Academy of Economic Studies

Doctoral School of Finance and Banking

Credit risk modeling – a macro perspective

MSc Student: Bogdan Chiriacescu

Supervisor: Ph.D. Professor Moisă Altăr

Bucharest, July 2010

Table of contents

Abstract

1. Introduction………………………………………………………………………………. 4

2. Literature review…………..…………………………..…………………………………….. 6

3. Methodology……………………………………………………………………………… 10

4. Data………………………………………………………………………………………... 14

5. Estimation, results and stress testing……………………..………………………………….. 18

   5.1 Household Sector…………………………………………………………………….. 19

   5.2 Corporate Sector ……………………………………………………………………... 28

6. Conclusions ……………………………………………………………………………...33

Bibliography ………………………………………………………………………………… 35

Annex 1 – Household Sector ……………………………………………………………….. 37

Annex 2 – Corporate Sector ……………………………………………………………… 40

Abstract

The importance of credit risk assessment and monitoring has increased since the recent financial turmoil. This paper presents a toolkit for credit risk modeling that follows the top-down approach proposed by Wilson (1997). The analysis is conducted separately for the household and corporate sector, by means of panel techniques and seemingly unrelated equations, using default aggregated data at county or business sector level. The results indicate that the determinants of default on bank loans for the household sector are unemployment, exchange rate, industrial production, indebtedness and interest rate spreads, while for the corporate sector the output gap, indebtedness and exchange rate are the main factors. Comparing the two models, it arises that default events from the corporate sector occur sooner than for the household sector in case of adverse macroeconomic developments. There are two possible explanations: i) there is no personal bankruptcy law for individuals in Romania, and ii) public administration appears to adjusts slower during recessions, an important part of the work force being part of this system. Furthermore, stress-testing analysis is performed on arbitrarily built portfolios by considering the impact of macroeconomic shocks on the probabilities of default over a one year time horizon.

1. Introduction

In recent years, one area that has emerged as an objective on which public authorities pay more attention is financial stability. An important lesson from this crisis is that price stability alone is not enough to achieve sustainable, non-inflationary growth and a high level of employment, as it is stated in the objectives of the most important central banks. Firstly, it has become obvious that general equilibrium models used by policy makers in setting the interest rates must be augmented with variables of a financial nature that control for the risks stemming from rapid credit expansions and asset price bubbles phenomena. Secondly, and equally important the health of the financial system seems to have caught the spotlight in international finance since the Lehman Brothers fall in September 2008. Governments have spent huge amounts of money to bail out banks that were in a bad shape. The costs were important and that has also contributed considerably to the increased budget deficits in some of these countries that are troubling now governments. It is more and more considered that “the invisible hand” appears to have failed in assuring the well functioning of the financial markets that somehow got it wrong by mispricing risks and ending in a financial turmoil. That is why public authorities are taking now a more proactive stance in building a new supervisory and regulatory framework in which the forward looking component is much more important.

In order to enforce financial stability it is obligatory to correctly evaluate the banking sector’s vulnerabilities as a whole. One method is to observe how resilient the system is to “exceptional but plausible events”, which is known as stress testing the system in abnormal conditions. In order to achieve this type of results, one needs to model first the way the system is behaving based on what has happened in the past.

The macro-prudential framework, consisting of risk modeling and stress-testing analysis, is now a standard instrument in central banks used for the purpose of assessing financial stability. Originally, it was a risk management tool developed by commercial banks in order to assess how their portfolio would react to a sudden crisis situation. It was proven though, that banks alone can’t manage risks properly. Some reasons worth mentioning relate to financial innovation exuberance that went a bit further than the controllable area and also somehow different goals for the management (short-term bonuses) and equity holders (long term wealth). Analysis therefore must be conducted by supervisory authorities at the aggregate level of the financial system taking into account all the inter-linkages in the system. Currently, there are advanced discussions about creating supranational bodies to also capture the risks emanating from the interconnections of multinational financial institutions that operate in different countries. Such an example is the European Systemic Risk Board.

The risks are of different nature (credit risks, market risks, liquidity risks, operational risks) and therefore are treated idiosyncratically. Credit risk is the most important of them, as the major source of losses incurred by banks comes from this side. It comes naturally to investigate which are the main drivers of credit risk at a systemic level. The burgeoning literature of the last decade concerning this topic presents strong evidence relating the business cycle and credit risk and the nonlinearities of the relationship. This is an argument for central banks to embark on a top-down approach in credit risk modeling, relating default events to macro variables, and subsequently to potential losses of the financial sector. By following such a procedure, first round effects of shocks to macroeconomic variables on financial sector are estimated. Trying to capture feedback (second round) effects from the economy and financial markets is at the moment at an incipient stage of research because of the complexities involved. Micro models (bottom-up approach) are the other available alternative, but they have a serious drawback being much more data intensive, with most of these data being confined to confidentiality issues (especially for the household sector).

Therefore, the goal of this paper is to provide a framework for credit risk modeling that looks distinctively at the corporate and household sectors in Romania. The methodology is the one proposed by Wilson (1997) and integrated in Credit Portfolio View. This involves the development of a multifactor model for systematic default risk that captures the relationship between default on bank loans and economic cycle. For the household sector, monthly aggregated data at county level is employed using panel estimation. In the case of corporate sector, the analysis is conducted industry wise o quarterly data and the estimation is performed by means of seemingly unrelated equations. The non-linear relationship between credit risk and macroeconomic conditions is modeled through the logistic transformation of the default rates. This is the most frequently used transformation to account for the fact that credit risk increases substantially more in times of high stress. The definition of default used in this paper follows the Basel II framework, and counts any credit obligation that is past due more than 90 days.

The explanatory variables that provide the best fit for the models are: i) unemployment, exchange rate, degree of indebtedness, interest rate spreads and industrial production for the household sector, and ii) output gap, indebtedness and exchange rate for the corporate sector. The risks stemming from the two sectors are compared by estimating a one year probability of default. Lastly, a stress test analysis is performed on arbitrarily built portfolios to capture the adverse effects of macroeconomic shocks on the potential loss of these portfolios.

The paper is organized as follows. Section 2 provides a review of the work done in this area of research. Section 3 presents the methodology employed. Section 4 describes the data used and its limitations. Section 5 presents the estimation results and main findings of the models, containing also the stress test analysis for the arbitrarily portfolios, and section 6 concludes the paper.

2. Literature Review

The area of credit risk modeling has grown in importance starting with the last decade and has become increasingly more abundant in the recent years. The main reason is that credit risk modeling and stress testing analysis are now standard tools for central banks in the process of assessing financial stability.

Most of the related studies investigating the macroeconomic determinants of credit risk have tried to analyze this relationship along at least a full business cycle. The reason behind it is that in periods of high stress the default rates are increasing a lot faster than in normal times. The various techniques used fall in three categories: time series analysis, panel data regressions or structural models, the number of the latter being somehow benign.

A substantial amount of literature devoted to credit risk macro models follows the methodology proposed by Wilson (1998). In his paper he describes a new approach that tabulates the loss distribution arising from correlated credit events for arbitrary portfolios of nonfinancial corporations, both at a regional and at industry sector level. The importance of having a loss distribution rather than a single potential loss is highlighted by the fact that counterparty defaults can be predicted with a degree of uncertainty and are not perfectly correlated. Wilson approach is to model expected and unexpected loss of a portfolio, taking into account the number and size of the loans in the portfolio, in contrast to considering a normal distribution or mean-variance approximations at portfolio level. This is important because having a heterogeneous portfolio in terms of size of loans, the loss distribution will be bimodal and highly skewed as opposed to unimodal and symmetric. Another important improvement consists in relating the loss distribution to the actual state of the economy, rather than based on the unconditional or historic averages of losses from default events, that do not reflect portfolio's true credit risk in resonance with present macroeconomic conditions. In this way, the model will capture the cyclical default events that represent the biggest part of risk for diversified portfolios. One important conclusion of Wilson's research is actually exactly the fact that the bulk of systematic or non-diversifiable risk of any portfolio can be explained by the economic cycle.

Pesola (2001) proposed a dynamic panel model to study the period with banking crisis that affected the Nordic countries during the 1990's. The dependent variable was loan losses divided by total lending or enterprise bankruptcies, while the explanatory ones were the lagged dependent variable along with lagged gdp growth rate, an income surprise variable combined with lagged indebtedness, change in real interest rate combined with lagged indebtedness and dummy variables for changes in regulations. Total indebtedness is the most important explanatory variable, being a proxy for the financial fragility. The surprise variables are compiled as the difference between the expected value of the variable and the actual outcome. The expected values are taken from the forecasts made by OECD. Empirical results suggest that high levels of indebtedness together with adverse economic shocks (surprises) are the causes of the banking crisis in Sweden, Norway and Finland.

Kalirai and Scheicher (2002) estimate time series regressions of total loan loss provisions for the Austrian banking sector with respect to a wide range of macroeconomic variables that are divided into six categories: cyclical indicators, price stability indicators, household indicators, corporate indicators, financial market indicators and external variables. Their methodology consists in running bivariate regressions using a single macroeconomic risk factor, with a lagged dependent variable and also a dummy variable to account for the change in the provisioning definitions for the analyzed period (1990 - 2001). Most prominent variables from each category are then selected to be used for stress testing purposes. These variables are industrial production, money (M1), Ifo business-climate index, real and nominal short-term interest rate, ATX index, DAX index, Euro STOXX index and exports.

Virolinen (2004) estimates a macroeconomic credit risk model for the Finish corporate sector using industry specific default rates that make it possible to have better results for the loss distributions than those obtained only with aggregated data. The logistic transformation is used for the industry specific default rates, whereas the preferred estimation method is seemingly unrelated regressions (SUR) that helps control for the contemporaneous correlation in the residuals. The results of the parsimonious model suggest a significant and quite robust relationship between industry specific default rates and key macroeconomic variables, in this case gdp, interest rates and corporate sector indebtedness. For agriculture though, explanatory variables appear to play only a marginal role in explaining default rates, whereas the importance of the interest rates varies across industries significantly. In order to perform stress test analysis, univariate autoregressive equations of order 2 (AR(2)) are used for the dynamics of the individual macroeconomic time series that explain default events. Lastly, Monte Carlo simulations are performed on a representative portfolio consisting of 3000 companies from all sectors of activity in order to obtain loss distributions for the corporate credit portfolio with a 1-year and 3-year time horizon.

Baboucek and Jancar (2005) use an unrestricted VAR model to empirically investigate the transmission mechanism between a set of macroeconomic variables describing the development of the Czech economy and the credit channel. The final model includes nine endogenous variables (real effective exchange rate, exports, monetary aggregate M2, imports, aggregate bank loans to clients, unemployment rate, consumer price index, domestic real three-month interest rate and nonperforming loans ratio as the share of non-performing loans in total bank loans). The model also includes seven exogenous variables, the more important being the two-week nominal interest rate (the Czech National Bank principal monetary policy tool) and another six dummy variable that control for breaks in the data. Finally, impulse response functions are used for stress testing purposes, so that the model constitutes a macroeconomic early warning system for the deterioration of the banks' loans quality. An advantage of the VAR approach is that it also captures feedback effects from credit risks to economy. Hoggarth and Zicchino employ the same technique to stress test the UK banking sector (2005).

Willem, Hoeberichts and Tabbae (2006) study credit risk for the Dutch banking sector by means of panel regression estimations with fixed effects that account for bank specific characteristics The contribution to the literature consists in the inclusion of a varying loss given default that results from estimating first the default rate dependent on macro variables and then loan loss provisions on macro variables and the default rate. Cross-border risks from foreign portfolios are also taken into account, as Dutch banks are also capital exporters. The results indicate gdp and interest rates to be the explanatory factors of default rates, while the different size of the fixed effects suggest that the Dutch banks have different sensitivities to macroeconomic developments, highlighting heterogeneity amongst banks in terms of risk profiles.

Jakubik (2007) develops a macroeconomic credit risk model for the Czech economy, using a latent one-factor Merton type model approach at the aggregate level. The number of defaults has a binomial distribution with conditional default probability to be estimated from the model on a one year time horizon and the given number of companies in the economy. The estimation of the model uses for the default variable the proportion of new bad loans in the total volume of loans in the economy. The default definition considered accounts for the classification of a loan as substandard or worse for the first time. There were various macroeconomic indicators tested for effects on the default rate, but finally only gdp, interest rates and inflation were included in the model. Following this, Jakubik and Schmieder (2008) employ the same framework, separately this time for the household and corporate sector, on two economies in different stages of development: Germany and Czech economy. Main findings of the study show that while the model is performing well for the corporate sector, it appears not to be as meaningful for the household sector. Another interesting outcome indicates that the variables that explain default rates for the two economies are more or less the same, even though the level and volatility of default rates have different paths: i) real exchange rate, inflation, gdp, credit to gdp ratio for the Czech corporates and ii) nominal interest rates, gdp, industry production and credit to gdp ratio for the German corporates.

3. Methodology

The methodology developed in this section follows the work of Wilson (1998) and Virolainen (2004) and tries to develop a framework to assess vulnerabilities stemming from the banking sector by building two distinct credit risk models, one for the household sector and the other for the corporate sector. This is a distinctive feature of the paper, as most of the work that has been done before regarding credit risk analysis is confined to the corporate sector only. In Romania the two sectors are equally important for the banking system, as their shares in nongovernment credit are very similar.

The essence of the model consists in linking the default events to macroeconomic factors in order to subsequently simulate future paths for the defaults rates by means of Monte Carlo methods and finally obtain values for the expected and unexpected losses for an arbitrarily portfolio from the loss distributions based on the actual macroeconomic environment. In addition, the model is also used to simulate the evolution of default probabilities in different scenarios concerning the dynamics of the macroeconomic conditions. An important point is that the estimated models will capture first round effects only of the macroeconomic shocks, and not necessarily be able to explain the complexities of feedback effects from the economy as a consequence of the potential financial sector distress. In order to address such an issue a bridge model would be needed linking the credit risk model with a macromodel, but at the moment this remains a challenge for future research.

Estimation of the models is performed on large portfolios of loans to ensure meaningful economic results. The analysis of the households' sector is conducted on a county level (a total of 42 counties in Romania), whereas for the corporate sector, industry specific portfolios are used for the five major business sectors (agriculture, industry, construction, trade and services).

First, average historic default rates are modeled with the logistic functional form that is extensively used in the literature to model bankruptcies. This logit transformation has the advantageous property that confines the default rates to the interval between 0 and 1. It is now widely accepted the idea that the relationship between default events and macro factors is non-linear, as the experience has shown that in high stress times extreme outcomes are more the rule than the exception (credit risk is by its nature not randomly distributed). That is represented as:

$$dr_{j,t} = \frac{1}{1 + \exp\left(y_{j,t}\right)}$$

where $dr_{j,t}$ is the default rate for the county / industry j at time t (depending whether the model refers to households or companies) and $y_{j,t}$ is a county / industry idiosyncratic macroeconomic index that stands as an indicator for the general state of the economy and whose parameters will be estimated. There is an inverse relationship between default rate and the state of the economy, represented explicitly through the macroeconomic index, in the sense that a better shape of the economy implies a higher $y_{j,t}$ and a smaller default rate $dr_{j,t}$.

Having the observed historic default rates available, but not the macroeconomic index we apply the inverse of the logistic function to the previous relationship and obtain:

$$y_{j,t} = \ln\left(\frac{1 - dr_{j,t}}{dr_{j,t}}\right) \tag{1}$$

Next, $y_{j,t}$ is assumed to be a function of various exogenous macroeconomic variables that determine the state of the economy. For the *household sector*, the model will be estimated by means of panel regressions techniques, in the form of:

$$y_{j,t} = \beta_0 + \beta_1 x_{1\,j,t} + \beta_2 x_{2\,j,t} + \dots + \beta_n x_{n\,j,t} + u_j + v_{j,t} \tag{2a}$$

where β_k is the set of coefficients to be estimated, $x_{k\,j,t}$ are the explanatory macroeconomic factors at time t that can be either specific to each county (eg unemployment, wages, degree of indebtedness) or common at the country level (exchange rate, interest rates, industrial production). There is a composite error structure, that consists of two parts: i) $v_{j,t}$, that is the traditional random error term being independent and identically normally distributed by assumption, associated with county j at time t, and ii) u_j that stands for the individual effects (random effects for this model, according to the Hauseman test performed to establish the correct estimation method) that allow for different intercepts among the 42 counties. This implies the existence of a structural default rate that varies across regions and that could be explained by a multitude of factors that generally are of a qualitative nature or can’t be easily quantified (the omitted variables problem). Among these factors one could mention the degree of education (schooling), financial culture or credit standards that banks apply according to their internal lending policy that is very different from county to county (in the sense that is much harder to be granted a loan in a less developed county than in Bucharest). This is an important advantage of panel methods compared to other estimation techniques and serves reasonably well the goals of this paper. On the other hand, a somehow restrictive assumption of the model is that the sensitivity of the default rates to common explanatory variables (like industrial production) is the same in different regions of the country, which is at least debatable.

The macro index for the corporate sector will be modeled by the use of a set of regression equations industry specific of the general form

$$y_{j,t} = \beta_{j,0} + \beta_{j,1}x_{1\,j,t} + \beta_{j,2}x_{2\,j,t} + \ldots + \beta_{j,n}x_{n\,j,t} + v_{j,t} \tag{2b}$$

where, again β_j are the coefficients to be estimated and $x_{k\,j,t}$ are the independent macroeconomic variables, that are either common for all business sectors (output gap, exchange rate) or particular to individual sectors (degree of indebtedness). The random error term $v_{j,t}$ is assumed independent and identically distributed. Unlike the specification of the households' model, the sensitivities of default rates to explanatory variables can vary across different industries and therefore individual specifications will be estimated.

The shared feature of the two specifications (for households and companies) is that they both can be regarded as a multi-factor model capturing the systematic non-diversifiable default risk as a weighed sum of macroeconomic variables plus individual shocks that are county or business sector specific.

Next step consists in describing the dynamics of the macro factors that have been selected in the specifications of the risk models for the two sectors, in order to simulate future paths for default rates (ie. probability of default). In the original methodology, Wilson (1998) indicates that AR(2) processes generally provide a fairly satisfactory fit for these macro factors. Though, this is not entirely the case with the actual data for the corporate sector. Therefore, to select the most appropriate ARMA (p,q) specification, the Box-Jenkins methodology together with information criteria are applied on the general form:

$$x_{j,t} = \alpha_{j,0} + \sum_{k=1}^{p} \alpha_{j,k}\, x_{j,t-k} + \varepsilon_{j,t} + \sum_{h=1}^{q} \varphi_{j,h}\, \varepsilon_{j,t-h} \tag{3}$$

where α_j are regression coefficients that will result from the estimation and $\varepsilon_{j,t}$ are the residuals assumed to be normally distributed as N(0,σ^2).

Equations 1 to 3 form together two systems of equations, one for each sector (households and companies), governing the joint evolution of the economic environment (through the chosen factors), industry/county specific default rates and also their associated structure of error terms E.

$$E = \binom{v}{\varepsilon} \sim N(0, \Sigma), \quad \Sigma = \begin{bmatrix} \Sigma_v & \Sigma_{v,\varepsilon} \\ \Sigma_{v,\varepsilon} & \Sigma_\varepsilon \end{bmatrix}$$

We have assumed so far that $v_j$and ε_j are serially uncorrelated and normally distributed, and their corresponding variance-covariance matrices are Σ_v and Σ_ε. An important point that must be made here is that matrix Σ_v is non-spherical (off-diagonal elements are non-zero) because of two reasons: i) default events in different counties/industries are highly correlated due to their common reactions to systematic risk factors (this is evidence from the data and testing procedures) and ii) in the companies case, there can be cross-sectoral dependence to macroeconomic shocks. In the same time errors $v_j$and ε_j are correlated because all the shocks on the macro factors are passed through the equations 2a and 2b. Therefore the variance-covariance matrix $\Sigma_{v,\varepsilon}$ is also non-spherical.

Finally, Monte Carlo simulations are performed based on the systems previously described in order to estimate a one year probability of default for the households and companies sectors. The simulation is conducted in the following manner: a) a vector of pseudo random variables (~N(0,1)) is generated and then transformed by the Cholesky decomposition of the variance-covariance matrix Σ into a vector of innovations that replicates the properties of the models' error terms, b) starting from some initial values of the macro variables and incorporating the simulations for the innovations in the models, simulated values will be obtained for macro factors, and subsequently for default rates. This procedure will be conducted 10.000 times for a one year ahead horizon in order to obtain the distribution of future default rates (hence the probability of default). Stress test scenarios of the macroeconomic conditions will also be conducted to observe the impact of default events on an arbitrarily build portfolio.

4. The Data

Default definition - general issues

The essence of this paper has at its centre default rates and tries to understand which are the main drivers of default on bank loans. Having said that, it is imperious necessarily to define the concept of default. The literature dedicated to studying credit risk shows the existence of various definitions for defaults, and highlights the difficulty of collection and sometimes interpretation of such data. There are three general concepts identified as the main categories of defaults:

i) legal definitions, also known as hard defaults, that refer to bankruptcies and solvencies and that are country specific. An important drawback that this approach poses to modeling is that court trials can be long-lasting and time-varying processes. This means that results obtained on data that are by nature heterogeneous in terms of time span can be misleading when drawing conclusions, even more so when analysis is conducted on different jurisdictions. Not to mention that in many countries, Romania included, there is no regulation in place regarding personal insolvency (that would protect households from creditors in case of impossibility of servicing the bank debt)
ii) payment incidents, also known as soft defaults, that have better defined characteristics and do not necessarily imply insolvency or bankruptcy. This refers mainly to companies that use payment instruments and are not able to fulfill the obligations that they commit to by issuing the instruments
iii) banking regulators definition specified by Basel II framework. This is regarded as the most comprehensive definition amongst the three and considers that a default has occurred when either or both of the following two events have taken place: a) the bank considers the obligor to be unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realizing security, and b) the obligor is past due more than 90 days on any material credit obligation to the banking group.

Therefore, the definition that will be used for the empirical work in next part of the paper will be the one suggested by Basel Committee. The reasons are multiple: a) it allows the use of the same definition both for the household and corporate risk models, hence insuring comparability of the two sources of risks on the banking sector, and b) creates scope for broader assessments of risks (possibly country wise).

Default rates

To correctly compute default rates, one has to account for the fact that sometimes for various reasons a debtor can enter the 90 days past due category and subsequently exit it and come with payments to date. So, in order to avoid counting the same debtor in default more than once, a sample of non-defaulters is built based on a previous one year record. The default rate is then easily computed at a county or industry level each period as the number of new defaulters in that period divided by the number of non-defaulters (with respect to the last year). The data used in constructing default rates for the two sectors (households and companies) are extracted from the Credit Risk Register database, that consists of all exposures of an individual or company that sum up to more than 20.000 Lei. In the case of households around 75% of total loans granted are part of this database, whereas in the case of companies the ratio is more than 95%.

Monthly data is used in order to model credit risk in territorial profile for households, starting from January 2006 till April 2010, which sum up to around 2000 observations (42 counties and 52 months). The model for companies is built with quarterly data because the number of default observations on a monthly basis is much smaller than in the first case and it would not be as meaningful. Quarterly data are available from q2 2006 till q1 2010.

The other sources of data are National Statistics Office and National Bank of Romania.

Household sector data

The explanatory variables of the model fall into two sub-groups: a) specific to each county, and b) common to all debtors in the country.

a) The degree of indebtedness of the population in a county is the ratio between the monthly debt service and monthly income. Monthly debt service is calculated using disaggregated data and accounting for individual characteristics of loans: i) type (consumer or mortgage loans), ii) initial maturity (short term, medium term or long term), iii) currency (leu or euro) and iv) remaining maturity expressed in number of months. It was assumed that repayment of loans takes the form of constant annuities and is determined using the formula:

$$payment = P * i * \frac{(1+i)^n}{[(1+i)^n - 1]}$$

where P is the outstanding value of the loan (principal), i is the monthly interest rate associated with each type of loan, currency and original maturity, and n is the remaining maturity expressed as number of months. The proxy used for the monthly income at county level is the total nominal net wages of the employees in each county, based on average wages and number of employees (information available from the National Statistics Office). Unemployment data on individual counties is also used.

b) Industrial production (% change, yoy) was chosen to account for the economic activity. One drawback of using this indicator as a proxy for the business cycle resides in the regional asymmetry of industrial development country wise. Inflation measured by the consumer price index (% change, yoy) was also tested for significance.

Nominal exchange rate should be a factor in the model as: i) more than half of the loans have been granted in foreign currency, ii) households credit expansion had very high growth rates in the periods of leu appreciation (around 80%, annual rates), and iii) the depreciation of the local currency exceeded 30% from the peak it reached in 2007. This altogether increases the pressure on those exposed to exchange rate risk.

Various measures of cost related variables, also used in other studies are empirically tested for significance in the model. Average interest rates charged on loans outstanding for the local and foreign currency stand for the total cost of credit. These tough, have two components: one that is the short-term market interest rate and the other a risk premium. In general, most loans are offered at a variable rate, which means that periodically the interest rate is reseted to reflect the dynamics of the money market interest rates and of banks aversion towards risk. These variable rates are linked to ROBOR 3M for local currency denominated loans, or EURIBOR 3M for the foreign currency denominated loans. Increased interest rates cause a higher debt service, that directly produce effects to default rates.

Unit root tests indicate that the dependent variable, industrial production and the degree of indebtedness are stationary.

Corporate sector data

Default data were available for the main six five industries: agriculture, industry, trade, construction and services (Appendix). Services sector is a broad category that also includes utilities and real estate services. Default rates exert a relative similar path, with the exception in construction. It is known that construction sector is the most sensitive to business cycle.

The output gap was used in order to capture the relationship between the corporate sectors defaults and the business cycle. This was estimated as the percentage deviations of gdp from potential output. Potential output was estimated with HP filter. This is also meant as a proxy for profits in each industry. Aggregate real gdp growth was also used in the estimation but provided poor results.

To measure corporate sector indebtedness, industry-specific variables are calculated by dividing the total loans outstanding of an industry by the annualized value added of that industry. This is to capture the financial leverage effects on default rates. The annualized value added deals with the seasonality problem of the series.

Finally, cost related variables are examined. As in the case of households most of the loans taken by companies are at a variable rate and linked to short-term market interest rates. An important part of them is granted in foreign currency. Therefore, the explanatory power of ROBOR 3M, EURIBOR 3M, nominal and real exchange rate or inflation is also analyzed.

Unit root tests indicate that the output gap and the degree of indebtedness are stationary.

5. Estimation, results and stress testing

Credit risk modeling for both sectors (households and corporations) is somehow restricted by the short series of relevant data. The analyzed period starting from 2006 doesn't comprise yet a full credit cycle. As long as the economy had been on an uptrend, the credit terms and standards were loose and the competition for market share amongst banks facilitated an important expansion of the non-government credit. Since the fall of Lehman Brothers tough, this tendency somehow has reversed and risk aversion increased considerably and also default events. During downturns, there are numerous factors that increase the pressure on those having bank debt. Intuitively, the first that react are those from the private sector, and that will be seen in the credit risk model of the companies where the lags of the explanatory variables are shorter than in the case of households. An explanation for this resides in the fact that the public sector adjusts slower during recessions. This also the case at the moment in Romania, with the government having to enforce important austerity measures (lay-offs from public administration, wage and pension cuts) in order to meet a balanced budget deficit after two years of too high deficits. This means that the pressure on the household sector will still last for a while. In these circumstances, it is possible that the model is biased more towards the good times and could underestimate the default rate for the last period of observations. So, an important drawback refers to the fact that data doesn't cover a full business cycle.

Main determinants that are used in the literature to explain the dynamics of default rates are: i) interest rates, ii) exchange rates, iii) the state of the economy measured by various variables, iv) level of indebtedness or v) inflation.

When building the model, a criteria followed was to try to include at least one variable from the five categories mentioned previously and to consider time lags that have a meaningful sense in explaining the attitude towards defaulting on bank debt. Another issue that has to be taken into consideration is the usage of a reasonable limited number of variables due to the degree of freedom considerations in the context of short time series. Having less variables would also be thoughtful in facilitating stress testing procedures. On the other hand, the previous consideration has to also account for any misspecification bias and ensure enough variables are included so as to provide the model with the best fit.

5.1 Household sector

The econometric tools used to estimate the credit risk model for the household sector involve panel techniques. The steps that must be followed begin by testing the null that the intercepts are equal. This would imply that there are no structural differences from county to county with respect to systematic default risk. If the null is accepted, the data are pooled and the estimation is done by means of simple OLS. If the null is rejected, a Hausman test is performed to see if the random effects estimator is insignificantly different from the unbiased fixed effects estimator. If the null is rejected, the fixed effects estimator is used instead, otherwise if the null is not rejected the random effects estimator is used. The random effects estimator has the advantage that it uses more information, in sense that controls for variation in the data, both within and between panels. This property makes it more efficient than the fixed effects estimator, being unbiased in the same time. The estimation method that allows for random effects is feasible GLS (general least squares).

The outcome of these tests indicates the presence of heterogeneity among counties and that the better method of estimation is by using random effects. Error testing also indicates the presence of contemporaneous correlation and cross-section heteroskedasticity. What this means is that the default rates are reacting together to shocks, but this reaction has different magnitude amongst counties. In order to compute robust standard errors, a White cross-section method for the coefficient covariance estimator will be employed. Serial correlation is not present at a 10% significance level, according to a Wooldridge test for autocorrelation in panel data.

The final model comprises unemployment, the degree of indebtedness, exchange rate, industrial production, and spreads charged by the banks over short-term markets interest rates both for the local and foreign currencies (mostly euro). Table 1 summarises the parameter estimates, their significance and also the reported adjusted R-squared and Durbin-Watson statistics. Non-stationary variables are introduced in first differences in the model.

Table 1 - Credit risk model for the household sector

Variables	Lag	Coefficient	Standard error	t-statistic
Constant		2.24 ***	0.50	4.41
D (Unemployment)	3	-0.22***	0.06	-3.67
Indebtedness	4	-0.60 ***	0.15	-4.05
D (Exchange rate)	6	-1.66 ***	0.42	-3.95
Industrial production	1	0.04 ***	0.004	8.90
D(Spread Euro)	12	-0.55 ***	0.13	-4.15
D (Spread Leu)	9	-0.03 **	0.01	-2.58
Adj. R-squared 0.71				
DW 1.81				

Significance level: * Significant at 10% level, ** Significant at 5% level, *** Significant at 1% level

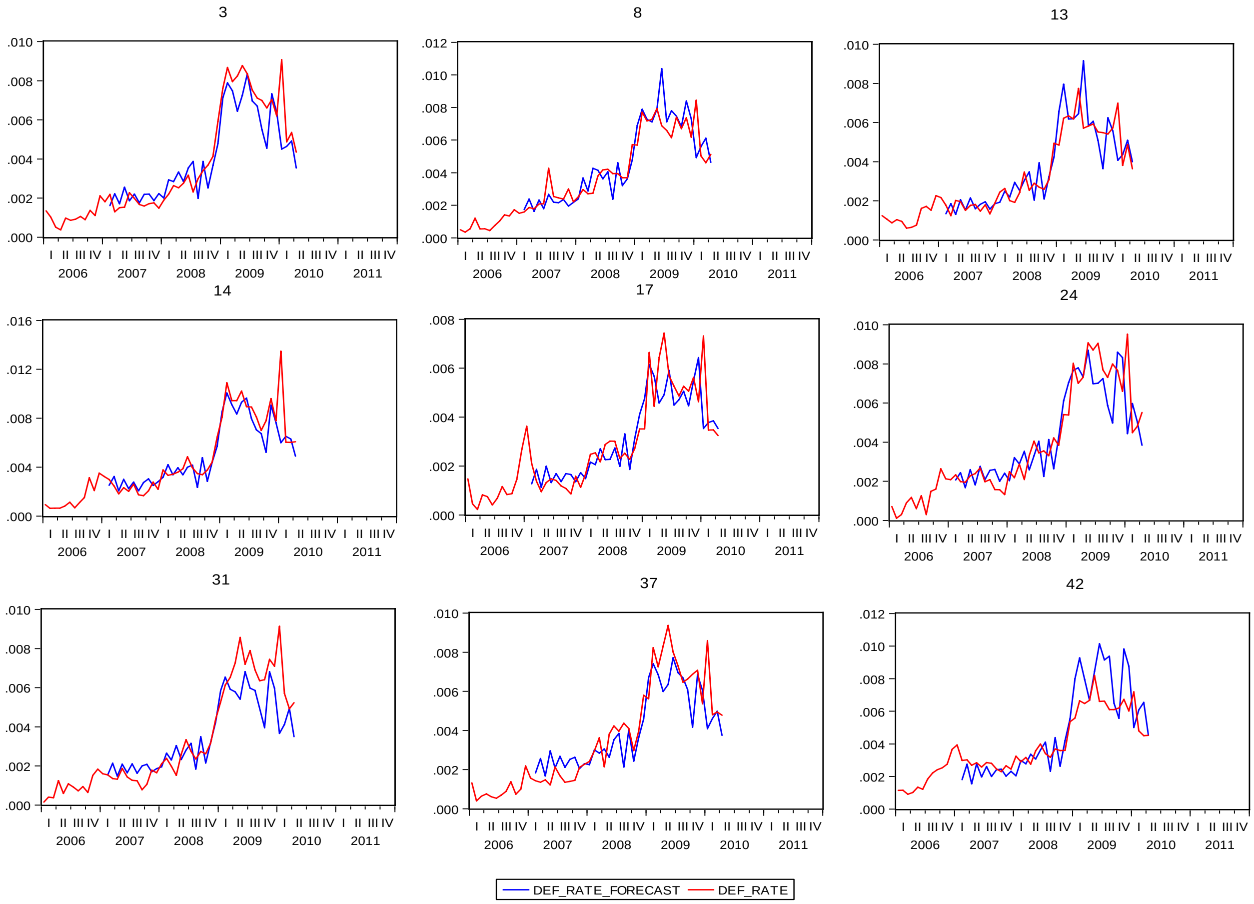
All variables are significant at a 1% level, except interest rate spread on local currency (at 5% significance level) and show the expected sign. There is an inverse relationship between default rates expressed by the logit transformation and the macro factors. Therefore a depreciation of the local currency or an increase in unemployment, interest spreads or indebtedness will result in a deterioration of default rates. In the same time, improvements in industrial production help in reducing it. The explanatory power of the model is satisfactory, as measured by R-squared (just above 70%).

Default rate seems to be explained by industrial production and the level of indebtedness on a longer term. Short time variation is also explained by interest rate spreads, unemployment and exchange rate. Comparatively, a 8% depreciation of the exchange rate has the same effect as a 1% increase in the spread for euro denominated loans. These two variables together seem to play an important part in explaining default rate. This is not surprising as the share of fx loans is more than 60% out of total loans granted to households.

The lag structure of the model has some interesting features. Having in mind the definition of default (90 days past due), the model suggests that unemployment has an immediate impact, in the sense that when an individual becomes unemployed, the next day will stop repaying the debt. In case of the degree of indebtedness and movements in the exchange rate, default events are prolongued with one and three months. A possible explanation is that for a few months, debtors might consider the situation to be temporarily and maybe reversible, and when that isn't the case the default occurs. Spreads exert longer lags because variable interest rates are commonly reseted every 3 to 6 months, depending on banks. Considering another couple of months to account for reasons related to expectations described previously, it makes the whole effect on defaults to take up to almost a year, also depending on the currency. Unlike the other variables, industrial production has a more contemporaneous effect, mainly due to the fact that it has the specific characteristic of being forward looking. An important advantage of the model resides in the lag structure, as it facilitates forecasting on shorter periods without having to make many assumptions about future developments of explanatory factors.

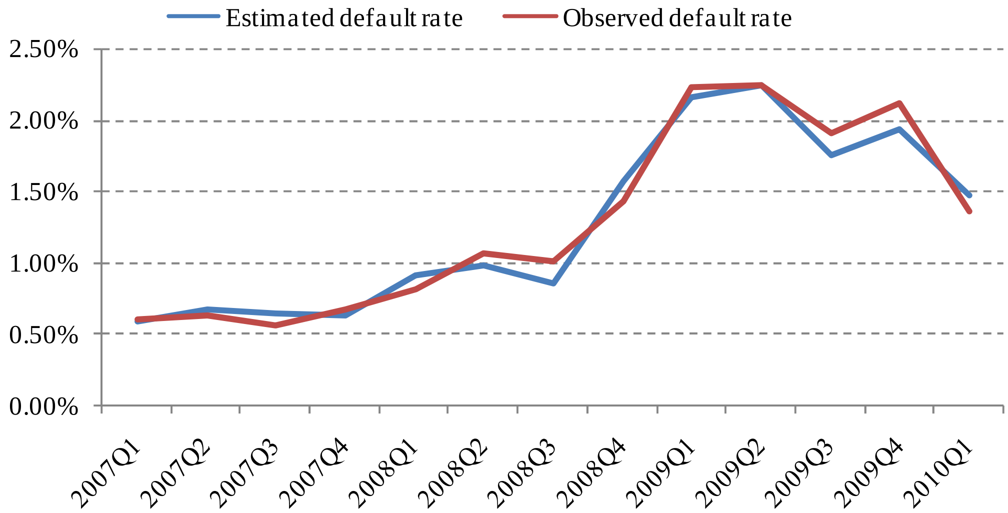
In order to see how the model performs an in-sample forecast was conducted. Figure 1 shows the behavior of the model for the first nine most important counties in terms of the volume of outstanding loans (they account for around 60% of total). One observation is that there is an increased volatility in the data for the second part of the analyzed period which makes the model have for some counties somehow more significant deviations from the actual data. In most cases default rates tend to be slightly underestimated by the model, especially during the last period characterized with high stress. There is though an important exception in Bucharest (panel 42) where the model seems to overestimate the actual default rates. This is also the case for Brasov (panel 8). This suggests that there is a higher capacity to withstand an adverse evolution of determinants of defaults. One possible explanation is that wealth and savings are higher in the capital.

Figure 1 - Performance of the model on a sample of nine counties (monthly frequency)



Another way to see the performance of the model is to compare aggregate default rates (at country level) with what the model would predict. The aggreated default rate is just the weighted average of counties' default rates. In order to eliminate some of the volatility in the data quarterly default rates were computed. Figure 2 indicates that when observing defaults on longer time periods (ie. at least quarters) the model performs reasonably well. For example, for the twelve months from may 2009 to april 2010, the default rate was 7,59% and the model prediction was 7,38%, approach that is in line with the scope of this paper (estimating the one year probability of default).

Figure 2 – Performance of the model on aggregated quarterly default rates, household sector



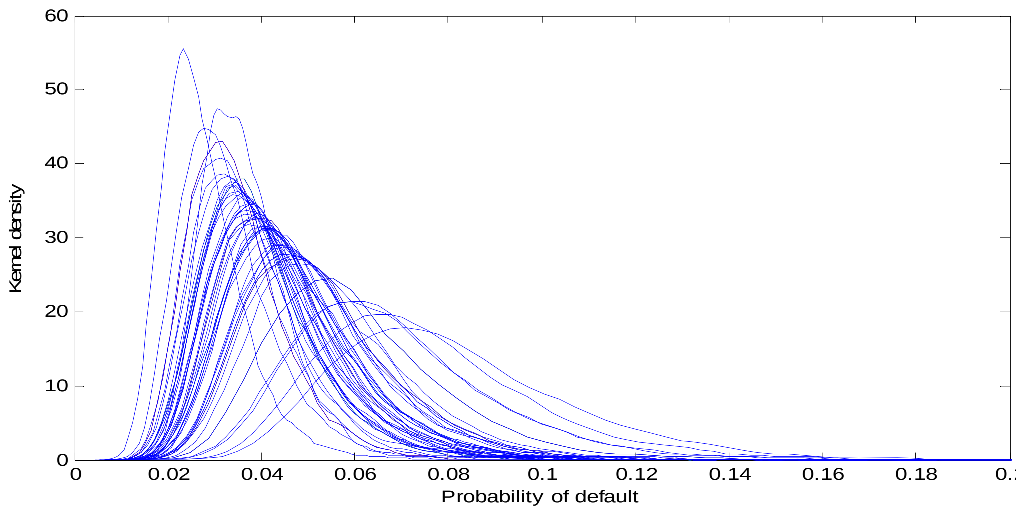
The last requirement before running the simulations is to add a dynamic component to the model by using univariate autoregressive equations to model the evolution of the macro factors. AR(2) specifications where chosen for the whole set of variables, as it provided the best results in terms of statistical significance (Tabel 2). AR(2) specifications are also most commonly used in the literature that deals with this kind of models (Wilson (1997) , Boss (2002), Virolainen (2004)). Covariance matrix of the error terms is also calculated. Correlation between the innovations of the micro factors is insignificant.

Tabel 2 - Estimation results for the ARMA models, household sector

	Exchange rate	**Unempl.**	**Ind prod**	**Spread Euro**	**Spread Leu**	**Indebtedness**
Constant	4.10	7.86	103.99	5.71	5.96	0.96
AR(1)	1.35	1.41	0.57	1.51	0.79	0.49
AR(2)	-0.38	-0.43	0.26	-0.54	0.09	0.35
Adj R^2	0.96	0.98	0.62	0.97	0.73	0.72
DW	**2.09**	**2.09**	**2.11**	**1.97**	**2.01**	**1.67**

The estimated system (credit risk model and AR(2) equations for the factors) will be used to simulate future county-specific default rates by means of Monte Carlo simulations. The system is solved stochastically. At each step in the simulation process a vector of pseudo random variables is drawn and transformed using the Cholesky decomposition of the initial variance-covariance matrix of the error terms, so as to obtain errors that replicate the correlations existing in the system. The procedure is iterated 10.000 times for a period of 12 months starting with may 2010. One year is the standard horizon used to observe default events. By cumulating the future default rates for each county, one year probability of default distributions will be generated (Figure 3).

Figure 3 - One year probability of default for the 42 counties



The resulting shapes of the distributions are skewed to the right, as expected, mainly because default rate is nonlinearly related to the macro variables. That means that the probability of default in times of high stress increases substantially more than in normal times. The probability of default varies in a range from 2,6% to 7,7% and indicates the heterogeneity of debt repayment among different regions in the country. The probability of default is approximated by the median of the distribution. Another observation that arises from the shapes of distribution is that the smaller the probability of default, the less uncertainty is about it (kurtosis gets smaller and smaller as probability of default increases). This is also related to the presence of nonlinearities in the model. Shocks will be amplified more in a riskier county.

The county individual probabilities of default will be used to determine the loss distribution of an arbitrary built portfolio. For the scope of this study and in order to avoid computational issues (working with too large datasets), this portfolio is formed by randomly selecting only 300 loans from the Credit Register Database for each county. It means the final portfolio will contain around 12,000 loans, out of a total of 1,3 million existing in the database. There are no other criteria used in the selection of loans, which should introduce bias towards the average size loans. This portfolio is not representative for the household sector.

The simulated probabilities of default are than applied to individual exposures in the arbitrary portfolio and its loss distribution is estimated over a one year time horizon. In determining expected and unexpected loss of the portfolio a fixed recovery rate is assumed throughout the simulations, so that LGD (loss given default) is set to 0.45. The expected and unexpected losses are estimated as the 50% and 99.9% percentiles of the loss distribution. The expected loss is calculated with the formula:

$$Expected\ loss = Exposure\ at\ default * Pd * LGD$$

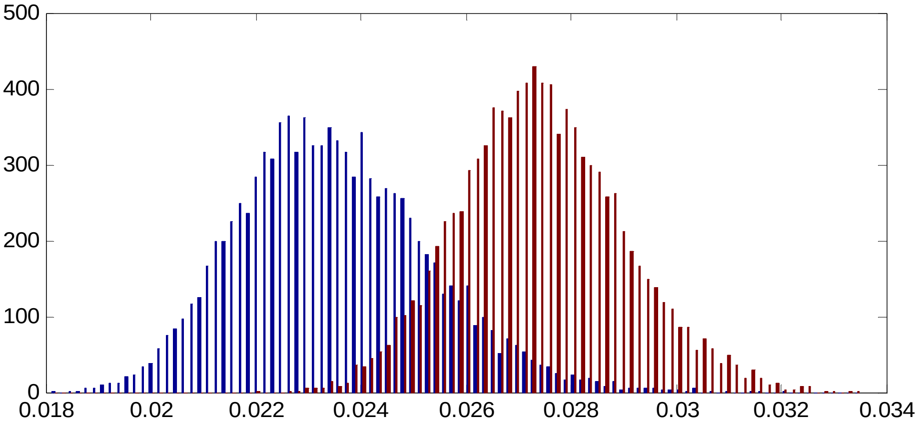
The AR(2) equations explaining the driving factors represent the baseline scenario used in simulations. They have the property of being mean reverting which for the baseline scenario seems reasonable. Therefore, an alternative scenario with a devaluing local currency and increases in unemployment is also considered to stress test the arbitrarily portfolio. This alternative scenario assumes that exchange rate will depreciate in the first two months to the levels of 4,5 and then 4,9 lei/euro and then keeping rather stable. This would imply a 17% depreciation of the local currency. In the same time unemployment will continue to rise as the public administration is expected to proceed restructuring the system. So, three consecutive jumps with 1,5%, 1% and 0,5% are introduced in the scenario followed by a constant 0,2% increase for the rest of the year.

The results depicted from the loss distribution in the two scenarios considered (figure 4), indicate that the expected loss of the portfolio in the baseline scenario is 2.34%, while the adverse scenario produces an additional increase of around 0.4%. The level of the loss is moderate. One reason is the lag structure. Spread interest rates have an impact on defaults after 9 and 12 months, which means that their evolution in the last 12 months will only produce effects during the next 12 months, such that their downward trend contributed positively to defaults.

Tabel 3 – The expected and unxpected loss for the arbitrary portfolio

	Baseline scenario	Alternative scenario
Expected loss	2,34%	2,73%
Unexpected loss (99,9%)	3,04%	3,25%

Figure 4 - Simulated baseline and adverse loss distributions of the arbitrary portfolio, one year horizon

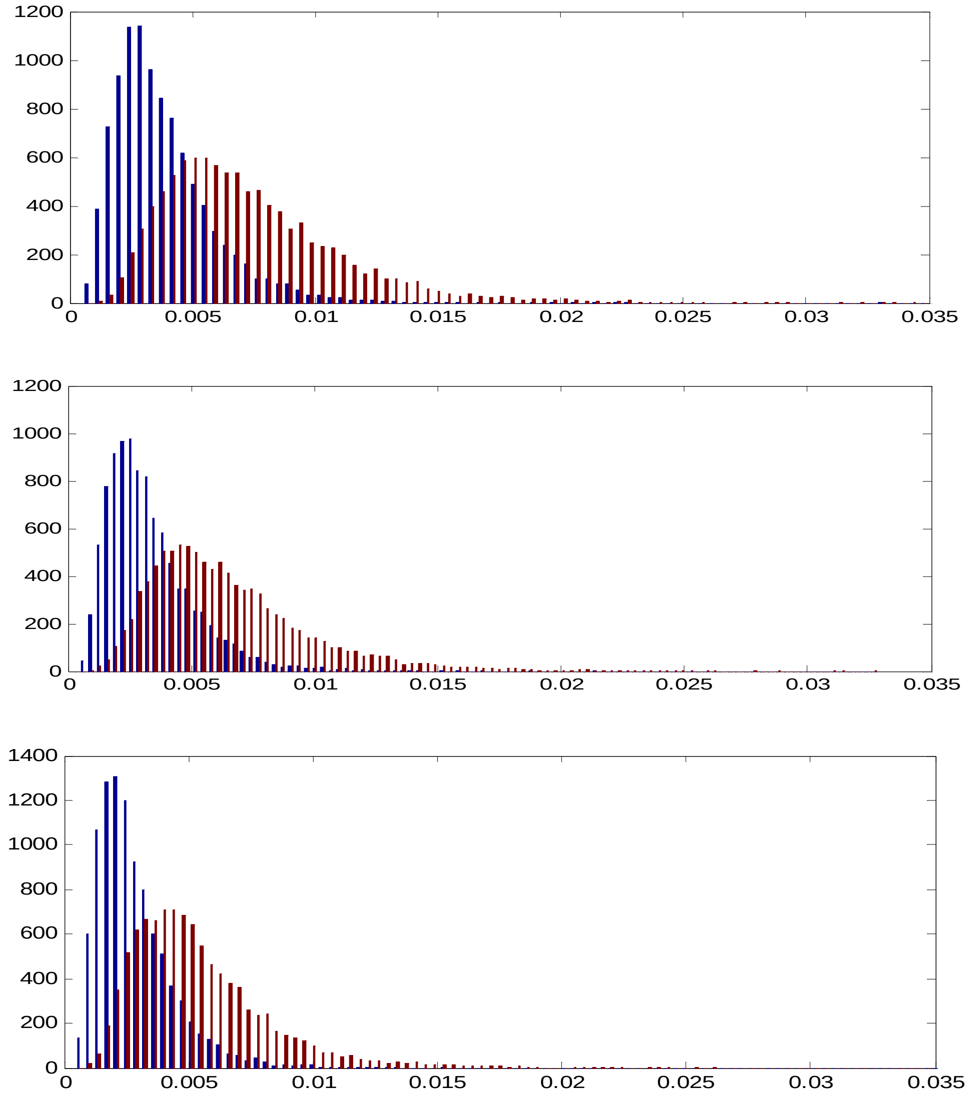


Replacing the AR(2) specification from the simulations with a specified path in the case of exchange rate and unemployment for the adverse scenario has contributed to a decreased uncertainty in the loss distribution.

The shape of the loss distributions has lost the asymmetry of probabilities of default because of the selection process of the loans in the arbitrarily portfolio. There is a bias towards the average smaller loans. The presence of large exposures in the portfolio would have resulted otherwise in high skewness and fat tail at the right end.

When cumulating defaults over a one year period the effect of a shock in certain months is not visible in the overall distribution. Figure 5 shows the effect of a 9% increase in the exchange rate on the probability of default distribution exactly in the month the model predicts its occurrence. This is highlighted in three counties from very different parts of the country (Neamt, Teleorman, Salaj) and produces an increase in the monthly default rate of 0.32%, 0.29% and 0.23%.

Figure 5 – Impact of a exchange rate shock on the probability of default in three counties Teleroman, Salaj and Neamt



5.2 Corporate Sector

The estimation of equations for the five industries' default rates was performed by means of seemingly unrelated regression (SUR). This method enhances the performance of the model by controlling for contemporaneous correlation among the default rates and is appropriate when the right hand side regressors are exogenous, as is in our case. One potential problem resides in usage of a static model with some non-stationary variables. Dynamic specifications were also tested but provided poor results, while cointegration methods can't be applied due to the short time period. Therefore, as in Virolinen (2004) which is confronted with the same issue, the static model will be used.

The results of the estimation are presented in table 4. The variables that have entered the final model account for the business cycle, financial leverage and exchange rates. Both output gap and real growth of gdp were tested, but output gap fitted the model better. The specification of the model is similar to that of Virolainen in terms of the variables used, with only one exception. The interest rate is replaced by the exchange rate. The degree of indebtedness is calculated as the ratio of debt of the specific industry to annualized value added by that industry.

Table 4 - Credit risk model for the corporate sector

	Agriculture	**Industry**	**Construction**	**Trade**	**Services**
Constant	7.97	10.98	10.61	9.23	8.77
	(11.52)	(28.08)	(13.07)	(15.52)	(16.27)
Output gap	1.49	4.31	4.86	6.48	5.24
	(1.01)	(7.24)	(2.86)	(4.37)	(4.33)
Indebtedness	-11.17	-17.34	-10.06	-8.20	-4.39
	(-2.77)	(-8.56)	(-1.65)	(-2.70)	(-2.11)
Exchange rate	-0.67	-0.85	-1.27	-0.80	-0.97
(-1)	(-2.71)	(-10.61)	(-4.01)	(-3.75)	(-5.09)
Adj R^2	**0.87**	**0.98**	**0.93**	**0.94**	**0.95**
SEE	**0.18**	**0.07**	**0.22**	**0.16**	**0.15**
DW	**2.13**	**2.43**	**1.67**	**2.03**	**1.32**

t-statistics in parantheses

All variables are statistically significant and have the expected sign. An increase in the output gap has positive effects on default rates (it lowers them), while an increase in indebtedness or a depreciation of the exchange rate translates into a higher default rate. Adjusted R squared indicates a good fit of the models. The Durbin Watson statistics and the Portmanteau test indicate that residual autocorrelation should not be problematic.

Different lag structures were tested, but it appears that the effect of output gap and indebtedness is contemporaneous, whilst the exchange rate has one lag. It is natural that output gap and corporate defaults move together, because on one hand more bankruptcies means less output, and on the other hand less output involves second round effects in the sense that more defaults occur due to the inter-linkages between companies and industries.

The financial leverage contains by its construction a lag component embedded in the series because the denominator represents the sum of value added of that industry for the last 4 quarters. Exchange rates impact the companies' debt servicing relatively similar to households where the lag was 6 months. Due to the fact that quarterly data is computed as the average of the quarter, it means that in terms of moths the lag can go from 1 to 6 months actually.

Output gap is more prominent for trade and services, but is not significant at all in explaining default rate in agriculture. A possible explanation relates to the fact that agriculture has only a minor contribution in total value added (around 7%).

Indebtedness is very important for the industry. During the last years many green field projects were developed and also old equipment of the existing companies was replaced by new technologies, all these investments requiring bank lending. Since our most important export markets have also been passing through a recession, the most vulnerable of these companies were affected.

Overall tough, it is clear that the companies react more quickly to adverse conditions than households in terms of nonperforming loans. There are two reasons for that: i) while companies benefit from the existence of a bankruptcy law, the households have no protection in this regard and try to service their debt even in tougher conditions up to the point where they are able to do it, and ii) in downturns real sector is the first to adapt, while the public sector where an important part of the households are working, always adjusts slower, even though eventually it does.

The advantage of this model consists in its relative simplicity. Only three variables are defining the macro index that influences the default rates, making stress testing scenarios easily to implement.

The dynamics of the macroeconomic conditions are explained with the help of ARMA specifications (table 5). For most of the variables, more than AR(2) specifications are appropriate.

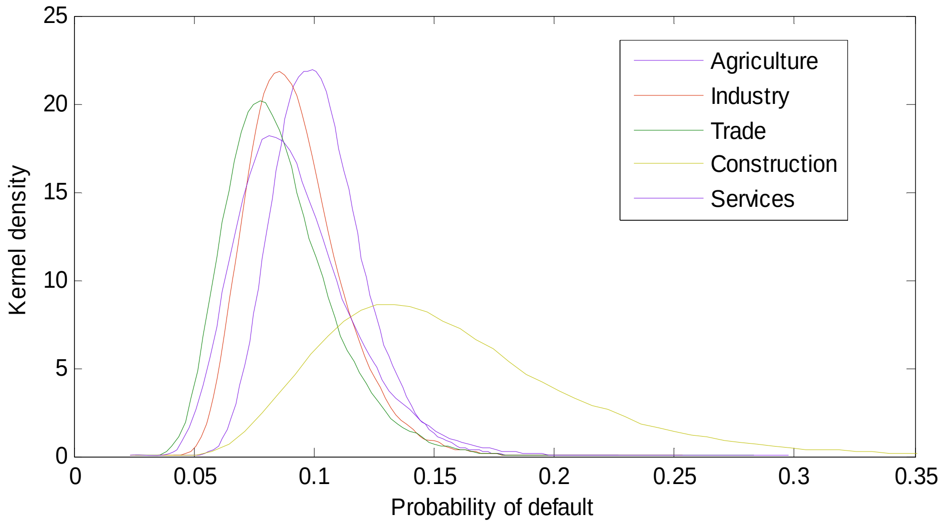
Tabel 5 - Estimation results for the ARMA models, corporate sector

	Output gap	**Exchange rate**	**Indeb. Agriculture**	**Indeb. Industry**	**Indeb. Construction**	**Indeb. Trade**	**Indeb. Services**
C	0.004	3.98	0.13	0.21	19.08	0.23	0.20
	(0.43)	(10.56)	(7.26)	(7.18)	(17.6)	(21.5)	(15.9)
AR(1)	1.65	0.74	1.34	0.47	1.05	0.53	1.09
	(8.34)	(3.32)	(8.26)	(7.17)	(6.39)	(7.01)	(5.55)
AR(2)	-0.86		-0.52		-0.26		-0.25
	(-3.79)		(-2.86)		(-2.70)		(-1.56)
AR(3)				-0.07			
				(-2.89)			
MA(1)		0.65	-0.77	-0.99		-0.74	-0.45
		(2.40)	(-6.76)	(-2.89)		(-6.27)	(-1.4)
MA(2)			0.88			-0.25	-0.54
			(7.83)			(-2.06)	(-1.35)
Adj R^2	0.89	0.78	0.88	0.91	0.77	0.94	0.88
DW	**1.74**	**2.19**	**2.14**	**2.42**	**2.26**	**2.07**	**1.96**

t-statistics in parantheses

The system of equations will be solved in a similar way as for the household sector. Monte Carlo simulations (50.000 this time) are used to generate future paths for default rates of the five analyzed industries over the next four quarters (from q2 2010 to q1 2011). By cumulating default rates over the four quarters every iteration, one year probability of default distributions will be generated for each industry (Figure 6).

Figure 6 - One year probability of default for the main economic sectors

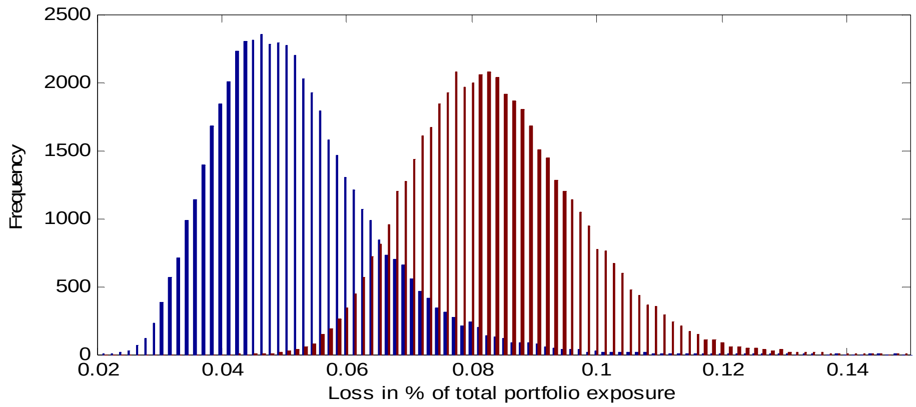


As in the case of households' model, the shapes of the distributions are asymmetric displaying fat tails to the right. The one year probability of default varies among the five industries from 8,3% to 14,8%, with the construction being the most risky sector, while the others are bellow 10%.

An arbitrarily portfolio will be build for the purpose of estimating its loss distribution. This portfolio will consist of 2500 loans, that are randomly drawn from each business sector in equal numbers. No other criteria are used in the selection process. The simulated probabilities of default for the six sectors are then applied to the individual loans in the arbitrary portfolio and its loss distribution is estimated over a one year time horizon. All other assumptions used for the household sector apply here.

Besides the baseline scenario for the macroeconomic environment an alternative scenario is also considered. To compare the results for the household and corporate sector, a similar adverse scenario is used in which the exchange rate is depreciating in the next two quarters to 4.5 and then 4.9 lei/euro.

Figure 7 - Simulated baseline and adverse loss distributions of the arbitrary portfolio, one year horizon



The results of the loss distributions in the baseline and adverse scenario (figure 7), indicate that the exchange rate risk for the arbitrarily portfolio is quite important in the case of non-financial companies vis-à-vis households. The expected loss increases with more than 3%, from 5% in the baseline scenario to 8,2% over a one year time horizon.

Tabel 6 – The expected and unexpected loss for the arbitrary portfolio

	Baseline scenario	Alternative scenario
Expected loss	4,99%	8,2%
Unexpected loss (99,9%)	10,97%	13,57%

6. Conclusions

The aim of this paper was to provide a framework for assessing credit risk distinctively for the household and corporate sectors in Romania, following a top-down approach.

It was found that determinants of defaults on bank loans are: a) unemployment, exchange rate, the degree of indebtedness, industrial production, and interest rate spreads charged by banks over the market interest rate for the household sector, and b) output gap, indebtedness and exchange rate for the corporate sector. Exchange rate and indebtedness are common explanatory factors, as the both sectors have accumulated substantial debt during the recent years, the biggest part being denominated in foreign currency.

The lag structure of the models indicates that the corporate sector is the first to react to adverse developments in the economy. Only exchange rates produce effects with a lag. The household sector appears to react somewhat slower to shocks. One possible explanation is the absence of a personal insolvency law in Romania to protect natural persons from creditors. Due to this individuals try to postpone the default event, hoping that shocks are temporary. This could be the case for exchange rate shocks. Another reason is that an important part of the population works in the public administration. Regarding the macro index as the overall state of the economy, and not only the variables that stayed in the final specification, it can then be said that the public administration adjusts slower during downturns in terms of restructuring than the real economy. This adjustment usually means layoffs and wage cuts, and therefore takes more time to affect defaults. The advantage of a longer lag in the model is that it is easier to perform a forecast.

Another finding of this study is that the probability of default over the next year, considering similar conditions, is higher for the corporate sector than for households. As expected, the riskiest of the business sectors is construction, whereas at the other end trade is the least risky.

Stress testing was applied to two arbitrarily build portfolios for each sector by considering exchange rate and unemployment shocks. The household portfolio was more resilient to adverse scenarios.

The relatively simple approach to modeling credit risk that was employed in this study presents both pros and cons. Advantages refer to the fact that having a small number of explanatory variables it is less costly in terms of the assumptions needed to perform forecasts for the probability of default. Embarking on a top-down approach is also less costly in terms of data constraints and even time. On the other hand, the disadvantage is that a macro perspective can't capture important details that are at a micro level only. The shortness of the series is another limitation of the model.

Another drawback of the estimated parsimonious models is that second round effects of financial distress to the economy are not captured. Bridge models that connect credit risk models with DSGE models could accommodate such an issue.

Bibliography

Alexander, C. and S. Elizabeth (2004), *The Professional Risk Manager's Handbook: A Comprehensive Guide to Current Theory and Best Practices*, PRMIA Institute

Baboucek, I. and M. Jancar (2005), "A VAR Analysis of the Effects of Macroeconomic Shocks to the quality of the Aggregate Loan Portfolio of the Czech Banking Sector", *Czech National Bank, Working paper 1*

Baltagi, B. H. (2005), *Econometric Analysis of Panel Data*, 3rd ed., John Wiley & Sons, West Sussex

Bangia, A., F. X. Diebold and T. Schuermann (2002), "Ratings Migration and the Business Cycle, with Application to Credit Portfolio Stress Testing", *Journal of Banking and Finance*, 26, 445-474

Boss, M. (2002), "A Macroeconomic Credit Risk Model for Stress Testing the Austrian Credit Portfolio", *Oesterreichische Nationalbank, Financial Stability Report 4*

Boss, M., G. Krenn, C. Puhr and M. Summer (2008), "Systemic Risk Monitor: A Model for Systemic Risk Analysis and Stress Testing of Banking Systems", *Oesterreichische Nationalbank, Financial Stability Report 11*

Bunn, P. and V. Redwood (2003), "Company Accounts Based Modeling of Business Failures and the Implications for Financial Stability", *Bank of England, Working Paper 210*

Cihak, M. (2007), "Introduction to Applied Stress Testing", *IMF Working Paper 59*

Endresz, M. V. and Z. Vasary (2008), "Macro Stress Testing with Sector Specific Bankruptcy Models", *Magyar Nemzeti Bank, Working paper 2*

Favero, C. A. (2001), *Applied Macroeconometrics*, Oxford University Press, Oxford

Green, W. H. (2000), *Econometric Analysis,* Prentice Hall, New Jersey

Hoggarth, G., S. Sorensen and L. Zicchino (2005), "Stress Tests of UK Banks using a VAR Approach", *Bank of England, Working papers 282*

Jakubik, P. (2007), "Macroeconomic Environment and Credit Risk", *Czech Journal of Economics and Finance,* 57(2), 60-78

Jakubik, P. and C. Schiemeder (2008), "Stress Testing Credit Risk: Comparison of the Czech Republic and Germany", *Financial Stability Institute, Award Paper*

Kalirai, H. and M. Scheicher (2002), “Macroeconomic Stress Testing: Preliminary Evidence for Austria”, *Oesterreichische Nationalbank, Financial Stability Report 3*

Kennedy, P. (2003), *A Guide to Econometrics*, 5th ed., The MIT Press, Massachusetts

Marcucci, J. and M. Quagliariello (2008), “Credit Risk and Business Cycle over Different Regimes”, *Banca D’Italia, Working paper 670*

Pesola, J. (2001), “The Role of Macroeconomic Shocks in Banking Crisis”, *Bank of Finland, Discussion Paper 6*

Sorge, M. and K.Virolainen (2006), “A Comparative Analysis of Macro Stress-Testing Methodologies with Application to Finland”, *Journal of Financial Stability*, 2, 113-151

Virolainen, K. (2004), “Macro Stress Testing with a Macroeconomic Credit Risk model for Finland”, *Bank of Finland, Discussion Paper 18*

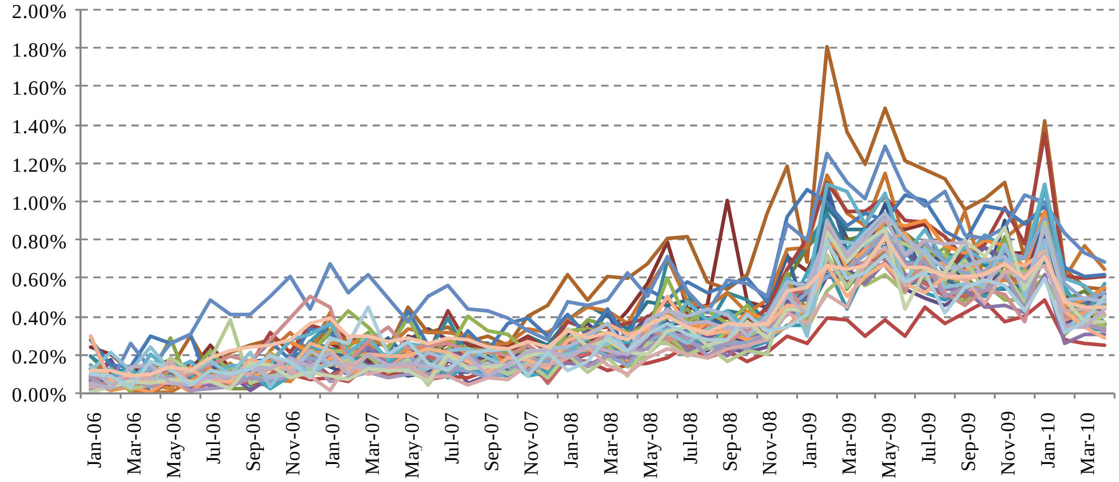
Willem, J., M. Hoeberichts and M. Tabbae (2006), “Modelling Scenario Analysis and Macro Stress-Testing”, *DNB Working Paper 119*

Wilson, T. C. (1998), „Portfolio Credit Risk”, *Federal Reserve Bank of New York, Economic Policy Review* 3

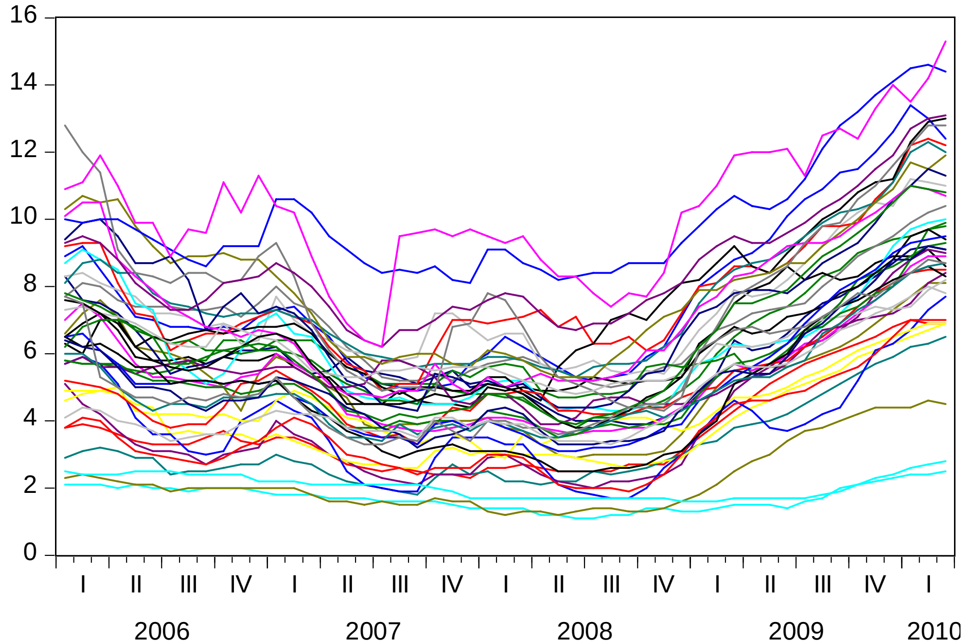
Wong, J.H., K. Choi and P. Fong (2008), “A Framework for Stress-Testing Banks’ Credit Risk”, *The Journal of Risk Model Validation*, 1, 3-23

Annex 1 – Household Sector

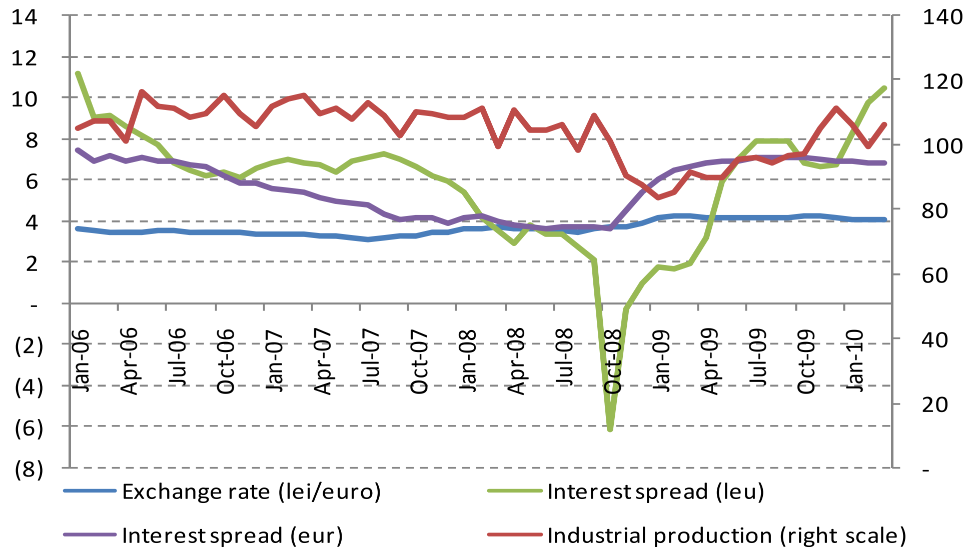
Monthly default rates at county level for the household sector (percent)



Unemployment rate at county level (percent)



Explanatory variables of the credit risk model



Hauseman test for panel estimation

```
. hausman fixed random

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+----------------------------------------------------------------
   L3D.unemp |   -.2158473    -.2153495       -.0004979        .0012429
  L4.debt_ser |     -.60104    -.6057552        .0047152        .0213106
L6D.exch_r~e |   -1.642384    -1.641909        -.000475        .0042672
   L.ind_pr_a |    .0392228     .0392084        .0000143        .0000905
L12D.sprea~r |   -.5487655    -.5481134       -.0006521         .003407
L9D.spread~u |   -.0304166    -.0303916        -.000025        .0002127
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(6) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        0.16
                Prob>chi2 =      0.9999
```

Residual Tests for Panel Estimation

```
xtcsd, pesaran abs
Pesaran's test of cross sectional independence =    86.401, Pr = 0.0000
Average absolute value of the off-diagonal elements =     0.478

Wooldridge test for autocorrelation in panel data
H0: no first-order autocorrelation
    F(  1,      41) =      2.480
           Prob > F =      0.1230
```

Estimation Results for the Credit Risk Model for the Household Sector

Dependent Variable: LOG((1/DEF_RATE)-1)
Method: Panel EGLS (Cross-section random effects)
Date: 06/24/10   Time: 14:44
Sample (adjusted): 2007M02 2010M04
Periods included: 39
Cross-sections included: 42
Total panel (balanced) observations: 1638
Swamy and Arora estimator of component variances
White cross-section standard errors & covariance (no d.f. correction)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.243536	0.508393	4.413000	0.0000
D(UNEMP(-3))	-0.216743	0.058985	-3.674561	0.0002
D(EXCH_RATE(-6))	-1.662185	0.420726	-3.950756	0.0001
IND_PR_A(-1)	0.039098	0.004388	8.909646	0.0000
D(SPREAD_EUR(-12))	-0.552203	0.132857	-4.156377	0.0000
D(SPREAD_LEU(-9))	-0.031326	0.012165	-2.575150	0.0101
DEBT_SER(-4)	-0.596054	0.147318	-4.046042	0.0001

Effects Specification

	S.D.	Rho
Cross-section random	0.204844	0.2873
Idiosyncratic random	0.322613	0.7127

Weighted Statistics

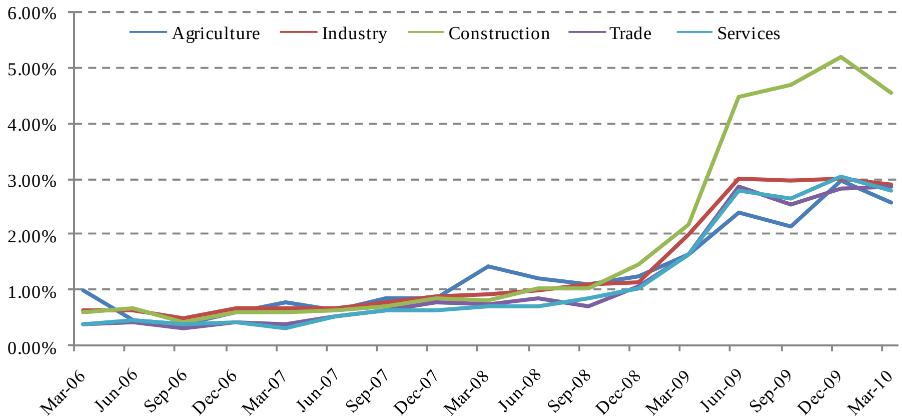
R-squared	0.712656	Mean dependent var	1.379118
Adjusted R-squared	0.711599	S.D. dependent var	0.600443
S.E. of regression	0.322456	Sum squared resid	169.5874
F-statistic	674.1887	Durbin-Watson stat	1.811271
Prob(F-statistic)	0.000000		

Unweighted Statistics

R-squared	0.648833	Mean dependent var	5.639802
Sum squared resid	234.0706	Durbin-Watson stat	1.312291

Annex 2 – Corporate Sector

Quarterly default rates for the corporate sector by industry (percent)



Estimation Results for the Credit Risk Model for the Corporate Sector

System: SUR
Estimation Method: Seemingly Unrelated Regression
Date: 06/29/10 Time: 13:37
Sample: 2007Q2 2010Q1
Included observations: 12
Total system (balanced) observations 60
Linear estimation after one-step weighting matrix

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	7.968200	0.691927	11.51595	0.0000
C(2)	1.487719	1.479978	1.005231	0.3208
C(3)	-11.17469	4.038035	-2.767358	0.0085
C(101)	-0.669086	0.247094	-2.707818	0.0099
C(4)	10.98608	0.391209	28.08239	0.0000
C(5)	4.311984	0.595787	7.237464	0.0000
C(6)	-17.34909	2.025522	-8.565247	0.0000
C(102)	-0.849937	0.080075	-10.61426	0.0000
C(7)	10.61132	0.811289	13.07958	0.0000
C(8)	4.863189	1.698903	2.862547	0.0067
C(9)	-10.06258	6.104966	-1.648262	0.1071
C(103)	-1.271111	0.317063	-4.009014	0.0003
C(10)	9.230262	0.594529	15.52534	0.0000
C(11)	6.479895	1.480599	4.376537	0.0001
C(12)	-8.195205	3.037284	-2.698202	0.0102
C(104)	-0.795655	0.211935	-3.754235	0.0006
C(13)	8.778585	0.539309	16.27748	0.0000
C(14)	5.236659	1.206691	4.339684	0.0001
C(15)	-4.394201	2.082482	-2.110079	0.0412
C(105)	-0.972321	0.190956	-5.091864	0.0000

Determinant residual covariance	9.41E-12

Equation: LOG((1/DEF_AGR)-1)=C(1)+C(2)*OUTPUTGAP+C(3)*IND_AGR +C(101)*CURS(-1)
Observations: 12

R-squared	0.905549	Mean dependent var	4.235727
Adjusted R-squared	0.870130	S.D. dependent var	0.493490
S.E. of regression	0.177841	Sum squared resid	0.253021
Durbin-Watson stat	2.137934		

Equation: LOG((1/DEF_IND)-1)=C(4)+C(5)*OUTPUTGAP+C(6)*IND_IND +C(102)*CURS(-1)
Observations: 12

R-squared	0.988429	Mean dependent var	4.213116
Adjusted R-squared	0.984090	S.D. dependent var	0.588349
S.E. of regression	0.074210	Sum squared resid	0.044057
Durbin-Watson stat	2.434552		

Equation: LOG((1/DEF_CON)-1)=C(7)+C(8)*OUTPUTGAP+C(9)*IND_CON +C(103)*CURS(-1)
Observations: 12

R-squared	0.953092	Mean dependent var	4.065928
Adjusted R-squared	0.935501	S.D. dependent var	0.849742
S.E. of regression	0.215805	Sum squared resid	0.372575
Durbin-Watson stat	1.671587		

Equation: LOG((1/DEF_COM)-1)=C(10)+C(11)*OUTPUTGAP+C(12) *IND_COM+C(104)*CURS(-1)
Observations: 12

R-squared	0.958390	Mean dependent var	4.386371
Adjusted R-squared	0.942786	S.D. dependent var	0.668016
S.E. of regression	0.159786	Sum squared resid	0.204253
Durbin-Watson stat	2.030519		

Equation: LOG((1/DEF_SER)-1)=C(13)+C(14)*OUTPUTGAP+C(15) *IND_SER+C(105)*CURS(-1)
Observations: 12

R-squared	0.966962	Mean dependent var	4.385999
Adjusted R-squared	0.954573	S.D. dependent var	0.683681
S.E. of regression	0.145718	Sum squared resid	0.169869
Durbin-Watson stat	1.318063		