2. MODELLING THE CONFIDENCE IN INDUSTRY IN ROMANIA AND OTHER EUROPEAN MEMBER COUNTRIES USING THE ORDERED LOGIT MODEL

Mariana GAGEA*

Abstract

The application of qualitative choice models is usually made by neglecting the analysis of autocorrelated and heteroscedastic errors. In the current paper, we aim to evaluate and mitigate the effects of violation of such a hypothesis using as example the modeling of confidence in industry in relation to the macroeconomic indicators for six countries of the European Union.

The ordered Logit model identified in the paper revealed the common macroeconomic factors which explain the formation of confidence in industry for the countries considered in the analysis. By mitigating the heteroscedasticity problems and specifying in the model the functional form of the error dispersion, the statistically significant improvement of the model performance was obtained.

Keywords: industrial confidence indicators, ordered logit model, Granger causality test, heteroscedasticity

JEL Classification: C12, C25, C87

I. Introduction

The econometrics of qualitative variables is related to the names of C. I. Bliss (1934), J. Berkson (1944) or D. McFadden (1969). Even if this field has recorded important developments in the last decades, clarifications from researchers are still expected on certain problems regarding the inclusion in the model of the interaction among terms, the number of variables which should be included in the model, the mitigation of hypotheses on autocorrelation and error heteroscedasticity.

The present paper aims to be a methodological-applicative research of the qualitative choice models, having two goals: 1) the building of a regression model that could

* *Alexandru Ioan Cuza University, Faculty of Economics and Business Administration, Blvd. Carol I nr. 22, Iaşi, 700505, Romania, E-mail: mariana.gagea@uaic.ro.*

explain the formation of managers' confidence in industry, for Romania and for other European Union member countries; 2) the mitigation of heteroscedasticity in the qualitative choice models.

The confidence indicators are obtained by means of data processing of business surveys, performed on a monthly and quarterly basis, in the following five sectors: manufacturing industry, constructions, consumption, retail trade and services. The business surveys are qualitative surveys, organised by the institute for statistics of every country in the European Union (the Directorate General for Eonomic and Financial Problems – DGECFIN) or the Organisation for Economic Cooperation and Development (OECD). The aim of the business surveys is the analysis of economic indicators evolution, by identifying the changes in specific sectors or in the economy, the performance of economic analyses and short-term forecasts. The indicators obtained are considered the key element of official statistical data. The main advantages of business surveys are the high frequency of data observations, the velocity of processing and the continuous harmonization of the methodology through the Business and Consumer Surveys program, which ensures the results comparability at international level.

The official European or national statistical bodies, the central banks, the research institutes use the confidence indicators in economic research having as main goals the forecast and identification of turning points of the gross domestic product or of the industrial production index, the research of correlations and causality between the confidence indicators and the reference series in the economic field under research, etc.

The relationship between the industrial confidence indicator and the industrial production index has been also studied by P. Bengoechea and G.P. Quiros (2004), who proposed a new methodology for dating the business cycle in the Euro Area economy by means of the industrial confidence indicator considered as a key variable in the identification of the current and future states of the Euro Area economy.

G. Bruno and M. Malgarini (2002) studied the predictive capacity of confidence indicators for the reference series for industry, retail trade, constructions and consumption by using a dynamic factor model. Other important research in this field was carried out by J. Vanhaelen, L. Dressea and J. De Mulder (2000), J. Goggin (2008).

The current paper entails a new approach of the analysis of confidence indicators, aiming to identify the factors that influence the formation of confidence in industry. The building of the best model is sought after, in order to explain the formation of entrepreneurs' confidence in industry in relation to the macroeconomic indicators. A comparative analysis of results will be also conducted for the EU member countries, by identifying certain features and generalities at the EU level, respectively.

The achievement of the second goal implies the verification of the existence of heteroscedastic errors in the regression model, the specification of its functional form in relation to the independant variables in the model and the assessment of the contribution of the heteroscedasticity correction to the improvement of the model performances.

The paper is structured in the following way: the second section briefly presents the main features, with constraints and advantages of the ordered Logit model; the third section comprises the identification of the confidence model in industry, with the approach specific to the regression model; the conclusions based on the results obtained in this paper are presented in the fourth part.

II. A Short Presentation of the Ordered Logit Model

The ordered models, besides the multinomial models, are part of the category of multiple choice models.

The dependent variable Y of an ordered model is defined in relation to a continuous latent variable *Y**, which is a linear combination of independent variables,

$$y_i^* = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i , \tag{1}$$

where: X_j represents the exogenous variables; β_j, $j = \overline{0, k}$, the regression coefficients; ε_i, $i = \overline{1, n}$, the error term. The deterministic component of the equation is denoted by z_i, $z_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij}$.

The ordered variable *Y*, having *p* categories, is defined as follows:

$$y_i = \begin{cases} 1, \, if \; y_i^* \le \delta_1 \\ 2, \, if \; \delta_1 < y_i^* \le \delta_2 \\ 3, \, if \; \delta_2 < y_i^* \le \delta_3 \\ \vdots \\ m, if \; y_i^* > \delta_{m-1} \end{cases} \tag{2}$$

For each statistical unit, we define the probability that the value of the Y variable should belong to a category s, $s = 1, 2, 3, \ldots, m$, of the form (Andrei and Bourbonnais, 2008: p. 329):

$$P_{is} = P(y_i = s) = F(\delta_s - z_i) - F(\delta_{s-1} - z_i) \tag{3}$$

where: $F(\cdot)$ represents the logistic distribution function, in the case of the logistic model, or normal, in the case of the Probit model, with $\sum_{i=1}^{m} P_i = 1$.

The estimation of β_j, $j = \overline{1, k}$ parameters, as well as of δ_s, $s = 1, 2, 3, \ldots, m$ values that define the categories of the variable *Y*, is performed by maximizing the likelihood function of the logit and normal distributions, respectively. The significance of the regression parameters is tested by means of the *t* Student test, the Wald test and LR test of the likelihood ratio.

The qualitative choice models are based on less restrictive hypotheses as compared to the hypotheses of the traditional regression analysis (Maddala, 1999: p. 15).

1. *The residual variables,* ε_i, $i = \overline{1,n}$ *, are independent and identically distributed (iid), of zero mean*.

2. *The error heteroscedasticity.* In the qualitative choice models, the heteroscedasticity affects the consistency of maximum likelihood estimators (Greene, 2003: p. 64) and represents a break in the independence hypothesis of irrelevant alternatives (IIA). The heteroscedasticity presence in the qualitative choice models is explained by the fact that the residues are obtained as a difference between the values of a discrete variable (the observed values y_i) and the values of a continuous variable (the estimated probabilities p_i).

The problem of heteroscedasticity can be mitigated through the correction of the Logit with the inverse of the error standard deviation, which means:

$$L_i^* = \beta_0 w_i + \sum_{j=1}^{k} \beta_j x_{ij}^* + v_i^* \tag{4}$$

where: w_i represents the weight; L_i^* the transformed or weighted logit; x_{ij}^* the *i* value of the transformed variable X_j; v_i^* the transformed error term, which is homoscedastic.

The SAS software program allows the mitigation of heteroscedasticity through the specification in the model of the functional form of the error variance in relation to the independent variables, considered a cause for heteroscedasticity. The QLIM procedure in the SAS software performs the estimation of the Logit and Probit models with corrected heteroscedasticity.

3. *The absence of error autocorrelation*. If the errors in the linear model of the latent variable from these models are autocorrelated of order one, the form of the likelihood function becomes a n-dimensional integer. The maximization of this function is almost impossible to determine. Under these circumstances, ignoring the error autocorrelation is common practice. Gourieroux, Monfort and Trognon (1980) showed that the estimators obtained in the presence of autocorrelated errors are consistent, while the formula of the standard deviation needs to be modified.

4. The absence of colinearity of independent variables. The qualitative choice models suffer from the same problems of multicolinearity as the traditional regression models do (Jula and Jula, 2009: pp. 163-167).

Limits and advantages. The ordered choice models have specific restrictions and advantages as well as some common with those of the binary choice models. The application of qualitative choice models is conditioned by the size of the analyzed sample. Generally speaking, one cannot establish the non-mobility, efficiency and normality properties of the estimators of the model parameters for samples of small size (Aldrich and Nelson, 1984: p. 53).

The effect of change in the independent variable, X_j , on the probability, $P(y_i = s)$, also dependes on the sign and size of the β_j parameter, but it is not entirely determined by it. This effect is not constant and its assessment needs some artifices of calculation. The interpretation of parameters is indirectly performed, similarly to the linear regression, only by means of the Logit.

Moreover, the ordered choice models are based on the critical hypothesis of slope parallelism attached to the independent variables, corresponding to different categories of the dependent variable for the same statistical unit (Borooah, 2001: p. 6). If this hypothesis is not met, i.e. the regression coefficients associated with an independent variable are different for various categories of *Y*, then the application of the ordered models is not appropriate and the multinomial Logit model is recommended.

The choice between the Logit and Probit models depends on the user's option. Generally, the Logit model is preferred, since it needs simpler calculations and the interpretation of the estimated parameters is easier.

III. Identifying an Ordered Logit Model of Confidence in Industry in Romania and in Other European Union States

The approach to the discrete-time regression models makes their analysis possible through the combination of econometric methods specific to the qualitative variables with those specific to time series. The selection of explanatory variables, an important stage in the construction of an econometric model, is made by means of Granger causality tests. Generally, the validation of discrete regression models is performed by testing the individual significance of the regression parameters, through the *t* test and the general significance of the model, by testing the LR likelihood ratio. In the present analysis, the above mentioned tests are completed with the heteroscedastic analysis and BDS tests (Brock – Dechert - Scheinkman), by means of which we verify if the errors are independent and identically distributed. The results of these tests will identify the models with heteroscedasticity problems.

During the next stage of our analysis, we attempt to improve the performance of the ordered Logit models through the mitigation of heteroscedasticity. The functional form of error dispersion is established in relation to one or several explanatory variables and we estimate the ordered Logit model with corrected heteroscedasticity.

In order to verify the model's robustness, we estimate and compare two Logit models, one for a period of control, January 2005 – December 2011, and a second one for the entire analyzed period, January 2005 – June 2012.

III.1. Description of the Data

Of the five economic sectors where the business surveys were performed, we analyze the most important sector: the industrial sector. The database on which relies the empirical research proposed in this paper contains the confidence indicators in industry and the macroeconomic indicators on which the economic agents' confidence is based in this field of activity.

In 2005, the development methodology of the business surveys at the level of National Institute for Statistics (NIS) underwent some changes. Taking into account this aspect, in order to ensure the data comparability, the time period recommended is 2005-2012, a relatively short period, with only 7 calendar years. The qualitative choice models are constrainted by the small size of the samples and we opted for the monthly, seasonally-adjusted data.

The statistical data series are formed of monthly indicators, registered at the level of the analyzed countries and at the level of the EU (27); between brackets we specified the symbols used in the paper:

- The confidence indicator in industry (ici_c);
- The index of the business environment (bc);
- The index of industrial production (ipi_c);
- The consumer price index (ipc_c);
- The unemployment rate (rs_c)
- The reference interest rate (rd_c; rd_ecb);
- The reference oil price (ob).

where: "_c" represents the country symbol (country) for which the analysis is performed.

Data changes. The confidence indicators show only the direction of the changes which will occur in the economic situation and not the size of these changes. Taking this aspect into consideration, we believe as appropriate the transformation of the variable attached to the confidence indicator in industry into ordered category variables (ici_c_ordered). In compliance with the significance thresholds established by NIS, we define the following categories:

- *Accentuated drop*, for values of ici_c lower than -40%, code 0;
- *Drop,* for values comprised within the interval [-40%, -15%), code 1;
- *Moderate drop,* for values within the interval [-15%, -5%), code 2;
- *Relative stability,* for values within [-5%, +5%), code 3;
- *Moderate growth,* for the interval [5%, 15%), code 4;
- *Growth* for the interval [15, 40), code 5;
- *Accentuated growth,* for values of ici_c higher than 40%, code 6.

The research is comparatively performed for Romania, Bulgaria, Greece, Spain, France and Germany. The statistical data have been provided by the following sources: Eurostat, the National Institute for Statistics – NIS, National Bank of Romania – NBR, Organization of the Petroleum Exporting Countries – OPEC.

III.2. Industry Confidence Factors of Influence

It is widely known that strong correlations between any two variables are insufficient to identify the cause and effect relations between them. In what follows, the identification of factors with significant influence on the confidence indicators is performed by means of Granger causality tests (Granger, 1969). The causality relation between variables can be verified in both directions, but the literature especially studies the influences of confidence indicators on the reference series in economy, and not in the opposite direction. The results presented are inconsistent.

Santero and Westerlund (1996) studied the correlation and Granger causality relations between the confidence indicators and macroeconomic variables (GDP, industrial production growth and some aggregate demand components) for a sample formed of 11 countries. The authors aimed to analyze the role of confidence indicators in the identification of stages of business cycles and in the forecast of economic changes, and not the explanation of the empirical behaviour of confidence indicators as well. The authors' conclusion was that the results obtained for individual countries were difficult to generalize, since the indicators conveyed different information and had a different time-relationship with economic variables in each country.

Other authors, such as Chopin and Darrat (2000), studied for the economy of the United States of America the role of confidence indicator in consumption in the forecast of some key macroeconomic variables. The results of Granger causality tests and of the VECM models indicated that the confidence indicator in consumption comprises useful information for the forecasting of some variables (incomes, interest rate), without being able to generalize the role of this indicator at the level of the entire economic activity. On the other hand, two factors contributing to the formation of consumer's confidence were identified: the inflation and the Dow Jones stock market index.

The previous conclusions are doubled by those obtained by Loría and Brito (2004), who ascertained that, for the United States economy, the confidence indicators in consumption is not Granger caused for the private consumption and investments, still the inverse causality is significant.

In the present paper, we aim to explain the behaviour of confidence indicators and the identification of their cause factors.

III.2.1. The application of the Granger test is preceded by the stationarity analysis, by means of the Augmented Dickey-Fuller (ADF) test. The results obtained are presented in Appendix 1.

All the confidence indicators in industry corresponding to the countries from the analyzed sample are integrated of order one, with the exception of Germany, which is stationary. Most of the macroeconomic indicators at national level are integrated of order one, while the consumer prices in Greece and Spain are integrated of order two.

At European and global level, 4 macroeconomic indicators are integrated of order one, the business environment in the Euro zone is stationary and the consumer price index for EU (27) is integrated of order two.

III.2.2. The Granger test is applied as follows: i) if the variables are stationary, the test is applied directly; ii) if the variables are not stationary, integrated of the same order, their cointegration is verified and then the Granger test is applied; iii) if the stationarity or cointegration conditions are not met, the causality verification is performed by means of the variant proposed by Toda-Yamamoto.

i) The confidence indicator in industry in Germany is stationary, as well as the business environment in the Eurozone. The pair of variables (cii_de, bc) meets the conditions of Granger causality test. The Granger test is applied and the probability attached to the Wald test smaller than the risk α of 0.05: $(p - value = 0.035) < (\alpha = 0.05)$ is obtained. The null hypothesis is rejected; the business environment in the Euro zone is Granger caused for the confidence indicator in industry in Germany.

ii) The cointegration between the confidence indicators in industry and the macroeconomic variables integrated of the same order I (1) is verified. The results obtained indicate that the confidence indicators are not cointegrated with the macroeconomic variables.

iii) The significant results of the causality test, Toda-Yamamoto version, for the pairs of variables which meet the conditions of this procedure, are presented in Appendix 2.

In brief, the causality tests applied above identified the following causal factors of the industry confidence indicator, by countries from the sample under analysis:

- Romania: ipi_ro, ipi_ue, rs_ue, rd_bce, bc, ob;
- Poland: ipi_pl, ipc_ue, rs_ue, rd_bce, bc;
- Greece: ipi_ue, rs_ue, rd_bce, bc, ob;
- Spain: ipi_es, ipi_ue, ipc_ue, rs_es, bc, ob;
- France: ipi_fr, ipi_ue, ipc_fr, ipc_ue, rs_ue, bc;
- Germany: ipi_ue, ipc_ue, rs_ue, bc

The causal indicators common in the formation of industry confidence for the analyzed countries are the indicators at the European Union level (the industrial production index (ipi_ue), the consumer price index (ipc_ue), the unemployment rate (rs_ue)) and the business environment in the Eurozone. The national industrial production index (ipi_c), the official interest rate fixed by BCE (rd_bce) and the official oil price (ob) are less frequent in the formation of confidence indicator in industry.

Based on the hypothesis that the industry confidence indicator reflects the gross domestic product, and the industrial production index, respectively, the causalities identified above are supported by laws and models demonstrated in the economic theory regarding the relations between the gross domestic product and key macroeconomic indicators such as: the unemployment rate, the consumer price index or the reference oil price.

III.3. The Estimation of the Ordered Logit Model

A regression model is identified describing the formation of industry confidence in relation to the cause factors identified in the previous subchapter. The variable *ici_c* is changed into the ordered variable *ici_c_ordered*, having the categories and codes presented in section II.1. The time period considered is January 2005 – December 2011.

In this section of the paper, the ordered Logit models without specified heteroscedasticity will be estimated. In the following sections, we attempt to identify and mitigate the error heteroscedasticity, by specifying in the ordered Logit model the functional form of the error dispersion. The following symbols of the two categories of models will be used: M1_c – for the Logit model without specified heteroscedasticity and M2_c – for the Logit model with corrected heteroscedasticity, where "c" represents the country's symbol.

The analysis is undertaken for each of the countries in the sample under consideration and comprises the following stages: i) *the selection of the independent variables* by means of the multiple linear regression of *ici_c* in relation to the cause variables, by means of the Backward method and the analysis of their colinearity, based on the indicators VIF and TOL; ii) the estimation and testing of an ordered Logit model that explains the formation of the variable ici_c_ordered in relation to the selected independent variables.

i) Following the selection of the explanatory variables through the above mentioned methods, from among the causal variables identified by means of the causality tests the analysis preserved: Romania: rs_ue, bc and ob; Poland: rs_ue and bc; Greece: rs_ue and bc;

Spain: rs_ue, rs_es, bc and ob; France: rs_ue, bc; Germany: ipc_ue, rs_ue, rs_de, bc.

We ascertain that the industrial production index and the consumer price index were eliminated from the analysis. The IPI and IPC colinearity in relation to the other independent variables in the linear multiple regression model is explained through the relations demonstrated in the economic theory, especially in relation to the unemployment rate.

ii) The ordered Logit models are estimated through the maximum likelihood method, using the SAS 9.1 software, by the QLIM procedure.

Some of the above selected explanatory variables were eliminated since the attached regression coefficients were not statistically significant.

For the final models, all the regression coefficients β_j, are statistically significant for a significance threshold of 5% (Table 1). The $\delta_i,\ i=1,2,\ldots m$ values, which underpin the definition of the categories of dependent variables ici_c_ordered, are statistically significant.

Table 1

The Coefficients of Ordered Logit Models without Specified Heteroscedasticity

Model	Dependent variable	Estimators	b_j	$s_{\hat{\beta}_j}$	t - test	
					Value	Probability
M1_RO	ici_ro_ordered	constant	19.569	5.006	3.910	0.000
		rs_ue	-1.785	0.516	-3.453	0.000
		bc	1.951	0.370	5.267	0.000
M1_PL	ici_pl_ordered	constant	31.359	7.577	4.14	0.000
		rs_ue	-3.311	0.813	-4.07	0.000
		bc	6.730	1.683	4.00	0.000
M1_EL	ici_el_ordinal	constant	21.287	3.195	6.66	0.000
		rs_ue	-1.608	0.296	-5.42	0.000
		bc	2.156	0.383	6.63	0.000

Model	Dependent variable	Estimators	b_j	$s_{\hat{\beta}_j}$	t - test	
					Value	Probability
M1_ES	ici_es_ordinal	constant	12.416	2.087	5.95	0.000
		rs_es	-0.302	0.059	-5.08	0.000
		bc	2.941	0.517	5.69	0.000
M1_FR	ici_fr_ordinal	constant	31.313	6.240	5.02	0.000
		rs_ue	-0.933	0.383	-2.44	0.014
		bc	7.687	1.453	5.10	0.000
M1_DE	ici_de_ordinal	Constant	49.909	10.296	4.85	0.000
		rs_de	-2.162	0.459	-4.71	0.000
		bc	11.647	2.366	4.92	0.000

The business environment in the Eurozone (bc) and the unemployment rate at EU level (rs_ue) significantly explains the formation of industry confidence for Romania, Poland, Greece and France. The explanatory model of the industry confidence for Spain and Germany are of the same form, but it comprises the unemployment rate at national level, rs_es and rs_de, respectively. The sign of the regression coefficients shows that the formation of industry confidence is negatively influenced by the unemployment rate and positively by the business environment in the Eurozone. With the exception of the Logit model for Romania, for all the other models there are significant differences between the absolute values of the two regression coefficients, while the higher absolute value of the coefficient attached to the business environment indicates the stronger influence of this indicator on the ici_c_ordered.

According to the LR test of the likelihood ratio (Table 2), all the models are significant on a whole: $LR_{Ma\_c} > (\chi^2_{0,05;2} = 5{,}991)$. The high values of the pseudo indicators – R^2 (Estrella, Adjusted Estrella and Mc. Fadden LRI) indicate that the variation in the variable ici_c_ordered is well explained through the specified qualitative choice model.

The most performant models, having the highest values of the pseudo indicators – R^2, are: M1_DE – Germany, M1_FR – France and M1_PL - Poland.

Table 2

The Synthesis of the Logit Models without Specified Heteroscedasticity

Indicator	M1_RO	M1_PL	M1_EL	M1_ES	M1_FR	M1_DE
Log Likelihood	-42.884	-22.379	-69.657	-56.281	-24.960	-29.880
AIC	95.769	52.759	153.377	124.562	59.920	73.761
SC	107.923	62.483	167.962	139.147	72.074	90.777
LR	74.149	131.03	103.43	120.230	143.22	216.85
Estrella	0.6946	0.9429	0.7981	0.866	0.955	0.993
Adjusted Estrella	0.6232	0.9192	0.7440	0.823	0.932	0.987
Mc. Fadden LRI	0.4637	0.7454	0.4221	0.457	0.7410	0.784

Note: The above results can be generalized at the level of all European Union countries as follows: the industry confidence, considered as a category variable, may be explained by means of an ordered Logit model, in relation to the uneployment rate (at national or European level) and to the business environment in the Eurozone.

III.4. Verifying the Model's Hypotheses

III.4.1. The analysis of error heteroscedasticity is performed by means of graphic methods. The errors of the Logit models are of zero mean, so that the dispersion is estimated through the mean square error: $\sigma_i^2 = M(e_i^2),\ i = \overline{1,n}$.

The graphic representation of the dynamics of the Logit models' mean square error corresponding to Romania, Poland, Greece and Spain (Figures 1-4) highlights heteroscedasticity problems, with a potential exponential relation of error dispersion in relation to one or several explanatory models. For the models corresponding to France and Germany (Figures 5 and 6) we believe that the errors do not have heteroscedasticity problems.

Figure 1
The Mean Square Error of the M1_RO Model

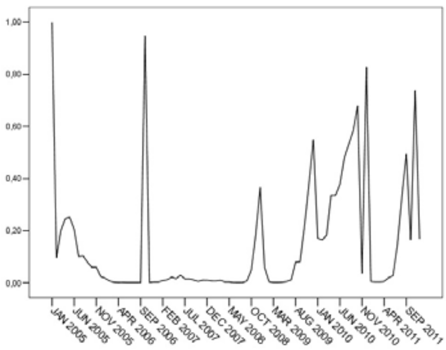


Figure 2
The Mean Square Error of the M1_PL Model

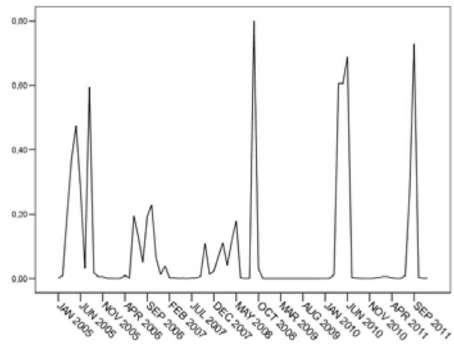


Figure 3
The Mean Square Error of the M1_EL Model

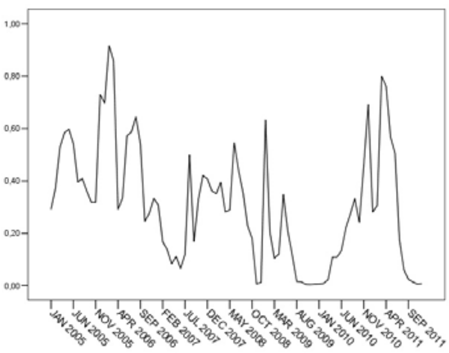


Figure 4
The Mean Square Error of the M1_ES Model

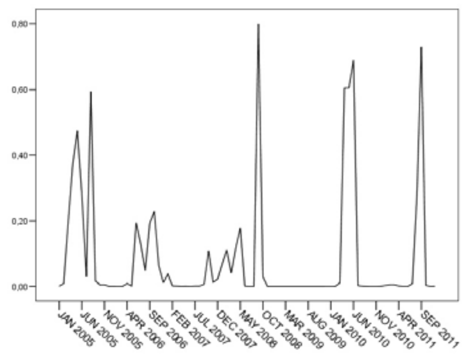


Figure 5

The Mean Square Error of the M1_FR Model

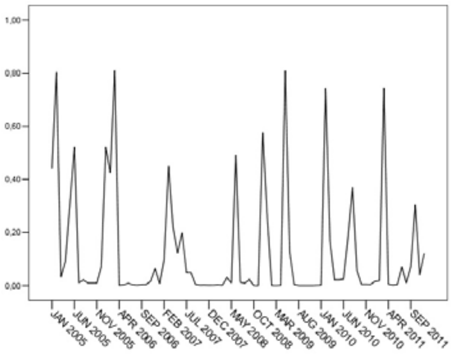
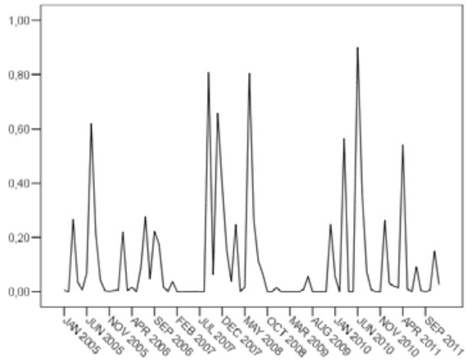


Figure 6

The Mean Square Error of the M1_DE Model



III.4.2. Verifying the hypothesis according to which the errors are independent and identically distributed (iid) is performed by means of the BDS test (Jula and Jula, 2012: pp. 33-42). In the current analysis, the sample is small ($n = 84$) and the test of the hypothesis iid of the errors of the Logit model will be conducted using the Eviews program, with the recommended techniques for such situations and employing the critical values recommended by Kanzler (1999). The correlation dimension will be at most equal to 6.

Table 3

BDS Statistics, Test Statistics and Probabilities, Calculated by Means of Bootstrap Techniques

Model \ m		2	3	4	5	6
M1_RO	BDS statistics	0.076	0.128	0.170	0.169	0.156
	test statistics	6.810 (0.000)	7.157 (0.000)	7.919 (0.000)	7.481 (0.000)	7.069 (0.000)
M1_PL	BDS statistics	0.062	0.100	0.1232	0.123	0.114
	test statistics	4.273 (0.000)	4.036 (0.000)	4.046 (0.001)	3.871 (0.004)	3.688 (0.006)
M1_EL	BDS statistics	0.100	0.167	0.211	0.233	0.239
	test statistics	21.402 (0.000)	22.399 (0.000)	23.724 (0.000)	25.185 (0.000)	26.789 (0.000)
M1_ES	BDS statistics	0.059	0.087	0.095	0.092	0.080
	test statistics	6.180 (0.000)	5.665 (0.000)	5.207 (0.000)	4.786 (0.000)	4.286 (0.004)
M1_FR	BDS statistics	0.030	0.036	0.017	0.008	0.003
	test statistics	2.534 (0.033)	1.870 (0.094)	0.753 (0.388)	0.325 (0.592)	-0.106 (0.848)
M1_DE	BDS statistics	0.015	0.013	0.030	0.037	0.037
	test statistics	1.334 (0.214)	0.759 (0.413)	1.382 (0.189)	1.629 (0.133)	1.670 (0.125)

For the models corresponding to Romania, Poland, Greece and Spain, the probabilities attached to the statistics of the test (Table 3) are smaller than the significance threshold of $\alpha = 0.05$, while the absolute value of the BDS statistics is higher than the corresponding critical value recommended by Kanzler. We decide to reject the null hypothesis, the errors of these models are not iid, which means that there is a non-linear dependence in their structure.

For the models of France and Germany, the probabilities attached to the test statistics are higher than the significance threshold of $\alpha = 0.05$ and the absolute value of the BDS statistics is smaller than the corresponding critical value recommended by Kansler. The null hypothesis is not rejected; the errors of these models are iid.

III.5. Mitigating the Heteroscedasticity

We aim to evaluate and mitigate the error heteroscedasticity in the models for which heteroscedasticity problems have been identified. For this purpose, the procedure is as follows: the Logit models with corrected heteroscedasticity are estimated by specifying the functional form of error dispersion in relation to the independent variables in the model: the statistical significance of the specified heteroscedasticity is verified, using the LR test; the model performance with and without corrected heteroscedasticity are compared.

For the heteroscedastic models, an exponential form has been identified, without a constant term of error dispersion, in relation to the following independent variables: the model M2_RO – bc and rs_ue; M2_PL – bc; M2_EL – rs_ue; M2_ES – rs_es.

Table 4

The Coefficients of the Ordered Logit Models with Corrected Heteroscedasticity

Model	Dependent variable	Estimators	b_j	$s_{\hat{\beta}_j}$	t-test Value	t-test Probability
M1.b Romania	ici_ro_ordered	constant	5092.622	93.767	54.31	0.000
		rs_ue	-469.253	22.861	-20.63	0.000
		bc	464.336	32.515	14.28	0.000
		H_bc	1.147	0.055	20.73	0.000
		H_rs_ue	1.032	0.557	1.85	0.000
M2.b Poland	ici_pl_ordered	constant	31.779	8.136	3.91	0.000
		rs_ue	-3.380	0.876	-3.86	0.000
		bc	6.542	1.646	3.97	0.000
		H_bc	-1.007	0.392	-2.56	0.010
M3.b Greece	ici_el_ordered	constant	13520	27.179	497.43	0.000
		rs_ue	-775.198	19.402	-39.95	0.000
		bc	2080.704	146.010	14.25	0.000
		H_rs_ue	1.531	0.032	47.64	0.000
M4.b Spain	ici_es_ordered	constant	3.522	1.809	1.95	0.051*
		rs_es	-0.081	0.043	-1.89	0.059*
		bc	0.826	0.438	1.89	0.059*
		H_rs_es	-0.203	0.077	-2.64	0.008

* The coefficients are significant for a threshold of 10%.

All the parameters of the models with corrected heteroscedasticity (the regression coefficients in the basic model, the coefficients in the the errors' heteroscedasticity function, the values $\delta_i, i=1,2,\ldots m$ are statistically significant for a significance threshold of at most 10% (Table 4).

The verification of the statistical significance of the correction of heteroscedasticity is performed by means of the LR test and by comparing the log-likelihood values corresponding to the model with and, without corrected heteroscedasticity, respectively.

For Romania, we obtain:

$$LR = -2\ln L(M1.a) + 2\ln L(M1.b) = -2x(-42.884) + 2x(-38.552) = 8.664$$

$(LR = 8.664) > (\chi^2_{0.05;\ 2} = 5.991)$, the null hypothesis is rejected; therefore, the specified heteroscedasticity is statistically significant, under the conditions of a taken risk of 5%.

Similar results are obtained for the models of Poland, Greece and Spain.

Table 5

The Synthesis of the Logit Models with Corrected Heteroscedasticity

Indicator	M2_RO	M2_PL	M3_EL	M4_ES
Log Likelihood	-38.552	-20.342	-65.563	-52.364
AIC	91.104	50.684	145.127	118.728
SC	108.120	62.838	162.143	135.744
LR	82.814	135.11	113.68	128.06
Estrella	0.750	0.953	0.838	0.890
Adjusted Estrella	0.657	0.926	0.782	0.845
Mc, Fadden LRI	0.518	0.769	0.464	0.551

All the performance indicators of the models with corrected heteroscedasticity (Table 5) are superior to the indicators corresponding to the model without corrected heteroscedasticity.

The empirical data are graphically compared with the adjusted data through the two forms of the ordered Logit models and we ascertain the following improvements of the second model:

- In the case of Romania: the model M2_RO corrects two adjustment errors of the M1_RO model (Figures 7 and 8). These errors correspond to the moments May 2005 and December 2008;
- In the case of Poland: three errors of the initial model are corrected, corresponding to the moments: March 2005, June 2006 and September 2006 (Figures 9 and 10);
- In the case of Spain: two errors of the M1_ES model are corrected, corresponding to the moments February 2008 and June 2010 (Figures 11 and 12);
- In the case of Greece: the M2_EL model corrects 11 errors of the M1_EL model, but incorrectly estimates other 8, having an advantage of 3 correct estimations.

Figure 7
The Initial and Adjusted Data by Means of the M1_RO Model

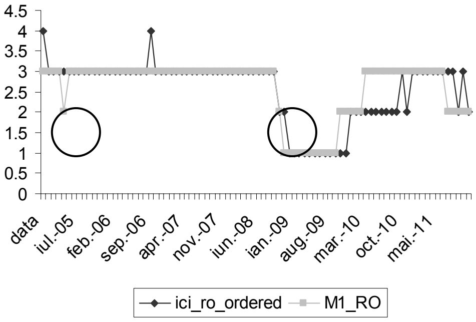


Figure 8
The Initial and Adjusted Data by Means of the M2_RO Model

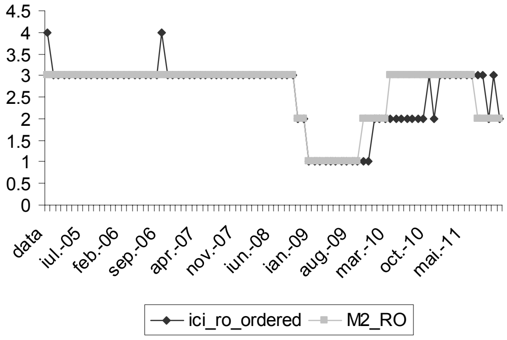


Figure 9
The Initial and Adjusted Data by Means of the M1_PL Model

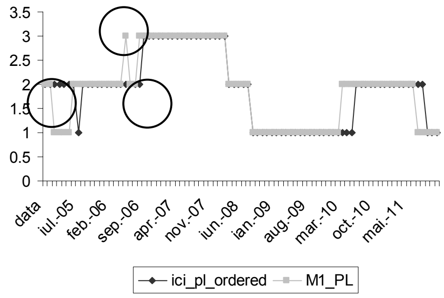


Figure 10
The Initial and Adjusted Data by Means of the M2_PL Model

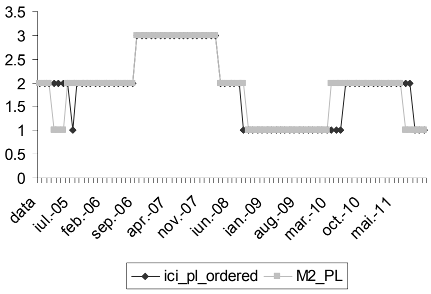


Figure11
The Initial and Adjusted Data by Means of the M1_ES Model

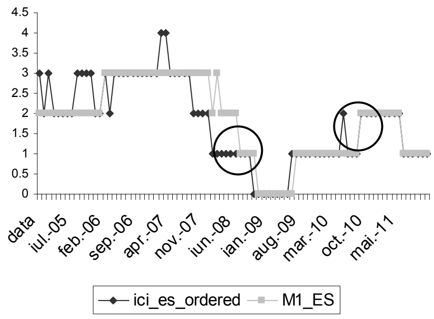
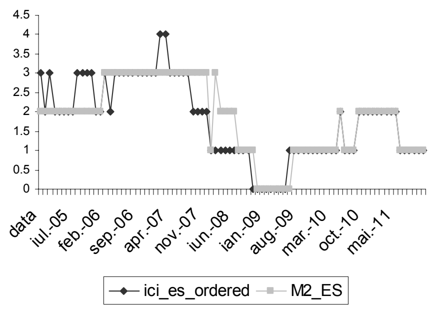


Figure 12
The Initial and Adjusted Data by Means of the M2_ES Model



Based on the graphic and numerical analyses, we may conclude that the Logit model with corrected heteroscedasticity is more performant than the Logit model without correction.

III.6. The Robustness of the Model

The ordered Logit model without heteroscedasticity is estimated for the period January 2005 – June 2012 and is compared with the corresponding model for the period January 2005 – December 2011.

According to the statistical tests, the estimated models are significant. The performance indicators were modified, for some of the models indicating better adjustments (Germany, Spain, Greece), while for others, weaker adjustments in comparison with control models corresponding to the period January 2005 – December 2011.

The verification of hypotheses regarding the errors of the ordered Logit model does not show significant differences betweent the models corresponding to the two periods. The BDS test indicates independent and identically distributed errors for the models corresponding to France and Germany. These four models, much like the pair models estimated for the control period, are affected by heteroscedasticity.

As a conclusion, the ordered Logit model that explains the formation of industry confidence in European Union countries, in relation to the unemployment rate (at European or national level) and to the business environment in the Eurozone, is validated by the results obtained for the two periods for the sample of countries under analysis.

IV. Conclusions

The discrete regression models are generally applied, under conditions of autocorrelation and heteroscedasticity of errors, which mitigates the models' performance. In the present paper, we attempted to highlight this aspect through the example of the ordered Logit model elaborated for the confidence indicator in industry in relation to macroeconomic indicators, in Romania and other European Union member countries.

The confidence indicator in industry, obtained by means of the business surveys, has the role of bringing more information about the evolution of gross domestic product or of industrial production index, for a much shorter period than the official statistical data. When selecting the explanatory variables of the confidence formation in industry we had in view the macroeconomic indicators that describe the situation of an economy, with a focus on the factors that underpin the theories of economic growth.

The results of the causality analysis indicated the explanation of the confidence formation in industry through the following variables: the industrial production index, the unemployment rate, the consumer price index, the business environment in the Eurozone and the reference oil price. The economic theory supports the causality relations identified between the formation of confidence in industry and the macroeconomic indicators. The model identified in the paper explains the formation of confidence in industry in relation to the unemployment rate and the business environment in the Eurozone, through an ordered Logit model.

The verification of the hypotheses formulated regarding the errors of the ordered Logit model highlighted significant problems about heteroscedasticity for some of the models. The mitigation of heteroscedasticity, with the specification in the model of the functional form of error dispersion led to the improvement of performance of all these models. The model's robustness was verified by comparing the estimated models for the analyzed period and for a control period.

The present research will be followed up by the extension of analysis to all the European Union countries, aiming to generalize the ordered Logit model of confidence indicator in industry in relation to unemployment rate. We will analyze the inclusion of new significant indicators which explain the formation of confidence in industry and we will develop analytical and forecasting simulations on the economic and social evolution in Romania and in the European Union countries. Another purpose is to shape a relevant methodology of analysis and forecast based on the qualitative choice models.

Acknowledgements

The author is grateful to Prof. Dorin Jula for his suggestions made to this paper. Acknowledgements are due to the President of the National Commission for Forecasting, Prof. Ion Ghizdeanu, and to the former Vice-President of the National Institute for Statistics, Cristian Nicolae Stănică, Ph.D., for their support in the collection of statistical data.

This work was supported by the project "Post-Doctoral Studies in Economics: training program for elite researchers - SPODE" co-funded from the European Social Fund through the Development of Human Resources Operational Programme 2007-2013, contract no. POSDRU/89/1.5/S/61755.

References

Aldrich, J.H. and Nelson, F.D., 1984. *Linear Probability. Logit and Probit Models.* Iowa City: SAGE Publications.

Andrei, T. and Bourbonnais, R., 2008. *Econometrie.* Bucureşti: Editura Economică.

Bengoechea, P. and Quiros, G.P., 2004. Useful Tool to Identify Recessions in the Euro-area. European Economiy. *Economic Papers*, 215. Available at: <http://ec.europa.eu/economy_finance/publications/publication566_en.pdf> [Accesed on April 2012].

Bliss, C.I., 1934. The Method of Probits. *Science*, 79, pp. 38-39.

Berkson, J., 1944. Application of the Logistic Function to Bioassay. *Journal of the American Statistical Association*, 39, pp. 357-365.

Borooah, V.K., 2002. *LOGIT and PROBIT, Ordered and multinomial Models*, California: SAGE Publications.

Bruno, G. and Malgarini, M., 2002. An Indicator of Economic Sentiment for the Italian Economy. *Documento di Livoro*, ISAE, Rome, no. 28.

Chopin, M. and Darrat, A.F., 2000. Can Consumer Attitudes forecast the Macroeconmy?. *American Economist*, 44(1), pp. 34-42. Available at:

<http://www.freepatentsonline.com/article/American-Economist/63299525.html> [Accessed 10 October 2011].

Goggin, J., 2008. An Analysis of the Potential of the European Commision Business and Consumer Surveys for Macroeconomic Forecating, *Quarterly Economic Commentary*, Winter. pp. 45-67.

Gourieroux, Ch., Monfort, A. and Trognon, A. 1980. Pseudo maximum likelihood methods: Application to poisson models. *Econometrica,* 52(3), pp. 701-720.

Granger, C.W.J., 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 37, pp. 424-438.

Greene, W.H., 2003. *Econometric Analysis*, Second Edition. New York: McMillan Publishing Company.

Jula, D. and Jula, N.M., 2012. *Prognoză Economică*, Bucureşti: Editura Mustang.

Jula, D. Jula, N.M., 2013. Economic Growth and Structural Changes in Regional Employment. *Romanian Journal of Economic Forecasting*, 16(2), pp. 52-69.

Jula, N. and Jula D., 2009. *Modelare economică. Modele econometrice şi de optimizare*. Bucureşti: Editura Mustang.

Kanzler, L., 1999. *Very Fast and Correctly Sized Estimation of the BDS Statistic*. Manuscript. University of Oxford. Available at: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=151669> [Accessed 25 February 2012] .

Loría, E. and Brito, L. (2004) , *Estudios de economía aplicada*, 22-3, art. 22306, Available at: <http://www.economia.unam.mx/profesores/eloria/articulos/art_culo_EEA.pdf> [Accessed 15 February 2012].

Maddala, G.S., 1999. *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

McFadden, D.L., 1975. The Revealed Preferences of a Government Bureaucracy: Theory. *The Bell Journal of Economics and Management Science*, 6(2), pp. 401-416.

Santero, T. and Westerlund, N., 1996. Confidence indicators and their relationship to changes in economic activity, *OECD Economics Department Working Papers*, no. 170, OECD Publishing.

Vanhaelen, J.J., Dresse, L. and De Mulder, J., 2000. The Belgian Industrial Confidence Indicator: Leading Indicator of Economic Activity in the Euro Area?, *Working Papers*, No.12, National Bank of Belgium.

Appendix 1

Unit Root Tests - Synthesis

Variable	Romania			Poland		
	ADF	ADF - 1st difference	I(d)	ADF	ADF - 1st difference	I(d)
ici_c	−1.265 (0.188)	−5.379 (0.000)	1	−0.919 (0.316)	−5.934 (0.000)	1
ipi_c	1.523 (0.968)	−8.826 (0.000)	1	2.533 (0.997)	−12.826 (0.000)	1
ipc_c	−2.262 (0.450)	−8.065 (0.000)	1	−2.849 (0.184)	−4.300 (0.000)	1
rs_c	−2.265 (0.448)	−10.918 (0.000)	1	−2.265 (0.448)	−1.978 (0.046)	1
Variable	Greece			Spain		
	ADF	ADF - 1st difference	I(d)	ADF	ADF - 1st difference	I(d)
ici_c	−0.304 (0.574)	−8.198 (0.000)	1	−0.731 (0.397)	−3.774 (0.000)	1
ipi_c	−2.056 (0.562)	−17.318 (0.000)	1	−1.149 (0.226)	−3.536 (0.000)	1
ipc_c	−3.410 (0.060)	−6.115* (0.000)	2	−2.953 (0.152)	−7.682* (0.000)	2
rs_c	−1.538 (0.932)	−6.097 (0.000)	1	−2.435 (0.360)	−2.155 (0.031)	1
Variable	France			Germany		
	ADF	ADF - 1st difference	I(d)	ADF	ADF - 1st difference	I(d)
ici_c	−1.265 (0.188)	−5.258 (0.000)	1	−2.459 (0.014)	-	0
ipi_c	−2.371 (0.039)	-	0	−2.506 (0.118)	−4.737 (0.000)	1
ipc_c	−2.545 (0.306)	−8.324 (0.000)	1	−1.918 (0.637)	−14.210 (0.000)	1
rs_c	−2.671 (0.251)	−2.067 (0.038)	1	−2.498 (0.328)	−3.204 (0.023)	1

Variable	ADF	ADF - 1st difference	I(d)	Variable	ADF	ADF - 1st difference	I(d)
ipi_ue	−2.054 (0.264)	−3.482 (0.000)	1	rd_bce	−2.147 (0.512)	−3.342 (0.001)	1
ipc_ue	−3.243 (0.083)	−3.750* (0.000)	2	bc	−2.569 (0.011)	-	0
rs_ue	−2.475 (0.339)	−2.139 (0.032)	1	ob	−3.306 (0.072)	−5.483 (0.000)	1

* ADF 2nd difference.

Appendix 2

Results of Toda-Yamamoto Version of the Granger Test

Country	Variable	Chi-square		Country	Variable	Chi-square	
		Value	Probability			Value	Probability
1	2	3	4	1	2	3	4
Romania	ipi_ro	10.929	0.004	**Spain**	ipi_es	10.319	0.005
	ipi_ue	15.957	0.000		ipc_ue	7.167	0.027
	rs_ue	12.612	0.002		ipi_ue	7.965	0.018
	rd_bce	9.073	0.011		rs_es	20.693	0.000
	ob	16.204	0.000		bc	6.217	0.044
	bc	21.612	0.000		ob	9.603	0.008
Poland	ipi_pl	12.845	0.016	**Germany**	ipi_ue	6.642	0.036
	ipc_ue	6.870	0.032		ipc_ue	7.860	0.019
	rs_ue	5.018	0.081*		rs_ue	5.392	0.052*
	rd_bce	9.118	0.010		rs_de	5.837	0.054*
	bc	18.639	0.000				
Greece	ipi_ue	12.755	0.002	**France**	ipi_fr	9.959	0.006
	rs_ue	6.808	0.033		ipi_ue	16.328	0.000
	rd_bce	6.467	0.039		ipc_fr	6.473	0.039
	ob	6.054	0.048		ipc_ue	10.988	0.004
	bc	12.417	0.002		rs_ue	8.752	0.012
					bc	11.900	0.002

* The tests are significant for a threshold of 10%.