МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2018
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2018
Proceedings of the Forty-seventh Spring Conference
of the Union of Bulgarian Mathematicians
Borovets, April 2–6, 2018

ON THE USE OF MACROECONOMIC FACTORS TO FORECAST PROBABILITY OF DEFAULT*

Vilislav Boutchaktchiev

A time series investigation of how the historic rates of default for corporate and retail credit exposures in three small Bulgarian banks is linked to the evolution of basic macroeconomic parameters. The result is a model which produces forecasts for probability of default based on economic outlook.

1. Introduction. The question of estimating the point-in-time probability of default for a credit exposure often arises in cases where a through-the-cycle rate of default is available.[1] Several studies, both in the academic (cf. e.g. [1], [4] or [3]) as well, as the professional literature (cf. [2]) have identified major challenges a bank would face when this technique of measuring the quantifiable credit risk is introduced for the first time.

The proposed time-series model studies the effect of the macroeconomic environment on the probability of default, splitting the exposures into two groups by "product" type, *corporate loans* and *retail*, each relating to GDP and unemployment rate, respectively. The model shows clear and significant dependence of the probability of default, from the macroeconomic variables, justifying our expectations, that economic growth correlates with decrease of cost of default.

The model is sufficiently sophisticated to catch some major economic effects postulated in the literature. For example an extended economic growth over a several periods leads to transference of decrease of PD for the corporate portfolio over future periods up to 18 months. Similarly the accumulation of the economic effect of decrease of unemployment over several consecutive periods leads to a "saturation" which manifests in stabilization of the PD for the retail portfolio.

The model's explicit form makes it accessible to be estimated by institutions with relatively short history of data as well, as low-default portfolios.

The quantitative parameters estimated in the model point toward some critical implications about the credit clientele typically served by the smaller bank.

Furthermore, I conjecture, that this technique can be implemented in calibration of probability of default curves for credit rating models, when the right assumptions are achieved (cf. [5]).

*2010 **Mathematics Subject Classification:** 62P20, 91B64.

Key words: Probability of default forecast, macroeconomic factors, time series modeling, IFRS 9 provisioning.

[1] The mandatory transition to IFRS 9 accounting standard raised this type of problem before the banks in Bulgaria in 2017. This was particularly burdensome for smaller institutions which have no established practice for standardized internal credit rating (IRB) which is made to reflect the fluctuation of the macroeconomic condition.

2. Data. The data used to estimate probability of default is propitiatory, selected from the portfolios of three small Bulgarian banks. Annual rates of default are computed at the end of every quarter over the period 2007q1 to 2016q4. Rate of default, r_t, is measured based on the size of the portfolio at the beginning of the period t:

$$r_t = \frac{N_t}{P_{t-1}}$$

Where N_t is the value of all exposures which were performing at the beginning of the period and became nonperforming at the end; P_t is the value of all exposures classified as performing at the end of the period t.

Loans which start the period as non-performing are excluded from this study, i.e., the effect of *cure rate* is disregarded.

The total population used varies between 2800 to 5300 expositions at each time valued between BGN 960 to 3980 million. At any given time, the corporate portfolio is between 9.2 to 12.7 times larger than the retail.

Table 1. **Data Summary.** Quarterly data for the rate of default for the two general portfolios — *Corporate* and *Retail*. Quarterly data of annual growth rate for the Real GDP and for Unemployment Rate. Seasonally adjusted data is used for the GDP; for Unemployment the data is smoothed using 4-lagged simple moving average

	Real GDP Growth Rate	Rate of Default (Corporate)	Growth Rate of the Unemployment Rate	Rate of Default (Retail)
N	38	39	38	39
min	.0035532	–.0659906	–.2590101	.011204
average	.0520655	.0174129	.022649	.0503441
max	1432881	.0710266	.600747	.1232733
std	0393395	.0275527	.2424151	.0009791

The total population of loans is split into two groups, according to "product" — *corporate* and *retail*. This is done according to the way the banks marketed their products to the clients. Basic investigation confirms that this separation is definitive. Essentially, it is in a bank's own interest that there is no reclassification of a given loan between the two categories. This is due, apart from various legal and accounting reasons, mostly to the expected source of repayment of the loan.[2]

Two kinds of outliers have been treated particularly. First, exposures that are disproportionately large have been excluded from consideration. A total of 9 loans at which at any time have exceeded, for corporate loans these include those above 2% of the total size of their portfolio; for retail the limit was 20%.

Second, a total of 3 periods with length 2-4 quarters were identified in which one of the banks have suffered corporate event of its own. These moments were marked with notably higher rates of default. To reduce the significance of these occasions these data points have been winsorized.

The macroeconomic data is publicly available from the Bulgarian National Statistical

[2]The population of loans exhibits very high degree of interrelation. Often a client with retail contract is a "related party" to a client with a corporate loan. This fact is a common practice for smaller banks in Bulgaria and its effect on PD would, no-doubt, be a topic of further investigation.

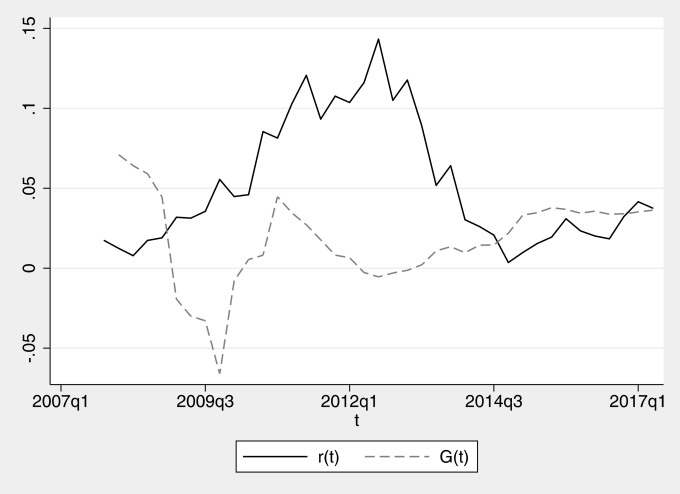

Fig. 1. **Time Series of Corporate.** $r(t)$ is the rate of default measured at the end of the period; $G(t)$ is the Real GDP growth rate. A negative correlation is visible over extended periods of time, but is not altogether statistically significant

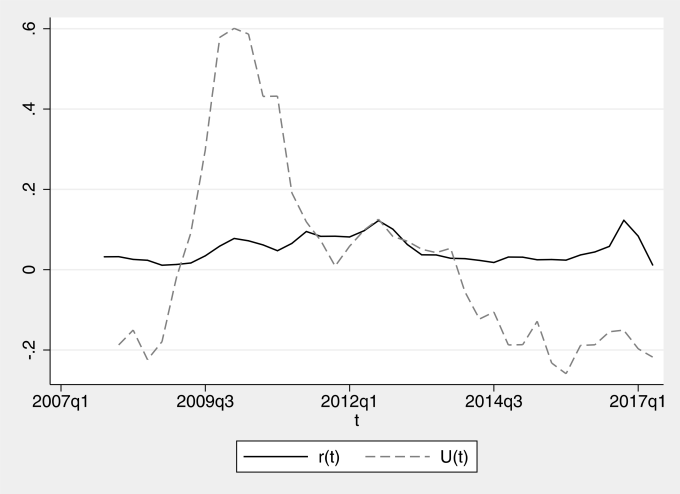

Fig. 2. **Time Series of Retail.** $r(t)$ is the rate of default measured at the end of the period; $U(t)$ is the annual growth rate of the unemployment rate.A positive correlation is visible over extended periods of time, but is not altogether statistically significant

Institute.[3] Two type of series were used.

- Real Gross Domestic Product. GDP is measured at prices of 2010. We used seasonally adjusted data, measured quarterly. The annual growth rate of the seasonally

[3] C.f. `http://www.nsi.bg/en` for variable construction.

adjusted GDP is denoted by $G(t)$.

$$G(t) = \frac{RGDP_t}{RGDP_{t-4}} - 1$$

- Unemployment Rate. We used the measurement at the end of each quarter. Furthermore, we smoothed the data series by calculating at each time point we took the average of the current and three previous data entries (4-period backward simple moving average). The annual growth of the unemployment rate, adjusted in this way is denoted by $U(t)$:

$$U(t) = \frac{UR_t}{UR_{t-4}} - 1$$

Figures 1 and 2 display the rate of default with the relevant macroeconomic factors, i.e., GDP and unemployment rate for corporate and retail, respectively. The summary statistics are provided in Table 1.

3. Model. The search for best predictive model was conducted under the following hypotheses.

1. The rate of default would correlate positively to the GDP in the corporate case. This assumption is based on the rationale that higher GDP is an indicator for better performing firms and, hence, larger capability to repay corporate loans. Similarly, in the retail portfolio would exhibit negative dependence on the unemployment rate since, by presumption, the main source for repayment of these loans is coming for the wages of the loan takers.

2. The rate of default r_{t_0} at time t_0, should be estimated based on the values of the macroeconomic factors measured at times $t < t_0$ and on forecasts at time t.

3. We expect that the rate of default is auto correlated. This mean, that the macroeconomic factor influences it directly, as well, as indirectly through the dependence of r_{t-i} from past macroeconomic conditions. It was our goal to isolate these two, and for this reason we were looking for an autoregressive model.

Notice, that if S_t denotes the information set available at time t, including the value of forecast of the macroeconomic factor x_t (whichever it is, in either of the two cases) then the *probability of default* q_t is nothing less than the forecast value:

$$q_t = E[r_t|S_t].$$

Corporate portfolio. We denote

$$\begin{aligned} y_t &= \ln r_t - \ln r_{t-4} \\ x_t &= G(t) - G(t-4) \end{aligned} \tag{1}$$

The variables x_t and y_t are stationary. Using the Box–Jenkins method we identify that autocorrelation of the two series is best accommodated by ARMA(1,1) model.

$$y_t = \beta x_t + \rho y_{t-1} + \theta \epsilon_{t-1} + \epsilon_t \tag{2}$$

Retail portfolio. Here the variable y_t is constructed in the same way as in (1), while x_t is used to denote

$$x_t = U(t) - U(t-4).$$

stationarity of the variables is verified using Dickie–Fuller test. The Box–Jenkins technique in this case suggest a simpler AR(1) model.

$$y_t = \beta x_{t-1} + \rho y_{t-1} + \epsilon_t \tag{3}$$

4. Estimation. The model was estimated using the language of Stata, which has an internal procedure for estimated ARIMA models using a maximum likelihood estimator.

The results are shown in Table 2. The coefficients have signs, as expected, and are statistically significant well beyond the 95% threshold. The Wald test for lack-of-fit in the case of Corporate has score $\chi^2(3) = 109.2$ and in the retail case $\chi^2(2) = 38.2$, producing p-value of satisfactory order.

Notice, that the x_t variable in the case of retail is lagged one period. We have tested the contemporary version which has a similar positive coefficient ($\hat{\beta} = 1.701$), with a smaller significance, however ($p = 0.04$).

Figures 3 and 4 show the juxtaposed values of the rate of default measured post factum and predicted a priori.

Table 2. **Estimation results.** The columns show the estimated values of the parameters in the following order $\hat{\beta}$, $\hat{\rho}$ and $\hat{\theta}$ (where applicable). The last column shows the p-value of the χ^2 test which concerns the altogether fit of the model. Beneath each estimated value is produced the respective p-value, based on z-score

Model	x_t	AR(1)	MA(1)	p-value for χ^2
(2)	-6.961092	.8643218	-.999993	0.0001
	(0.044)	(0.001)	(0.001)	
(3)	1.696857	.6970284	—	0.0001
	(0.009)	(0.001)		

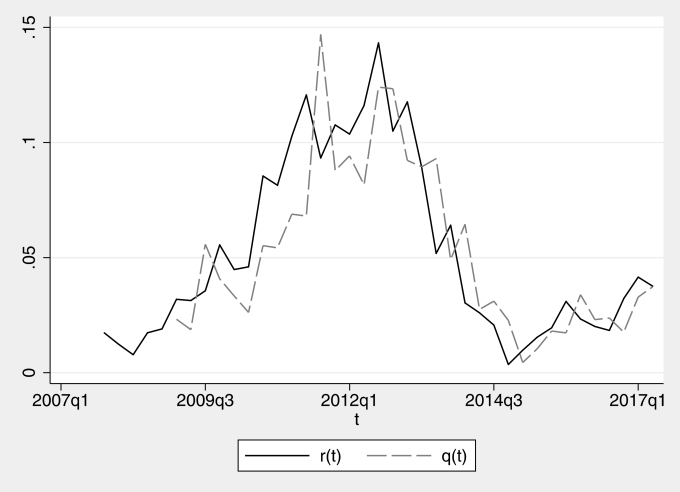

Fig. 3. **Corporate Probability of Default.** Rate of default r_t and probability of default q_t are shown on the same graph

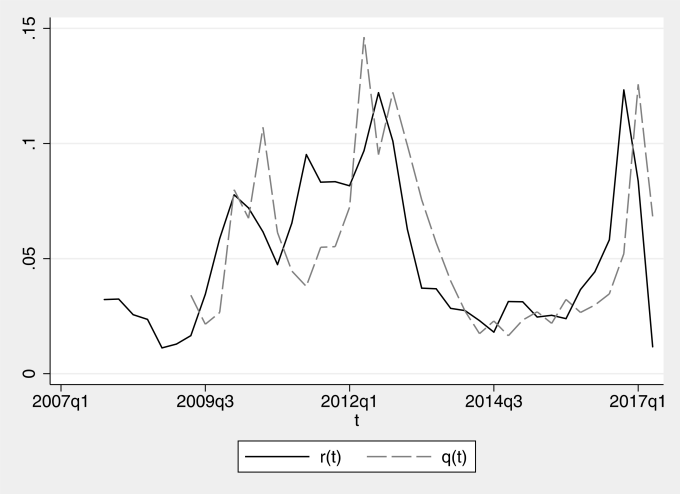

Fig. 4. **Retail Probability of Default.** Rate of default r_t and probability of default q_t are shown on the same graph

5. Results. Our test confirms the strongly significant dependence of the probability of default from macroeconomic conditions. In the case of corporate loans, the growth of PD is negatively related to the growth of GDP, while in the retail case a direct relation exists between the growth of PD and of the unemployment rate.

The lag of one quarter in the case of the retail model (3) is a phenomenon that may cause further interest. It may well be explained by the social legislation, which allows for unemployment benefits to be collected in the extend of up to 9 month in this country. Alternative explanation might come from the high cost of personal bankruptcy for these particular clients.

Furthermore, the result allows separation between the direct effect of the macroeconomic conditions from the general cyclical evolution of the time series of the probability of default. Consider for example the case of retail and denote $\hat{y}_t = \hat{\beta} x_{t-1}$ and $\tilde{y}_t = \hat{\rho} y_{t-1}$. Then the prognosis of the PD log-growth based on the information set S_{t-1} available at time $t-1$ can be expressed

$$E[y_t|S_{t-1}] = \hat{\beta} x_{t-1} + \hat{\rho} y_{t-1} = \hat{y}_t + \tilde{y}_t.$$

Here $\hat{y}_t$ represents the direct effect of the dynamics of the macroeconomic factor over the dynamics of the PD, and $\tilde{y}_t$ — the evolutionary component of the PD of the considered portfolio. Similar considerations apply to the corporate case.

The time series of probability of default exhibit very strong positive autocorrelation, which produces several notable results. First, this has the effect of long term inheritance of macroeconomic pressures. For example, if a sizable fall of unemployment brings down the rate of default at time $t-1$, this effect will be felt at time t. Second, positive autocorrelation might be an indication of the commonly acknowledged fact that the exposures in a small bank are strongly interrelated. Finally, it may suggest, moreover, a significant lack of diversification within the population of loans in a bank of this size, which brings it to outstanding credit risk.

REFERENCES

[1] S. Figlewski, H. Frydman, W. Liang. Modeling the effect of macroeconomic factors on corporate default and credit rating transitions. *International Review of Economics & Finance*, **21**, Issue 1 (2012), 87–105, ISSN 1059-0560.
https://doi.org/10.1016/j.iref.2011.05.004

[2] Global Public Policy Committee of representatives of the six largest accounting networks(GPPC), The implementation of IFRS 9 impairment requirements by banks. June 2006.

[3] M. Laurin, O. Martynenko. The Influence of Macroeconomic Factors on the Probability of Default, 2017.

[4] D. Simons, F. Rolwes. Macroeconomic Default Modeling and Stress Testing. International Journal of Central Banking, **5** (2009), 177–204.

[5] D. Tasche. The art of probability-of-default curve calibration. *Journal of Credit Risk* **9** (2013), 63–103. 10.21314/JCR.2013.169. https://arxiv.org/pdf/1212.3716.pdf.

Vilislav Boutchaktchiev
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Acad. G. Bonchev Str., Bl. 8
1113 Sofia, Bulgaria
e-mail: vboutcha@math.bas.bg

ВЪРХУ УПОТРЕБАТА НА МАКРОИКОНОМИЧЕСКИ ФАКТОРИ ПРИ ПРОГНОЗИРАНЕТО НА ВЕРОЯТНОСТ ОТ НЕИЗПЪЛНЕНИЕ

Вилислав Бучакчиев

Чрез изучаване поведението на времеви редове е изследвано влиянието на основни макроикономически показатели върху вероятността от кредитно неизпълнение. Извадката обхваща тримесечни данни в интервал от десет години за неизпълнението на кредити в три малки български банки. В резултат е произведен модел за прогнозиране на вероятност от неизпълнение въз основа на исторически данни и макроикономически перспективи.