixed Capital Formation in the Romanian Economy”, Romanian Journal of Economic Forecasting, vol IX, no. 4/2008, pp. 196-206. A communication was presented at the 5th European Conference on Complex Systems, Jerusalem, September 14-19, 2008.

Keywords: complexity, predictability, consumption, complexity pursuit, time series forecasting
JEL Classification: E20, E27, C22, C51, C63

1. Introduction

There are many and diverse ways of measuring the complexity, from the complexity in Kolmogorov sense, based on entropy, to measures of interdependency between the components of the analysed system [Solomon, S. (1998), Shnerb, N., Bettelheim, Y., Louzoum, Y., Agam, O., Solomon, S., (2001), Stone, James V., (2004), Lopes, J. Carlos, Dias, J., Fereira do Amaral, J., (2007)]. A simple and intuitive measure of complexity is the inverse of the predictability (Stone, 2004). Starting from the idea that a complex phenomenon is the result of the cumulative effects of many simple ones we first decompose it into its principal components, which are statistically uncorrelated, and find their sources of complexity as some latent variables.

We consider the economy as a complex phenomenon, which can be described by the GDP evolution and its components: gross capital formation, consumption (public and domestic), export and import (the GDP decomposition by expenditure). Each of these components is determined by specific economic processes described by different variables.

In Scutaru, Sâman, Stănică (2008), we have constructed an empirical model for the case of the Gross Fixed Capital Formation, considering the group of macroeconomic indicators that influence it, and computed the principal components, which were decomposed into some factors which were common sources of evolution of the entire group and also predictable. Evaluating the principal components is not new for the Romanian economy: in Klein, Roudoi, Eskin, Albu, Stănică, (2004) and Klein, Roudoi, Eskin, Albu, Stănică, Nicolae, Chilian, (2004), the GDP was evaluated by expenditure and production, in order to estimate a predictive model, using the principal component analysis, respectively, and in Klein, Roudoi, Eskin, Nicolae, (2004) the impact of oil price on GDP is studied using the same techniques.

In this paper, we construct such an empirical model for consumption (domestic and public and consumption), considering specific variables.

2. The Complexity Pursuit Algorithm

Complexity pursuit (Hyvärinen 2001, Hyvärinen and Oja, 2000) is a class of methods, which seek for minimally complex components of time series, and it derives from information theoretic measures of complexity. We use a complexity pursuit algorithm for separating minimally complex and, therefore, most predictable sources in time series that was recently presented by Stone (2004), which minimizes a measure of Kolmogorov complexity. The method assumes that any linear combination of variables is more complex than the simplest of them. This conjecture is the basis for separating mixtures into their sources by seeking the least complex signal obtained from the mixture (Stone, 2004).

We employ a vector space model for representing the GDP components[4]. Each group is formalized as data matrix X of size TxM containing the component vectors as its columns.

The data is first whitened by PCA, we denote by z(t) this preprocessed data, and b now corresponds to an estimate of a row of the inverse of the mixing matrix for whitened data.

$$X = B\mathbf{x}Z \qquad \textbf{(1)}$$

The new data matrix Z and its columns z_i, i=1,…, M are the inputs for the complexity pursuit algorithm. The time-structure of the GDP components, or the minimum complexity projections can be found by projecting Z onto the directions W = (w_1,…, w_M) given by the complexity pursuit algorithm described in the next paragraphs.

The data model assumes that the observations z(t) are linear combinations of some non-observable components s (t):

$$Z = A\mathbf{x}S \qquad \textbf{(2)}$$

where: Z = (z_1,…, z_M) is the vector of observed random variables, S = (s_1,…, s_M) is the vector of the predictable nonobservable components which we search for, and A is an unknown constant mixing matrix, also to be found.

A moving average model with long memory is assumed to model each component s_i = w_i^T Z; as a exponentially weighted sum of past values:

$$\tilde{\mathbf{s}}_i(t) = \sum_{k=0}^{t-1} \lambda^{n-k-1}(1-\lambda)\mathbf{s}_i(k) + \lambda^t \tilde{\mathbf{s}}_i(0) \qquad (3)$$

for some λ, $0 \le \lambda \le 1$.

The value of λ determines how much memory of past values of $s_i(t)$ is needed to forecast $\tilde{\mathbf{s}}_i(t)$. Large values for λ means a long memory process.

It could be said that the predicted value $\tilde{s}_i(t)$ is a sum of values measured up to time (t-1), so that recent values have a larger weighting than those in the distant past:

$$\tilde{\mathbf{s}}_i(t) = \lambda\,\tilde{\mathbf{s}}_i(t-1) + (1-\lambda)\,\mathbf{s}_i(t-1)\,. \qquad (4)$$

We found the vector w for each component s = w^T z, which minimizes its complexity, which in the same time maximize its predictability. The measure F of s predictability is defined (Stone, 2004) as a ratio of total variance of s=(s(t)), t=1…T by the measure of its “smoothness” calculated as a moving average. The complexity is measured in terms of temporal predictability so that lower complexity corresponds to higher predictability:

[4] What we presented in this paper is part from a larger study which explores training models of GDP components (population consumtion, goverment consumtion, gross capital formation, export, import) from the assessment of their predictability / complexity, using the same methodology illustrated here for gross fixed capital formation.

$$\mathrm{F}(\mathbf{w}_i, \mathbf{z}) = \ln \frac{\sum_t (\overline{\mathbf{s}_i} - \mathbf{s}_i(t))^2}{\sum_t (\tilde{\mathbf{s}}_i - \mathbf{s}_i(t))^2} = \ln \frac{V_i}{U_i} = \ln \frac{\mathbf{w}_i \overline{\mathbf{C}} \mathbf{w}_i^T}{\mathbf{w}_i \tilde{\mathbf{C}} \mathbf{w}_i^T} \quad (5)$$

where: $\overline{\mathbf{C}}$ and $\tilde{\mathbf{C}}$ are MxM matrix of covariance:

$$\overline{C}_{ij} = \sum_t (z_{ti} - \overline{z}_{ti})(z_{tj} - \overline{z}_{tj})$$
$$\tilde{C}_{ij} = \sum_t (z_{ti} - \tilde{z}_{ti})(z_{tj} - \tilde{z}_{tj}) \quad (6)$$

The gradient update of the weights w in order to extract the most predictable latent component s that maximize F is given by the next formula (Stone, 2004):

$$\mathbf{w}_i = \mathbf{w}_i + \eta \nabla_{\mathbf{w}_i} F$$
$$\nabla_{\mathbf{w}_i} F = \frac{2\mathbf{w}_i}{V_i} \overline{\mathbf{C}} - \frac{2\mathbf{w}_i}{U_i} \tilde{\mathbf{C}} \quad (7)$$

where: η is a small constant, the step uphill along the direction of the gradient.

To estimate all M projections w_i we can solve a generalized eigenproblem whose solutions are eigenvectors of the matrix $\tilde{\mathbf{C}}^{-1} * \overline{\mathbf{C}}$.

Thus, a matrix W = (w_1, …, w_M) is found which verifies:

$$S = ZxW \quad (8)$$

The algorithm to get multiple forecasts:

1. Equation (4) applied for time T +1 give the sources values for time T+1: $\tilde{\mathbf{s}}_i(\mathrm{T}+1) = \lambda \tilde{\mathbf{s}}_i(\mathrm{T}) + (1-\lambda)\mathbf{s}_i(\mathrm{T})$
2. Calculate Z = W^{-1}xS, where S=(s_1, …, s_M) and s_i=(s_i (t)), t=1…T+1.
3. If we want next forecast value, then T = T+1; go to pass 1.

This algorithm calculates several forecasts values starting from the same input data, without incorporating the new observations, so expect greater errors as they are propagated through steps T to the next.

3. Experimental Results on Domestic and Public Consumption in the Romanian Economy

Data from the statistics are compiled by the National Institute of Statistics, from the first quarter of 2000 until the second quarter of 2007; we are working with the index chain of variables expressed in real terms. The GDP components group were created by the expenditure decomposition method, for each component was built a group of variables that expressed by our opinion, the most powerful influences on the component.

3.1. Domestic consumption

The G1HC Group (domestic consumption) includes variables selected for proper assessment of the main components of the population's consumption: the share (ratio) of social expenditure in total expenditure of the general consolidated budget (RSE); consumer price index (total - and the IPCT and non-food goods - IPCANA); the exchange rate (ERE); the volume of retail trade (ICA); the volume of production machinery and electrical appliances (IMAEL); net taxes (NT), the volume of industrial production (IPI) and gross salary per economy, in real terms (SBREC).

The components are: HCPC1, HCPC2, HCPC3 and HCPC4, which correspond to the eigenvalues of 3.429, 2.597, 1.157 and 0.759; the proportion of variance explained by the first three main components is 80% and if we add the fourth we obtain 88% of variance explained. Analysis of the main components indicates that HCPC1 depends on: the weight (ratio) of social expenditure in total general consolidated budget expenditure (-0.449), the volume of retail trade (-0.51); production volume of electrical machinery and apparatus (-0.25); net taxes (-0.40); volume of industrial production (-0.43) and gross salary per economy in real terms (-0.34) and does not depends on: Consumer price index (total amount, 0.08 and the goods and nonfood, 0.09) and exchange rate (0.07).

Equations of the four main significant components are:

HCPC1 =-0.45*IRSE + 0.08*IIPCT + 0.07*IERE - 0.51*IICA - 0.25*IIMAEL - 0.39*INT + 0.08*IIPCANA - 0.43*IIPI-0.34*ISBREC

HCPC2 = 0.12*IRSE + 0.59*IIPCT + 0.51*IERE - 0.03*IICA + 0.14*IIMAEL + 0.03*INT + 0.59*IIPCANA + 0.08*IIPI + 0.05*ISBREC

HCPC3 = 0.14*IRSE - 0.24*IERE-0.24*IICA+0.54*IIMAEL-0.48*INT + 0.04*IIPCANA - 0.20*IIPI + 0.54*ISBREC

HCPC4 = -0.26*IRSE - 0.20*IIPCT + 0.26*IERE + 0.02*IICA + 0.63*IIMAEL - 0.19*INT - 0.13*IIPCANA + 0.39*IIPI - 0.47*ISBREC

Our experiments showed that choosing a moving average model:

$\tilde{\mathbf{s}}_i(t) = \lambda \tilde{\mathbf{s}}_i(t-1) + (1-\lambda)\mathbf{s}_i(t-1)$ was successful and that $\lambda = 0.85$ was the most suitable.

We calculated two types of measures for deviations from this model:

$$M1(\mathbf{s_i}) = \frac{\tilde{\mathbf{s}}_\mathbf{i}(T) - \mathbf{s_i}(T)}{\mathbf{s_i}(T)} * 100$$

$$M2(\mathbf{s_i}) = \frac{\tilde{\mathbf{s}}_\mathbf{i}(T) - \mathbf{s_i}(T)}{stdev(\mathbf{s_i})} \tag{9}$$

where: M1 represents the percentage error of the model and M2 represents the number of standard deviations away from real data.

The criterion adopted for analyzing the quality of prediction is presented in Table 1.

Table 1

M1 Criterion for the Evaluation of a Forecasting Model

M1 - Absolute value	Forecasts Classification
<10	high accuracy
10-20	good accuracy
20-50	reasonable accuracy
>50	unreliable

In order to evaluate the model, we divide the data into two sets: the set for training *Tr* with data from 2000:Q1 – 2006Q1, and the set for testing the model *Test* from period 2006:Q2 – 2007:Q2. This is a medium-term dynamic forecast.

We search for λ that minimize M(λ) on set *Tr*,

$$\text{where } M(\lambda) = \min \sum_{i=1}^{n} |M1(z_i)| * eigenval(i) \qquad (10),$$

z_i is the *i*th principal component and eigenval(i) is the *i*th eigenvalue of the correlation matrix for the variables. Some values for M1 an M are presented in Table 2 and in Figure 1.

In our analysis, we eliminate from the beginning the last two principal components because they are responsible for only 1% of the explained variance.

Table 2

Values of M1 and M for Principal Components of G1HC group

λ	HCPC1	HCPC2	HCPC3	HCPC4	HCPC5	HCPC6	HCPC7	$M(\lambda)$
0.01	23.38	2.2	-69.7	-179.42	370.13	15.89	-18.7	459.35
0.1	-15.47	-1.54	100.22	8.13	-28.25	2.19	33.97	200.37
0.2	-14.01	-1.32	84.3	-2.21	-22.64	2.12	31.87	168.75
0.3	-12.67	-1.15	68.24	-11.61	-17.33	3.57	29.91	150.40
0.4	-11.46	-1.03	53.25	-20.35	-12.55	6.04	28.03	133.61
0.5	-10.39	-0.95	39.39	-28.5	-8.47	9.1	26.24	118.76
0.6	-9.48	-0.89	26.91	-35.83	-5.1	12.34	24.66	105.84
0.7	-8.73	-0.78	15.95	-41.63	-2.23	15.08	23.51	94.15
0.85	-7.89	0.09	1.95	-47.39	1.66	14.53	22.77	77.07
0.9	-7.44	0.6	-3.27	-54.07	1.69	9.93	22.01	81.79
0.99	-3.21	3.86	-24.3	-136.52	-15.42	-29.19	9.42	170.76
1	-1.94	4.52	-29.15	-161.48	-21.22	-39.98	5.34	197.28

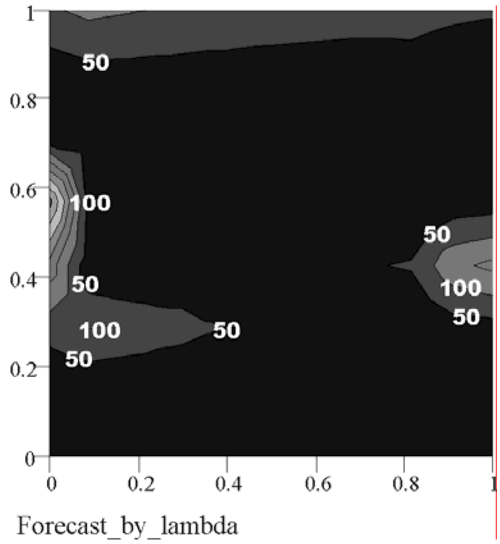
Source: Authors' calculations.

We select the model for $\lambda = 0.85$ so that the moving average model described by equation (3) and (4) is a process with long memory, assigning weight on the past values.

The identified predictable sources $\tilde{\mathbf{s}}_i(t)$ deviate by at most 1.33 standard deviations from $\mathbf{s}_i(t)$ and the forecasts have quite a reasonable accuracy.

Figure 1

Values of M1 for Principal Components of G1HC group



But the criteria for selecting the best model are the measure M for the principal components.

For $\lambda = 0.85$ we obtain high accuracy for the first three principal components (z_i, i = 1,3), which represent 80% of the explained variance and reasonable enough for the fourth principal component (percentage error is 48.78%) we select this best model for $\lambda = 0.85$.

Table 3

Values of M1 and M2 for predictable sources (s_i) and Principal Components (z_i) for $\lambda = 0.85$

$M1(s_i)$	$M2(s_i)$	$M1(z_i)$	$M2(z_i)$
1.91	0.74	-7.89	-1.15
1.15	-0.28	-0.09	-0.04
-21.56	-0.33	1.95	0.02
48.78	-1.33	-47.39	-0.61
12.02	-1.15	1.66	-0.05
-2.67	0.31	14.53	-0.31
-8.75	-0.85	22.77	-1.14
-2.15	-0.3	-72.04	-1.48
-24.16	0.92	57.98	-1.17

Source: Authors' calculations.

On the set used for testing, *Test*, we evaluate the model, calculating the M1 and M2 values for the principal components as described by the model.

The values of percentage error M1 for the medium-term dynamic forecasts are given in Table 4.

Table 4

Values of M1 for Principal Components from G1HC group in the period 2006:Q2 to 2007:Q2

	HCPC1	HCPC2	HCPC3	HCPC4	HCPC5	HCPC6	HCPC7
2006:Q2	-7.89	-0.09	1.95	-47.39	1.65	14.53	22.76
2006:Q3	1.88	2.00	1.14	182.62	-16.22	-27.46	-20.70
2006:Q4	-6.73	0.35	16.30	-143.11	-8.08	101.67	15.38
2007:Q1	16.42	2.52	-65.19	-82.60	337.54	112.60	-8.42
2007:Q2	-6.19	0.93	612.88	-379.67	-7.30	-29.94	-9.48

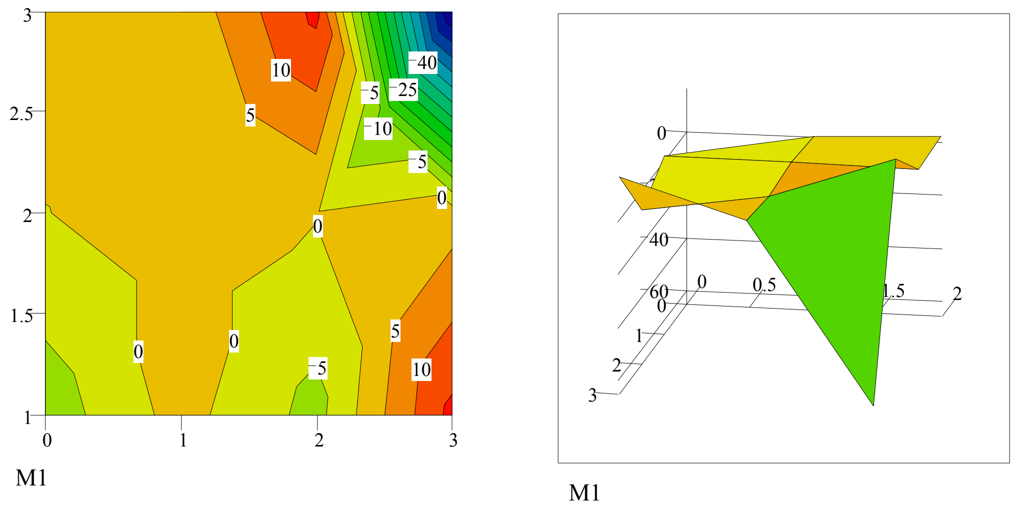
Source: Authors' calculations.

From the analysis of the error values and also based on the criteria presented in Table 1, one may say that the models successfully produced good accurate forecasts for the first three principal components for three periods ahead. When we apply the model without incorporates the new information available in the real data (observations) and use the model to forecast for more than one period ahead we must not expect good accurate results.

Figure 2 presents the M1 percentage error for the first three principal components on the forecasting horizon 2006:Q2–2007:Q1, where we have high and good accuracy for the forecasting model.

Figure 2

Graph of the M1- measure for the first 3 Principal Components of G1HC group



In order to get the real image of the goodness of the forecasting procedure, we must analyze the M2 measure in Table 5.

Table 5

Values of M2 for variables from G1HC group in the period 2006:Q2 to 2007:Q2

	HCPC1	HCPC2	HCPC3	HCPC4	HCPC5	HCPC6	HCPC7
2006:Q2	-1.15	-0.04	0.02	-0.61	-0.05	-0.31	1.14
2006:Q3	0.25	0.75	1.03	0.5	0.6	0.9	1.56
2006:Q4	-0.98	0.13	0.14	2.31	0.28	-1.26	0.83
2007:Q1	1.94	0.93	-1.07	-1.66	-2.49	-1.23	0.56
2007:Q2	-0.89	0.35	1.04	0.81	0.22	0.91	0.65

Source: Authors' calculations.

One may observe in Table 5 that the first three principal components deviate from the real value by at most 1.94 standard deviations. For HCPC3 on 2007:Q2, where we have the maximum on M1, the M2 measure is 1.04 and for HCPC1 on 2007:Q1, where we have the maximum on M2, the measure M1 shows good accuracy.

Thus, we propose that the evaluation of the accuracy of the model combines the two measures defined in (9), and to consider as principal criteria M1 as defined in Table 1 and to accept as reasonable accuracy even if this criteria is not fulfilled if M2 is approximately 1 in absolute value.

3.2. Public consumption

The G1GC Group (public consumption) includes variables selected for proper assessment of the main components of government consumption: the share (ratio) of incomes and expenditures of the general consolidated budget (RSB), which represent budget deficit; the volume of construction activity (ICONSTR); the volume of industrial production (IPI); the volume of retail trade (ICA); the exchange rate (ERE); broad money M2 in real terms (M2R); the degree of coverage of imports by exports (GXM), representing the net trade balance; the consumer price index (total – IPCT); the price index of industrial production (IPPI); the gross salary in administration, health, education, in real terms (SBRADM); the share (ratio) of expenditure with social assistance in total expenditure of the general consolidated budget (RSE), representing the level of social assistance to the population.

The components are: GCPC1, GCPC2, GCPC3, which correspond to the eigenvalues of 4.606, 2.628 and 1.167; the variance explained by first three main components is 81%. Analysis of the main components indicates that GCPC1 cumulates the influence of the net general consolidated budget and the social assistance to the population, the volume of construction activity, the industrial production, the volume of retail trade, broad money and trade balance. The second principal component, GCPC2, mostly includes the influence of the exchange rate and the domestic prices: the consumer price index and the price index of industrial production. The third principal component depends on the gross salary in administration, health, and education, in real terms, on the budget deficit and on the social assistance to the population.

The equations of the four main significant components are:

GCPC1=-0.35*IRSB+0.42*IICONSTR+0.39*IIPI+0.43*IICA+0.02*IERE+0.21*IM2R-0.39*IGXM+0.02*IIPPII-0.04*ISBRADM+0.39*IRSE

GCPC2=0.08*IRSB+0.02*IICONSTR+0.11*IICA-0.54*IERE-0.15*IM2R-0.06*IGXM-0.56*IIPCT-0.58*IIPPII+0.01*ISBRADM+0.02*IRSE

GCPC3=0.35*IRSB+0.19*IICONSTR+0.27*IIPI+0.11*IICA+0.08*IERE+0.22*IM2R++0.19*IGXM-0.15*IIPCT+0.05*IIPPII-0.71*ISBRADM-0.34*IRSE

Table 6

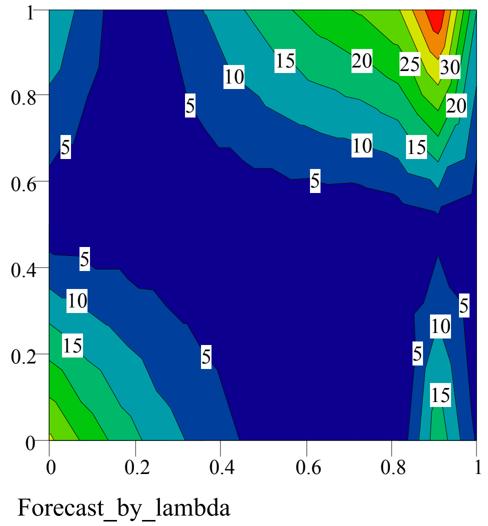
Values of M1 and M for Principal Components of G1GC group

λ	GCPC1	GCPC 2	GCPC 3	$M(\lambda)$
0.01	30.95	-1.39	14.69	170.88
0.12	23.21	-1.20	7.57	122.78
0.22	17.04	-1.02	1.08	83.01
0.32	11.98	-0.85	-4.86	65.58
0.42	7.90	-0.73	-10.18	55.37
0.52	4.64	-0.67	-14.90	48.12
0.62	2.02	-0.66	-19.11	43.09
0.72	-0.10	-0.60	-22.74	40.19
0.82	-1.65	-0.22	-25.58	51.06
0.92	0.82	1.18	-29.76	56.77
0.98	18.43	2.88	-44.67	167.36
0.52	4.50	-0.67	-15.12	47.82

Source: Authors' calculations.

Figure 3

Values of M1 for Principal Components of G1GC group



The identified predictable sources $\tilde{\mathbf{s}}_i(t)$ deviate by at most 0.23 standard deviations from $\mathbf{s}_i(t)$ and the forecasts have high accuracy.

But the criteria for selecting the best model is the measure M for the principal components, so we select $\lambda = 0.72$ which gives high accuracy for the first three principal components (z_i, i = 1,3), which represent 81% of the explained variance.

Table 7

Values of M1 and M2 for predictable sources (s_i) and Principal Components (z_i) for $\lambda = 0.72$

$M1(s_i)$	$M2(s_i)$	$M1(z_i)$	$M2(z_i)$
0.04	0.02	-0.1	-1.46
0.42	0.12	-0.6	0.29
0.9	0.23	-22.74	1.16
379.48	1.19	-523.29	0.55
309.97	0.97	-1.98	-0.62
19.11	-0.6	-4.85	-1
0.16	0.03	2.91	1.66

Source: Authors' calculations.

On the set used for training, *Tr*, we evaluate the model, calculating the M1 and M2 values for the principal components as described by the model.

The values of percentage error M1 are given in Table 8.

Table 8

Values of M1 for Principal Components from G1GC group in the period 2006:Q2 to 2007:Q2

	GCPC1	GCPC2	GCPC3
2006:Q2	0.93	0.31	0.11
2006:Q3	-0.04	-0.47	-22.77
2006:Q4	-14.85	-0.34	-834
2007:Q1	102.04	1.52	-44.12
2007:Q2	-28.33	3.43	-33.75

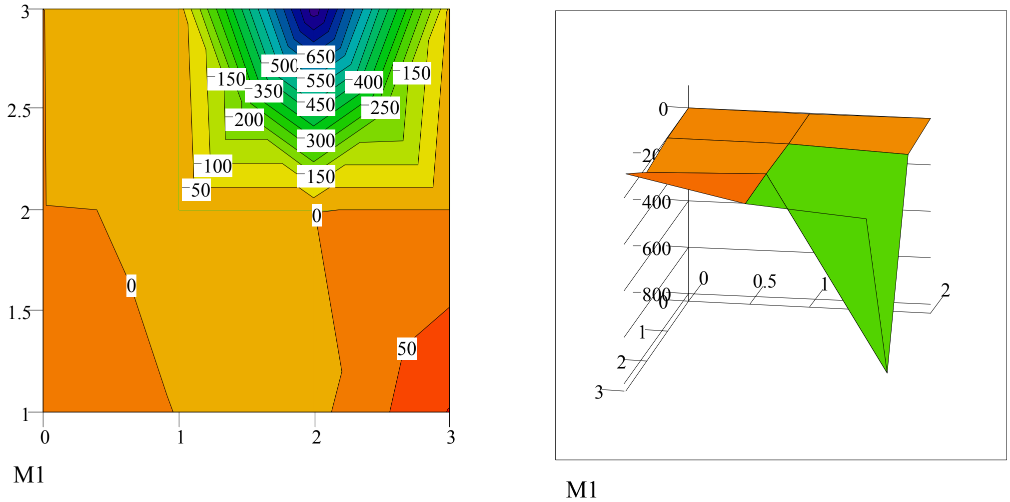
Source: Authors' calculations.

From the analysis of the error values we get the same kind of results as for the domestic consumption: the models successfully produced good accurate forecasts for the first three principal components for two periods ahead.

This could be easily observed in Figure 4, which presents the M1 percentage error for the first three principal components.

Figure 4

Graph of the M1- measure for the first 3 Principal Components of G1GC group



4. Conclusions

In this paper, we have constructed empirical models for domestic and public consumption.

We found evidence that these empirical models have applicative relevance. It was shown that these models are sufficiently accurate to predict for two or three periods ahead, on a medium-term dynamic forecast.. When we apply the model without incorporating the new information available in the real data (observations) and use the model to forecast for more than two periods ahead we must not expect good accurate results.

Our approach focuses on the selection procedure for lambda, which is the parameter from the moving average model of predictable sources found with the algorithm of complexity pursuit (Hyvärinen, 2001; Stone, 2004) as a generalization of projection pursuit to times series that consist in estimating projections of data whose complexity measure is minimized.

References

Hyvärinen, A. and Oja, E. (2000), "Independent Component Analysis: Algorithms and Applications", *Neural Networks*, 13(4-5):411-430.

Hyvärinen, A., (2001), "Complexity pursuit: separating interesting components from time-series", *Neural Computation*, 13(4):883–898.

Klein, Lawrence R. and Andrei Roudoi, Vladimir Eskin, Lucian-Liviu Albu, Cristian Stănică, (2004), "Quarterly GDP Data Correction Using Principal Components Analysis. The Case of the Romanian Economy – GDP Expenditures Side", *Romanian Journal of Economic Forecasting*, Supplement 2004.

Klein, Lawrence R. and Andrei Roudoi, Vladimir Eskin, Lucian-Liviu Albu, Cristian Stanica, Mariana Nicolae, Mihaela Nona Chilian, (2004), "Principal Components Model of the Romanian Economy GDP – Production Side", *Romanian Journal of Economic Forecasting*, Supplement 2004.

Klein, Lawrence R. and Andrei Roudoi, Vladimir Eskin, Mariana Nicolae, (2004), "Principal Components Model of the Romanian Economy. Study of the Oil Price Impact upon GDP", *Romanian Journal of Economic Forecasting*, Supplement.

Lopes, J. Carlos, Dias, J., Fereira do Amaral, J., (2007), "Economic Complexity as Input-Output Interrelatedness: A Comparison of Different Measurement Methods", paper submitted at the 16th International Input-Output Conference, Istanbul, Turkey, 2-6 July 2007.

Shnerb, N., Bettelheim, Y., Louzoum, Y., Agam, O., Solomon, S., (2001), "Adaptation of Autocatalytic Fluctuations to Diffusive Noise", *Phys. Rev.* E 63(2), http://xxx.lanl.gov.abs/cond-mat/007097

Solomon, S. (1998), "Stochastic Lotka Volterra Systems of Competing Auto-catalytic Agents lead generically to Truncated Pareto Power Wealth Distribution of Market Returns, Clustered Volatility, Booms and Crashes", *Decision Technologies for Computational Finance*, edited by A.-P. Refenes, A. N. Burgess and J. E. Moody (Kluver Academic Publishers, http://xxx.lanl.gov/abs/cond-mat/9803367.

Scutaru, C., Saman, C., Stănică, C. (2008), "Predictability and Complexity in Macroeconomics. The case of Gross Fixed Capital Formation in the Romanian Economy", *Romanian Journal of Economic Forecasting*, 9(4): 196-205.

Stone, James V., (2004), "*Independent Component Analysis. A Tutorial Introduction"*, The MIT Press Cambridge, Massachusetts, London, England.

Yaari, G., (2006), *The AB Model, the GLV, and a Data set from Poland*, Presentation of the Ph.D. thesis, under the supervision of Prof. Sorin Solomon, Racah Institute of Physics in the Hebrew University of Jerusalem, Israel; Multi-Agent Systems Division, ISI Torino, Italy.