Model of Credit Losses for Multiple Portfolios

Petr Gapko[1], Martin Šmíd[1]

Abstract

We propose a dynamic model of credit risk for multiple portfolios with two factors for each portfolio. The model estimates the interconnectedness of the portfolios and enhances the current research by introducing dynamics, incorporating the external (macroeconomic) influence and enabling to set realistic parameters of a loan portfolio. We follow the common approach that the credit risk on a loan portfolio can be decomposed into a probability of default and a loss given default and assume that both are driven by two underlying factors: one common for all borrowers in the portfolio and one individual for each single borrower. We suggest on a set of two large mortgages portfolios (one residential and one commercial) econometric estimation of the model and show realistically how the factors are related, how the portfolios are interconnected and how the credit risk is influenced by macroeconomic environment, which enables a comprehensive approach to stress testing.

Keywords: credit risk, mortgage, loan portfolio, dynamic model, estimation, interconnectedness, cointegration

JEL Classification: G32

Commented [H1]: Vsechny vety vyse v tomto abstraktu jsou srozumitelne. Jenom bych vice zduraznil novelty. Napsali jste co delate, ale pokud clovek nedela primo v teto literature, tak nepochopi, co je pridana hodnota clanku.

Commented [H2]: Proc suggest? Predpokladal bych, ze udelate odhad. Pak bych nepsal suggest, ale napsal bych We estimate.

Commented [H3]: Tohle je zvlastni spojeni. Show realistically. Proc by jste neco ukazovali nerelealisticky. Obecne ta posledni veta by si zaslouzila prepsat.

Commented [H4]: Obecne nerozumim, proc pouzivate kointegraci a zamerujete se jen na long-term relationship. Pokud zkoumate credit risk, tak vas zajima kratkodoba dynamika a patrne I dlouhodoby vztah. Vase regrese jsou ale jen o dlouhodobem vztahu.

[1] Econometric Department, Institute of Information Theory and Automation, Czech Academy of Sciences

1. Introduction

In 2009, when the financial crisis fully hit the US economy, losses from real estate loans in the US increased ten times, compared with the relatively quiet period ending in 2007. The aim of our paper is to investigate the risk factors, which drove the increase of delinquencies and charge-offs of residential and commercial real estate loans and estimate the interconnectedness of the residential and commercial mortgage portfolios.

To this end, we propose a credit risk model based on Merton's assumption that credit risk is driven by underlying risk factors (Merton, 1974). We derive a Merton-Vasicek type of the loss distribution (Vasicek, 1987) with several extensions. Our model converts default (or delinquency) rate and a real experienced percentage loss (or loss given default) on a portfolio into underlying factors. The default is underlined by a factor common for all borrowers in the portfolio and a factor specific for each single borrower. Similar assumption also holds for the loss given default (which in our framework is based on the value of an underlying collateral). Our proposed methodology might be compared with the approaches of Frye (Frye, 2000), Pykhtin (Pykhtin, 2003), Jimenez & Mencia (Jimenez & Mencia, 2009) or Witzany (Witzany J. , 201[illegible]e common factors can be easily explained as an influence of external environment (e.g. macroeconomic situation) and the individual factors might be interpreted as an individual's ability to maintain asset base (or, more naturally, to earn money) and a specifics of a collateral, respectively. Additionally, we show how the credit risk of multiple portfolios can be assessed jointly, taking into account the interconnectedness of the portfolios.

Our approach brings a further extension of the abovementioned frameworks in three ways. First, we introduce dynamics in the underlying risk factors. Second, we use the cointegration analysis to find a relationship between common factors and the external (macroeconomic) environment. Finally, we introduce a multi-generation approach, i.e. our model is capable of capturing individual parameters of individual vintages of borrowers. Also the approach to the estimation of the interconnectedness of multiple portfolios is an enhancement.

We estimate the model on a dataset of US nationwide residential and commercial real estate loan portfolios 30+ delinquencies (loans more than 30 days past due) and charge-off (net charge-offs of loans from books) rates. The cointegration analysis of underlying risk factors and macroeconomic variables such as GDP, unemployment, house price index, etc… clearly shows that the risk performance of the examined portfolio is linked to the macroeconomic situation. Also, the cointegration analysis reveals that the two analyzed portfolios performance evolves dependently.

The recent research clearly proved that there is an obvious relationship between the state of the economy and the credit risk. Hamerle et al. (Hamerle, Dartsch, Jobst, & Plank, 2011) showed on a bond portfolio the necessity of taking into account changes in macroeconomic environment. Similarly, (Sommar & Shahnazarian, 2009) used the vector error correction model to estimate the dependency of expected default frequency of a portfolio of nonfinancial listed companies on several macroeconomic factors, from which they found the most influencing the interest rate. The mentioned results are in line with the findings of Pesaran et al. (Pesaran, Schuermann, Treutler, & Weiner, 2003) or Virolainen (Virolainen, 2004).

Commented [H5]: Neni to spatny zacatek, ale to spojeni je trochu zavadejici jako by implikovalo, ze az v roce 2009 byla US economy plne zasazena. Nekdo by ale mohl argumentovat, ze byla plne zasazena uz v 2008 – fall of Lehman Brothers.

Commented [H6]: To je zbytecne. Neni jasne relatively to what.

Commented [H7]: Tohle je dobry odstavec, protoze hned rikate na zacatku paperu, co delate.

Commented [H8]: Tohle vylepsit, toto slovni spojeni je hrozne obecne. Ke vsemu nejak jde se porovnat. Chcete rict, ze tyto papery maji conceptually similar framework?

Commented [H9]: Pryc se situation, misto toho conditions.

Commented [H10]: Tohle vic rozepsat. Myslite tim dynamics, ze tam chybi lagged dependent variable v tech modelech? Nebo neco jineho? Je to nejasne, protoze hned dalsi veta mluvi o kointegraci, kde zadna lagged dependent variable v kointegracnim vztahu neni.

Commented [H11]: Cointegraci dela ale I Pesaran.

Commented [H12]: Tento odstavec je naprosto klicovy, udelal bych z nej 2 odstavce. Nejsem si jisty, zda kointegraci v tomto odstavci zminovat, protoze mi prijde, ze uz byla pouzita Pesaranem. Spis bych psat, ze paper ma nekolik teoretickych extensions vuci predchozi literature but importantly tests the theory on empirical data using cointegration methods.

Commented [H13]: Tohle neni spatne, ale kazdy to vi a urcite na to byla rada empirickych odhadu udelana. Nedalo by se to psat, ze we test the predictions from our theoretical model on data. Jenom kolega z IES ma na to 2 papery:
https://ideas.repec.org/a/kap/openec/v26y2015i3p525-550.html
https://ideas.repec.org/p/fau/wpaper/wp2014_26.html

Paper je obecne velmi malo navazany na existujici literature o credit risku a jeho determinantech.

The paper is organized as follows. In the following section we provide a description of the model methodology and comparison with several similar models. In the section 3 we describe the empirical analysis, data, which the analysis was based on, and our results. Finally, the section 4 concludes.

2. The Model for a Single Portfolio

As it was already premised, we consider a multi-generational (hypothetical) portfolio of loans (with the same amount lent for simplicity). In particular, we assume that N^τ new loans arrive into the portfolio at each time $\tau \geq 1$. The duration of the loans is assumed to be 30 years.

In line with (Vasicek, 1987) we assume that the wealth of the i-th debtor from the τ −th generation of loans (i.e. of those with the first repayment at time τ) follows a "trended" Geometrical Brownian Motion with a common factor as a "trend", i.e.,

$$A_t^{\tau,i} \exp\{Y_t + Z_t^{\tau,i}\}, \qquad t \geq \tau,$$

where

- Y (common factor) is an general stochastic processes
- $Z^{\tau,i}$ (indivudual factor) is a stochastic process such that
  - $Z_\tau^{\tau,i} \sim N(0, \sigma_1)$ for some $\sigma_1 > 0$,
  - $Z_t^{\tau,i} = \phi Z_{t-1}^{\tau,i} + U_t^{\tau,i}, U_t^{\tau,i} \sim N(0, \sigma)$, $t > \tau$, for some constants $\phi \in R,\ \sigma > 0$

Analogously to the wealth and similarly to (Frontczak & Rostek, 2015) we assume that the price of a collateral securing the i-th loan from the τ-th generation fulfils

$$P_{\tau-1} = 1, \qquad P_t^{\tau,i} = \exp\{I_t + E_t^{\tau,i}\}, \qquad t > \tau,$$

where

- I is another common factor
- $E^{\tau,i}$ is a stochastic process fulfilling $E_t^{\tau,i} = \psi E_{t-1}^{\tau,i} + V_t^{\tau,i},\ V_t^{\tau,i} \sim N(0, \mathrm{\rho})$, $t > \tau$, for some constants, $\psi \in R$ (here, $E_\tau^{\tau,i} = 0$ by definition).

As it is usual in factor models, we assume that

- $U_1^{1,1}, V_1^{1,1}, U_1^{1,2}, V_1^{1,2}, \ldots U_1^{1,N^1}, V_1^{1,N^1}, U_1^{2,1}, V_1^{1,2}, \ldots V_T^{1,N^1}, U_1^{2,1}$ ... are mutually independent, independent of Y, I.

We say that the i-th loan from the τ −th generation defaults at time t if the wealth of the corresponding debtor does not suffice for repaying the mortgage, namely if

$$A_t^{\tau,i} < (t - \tau + 1)b$$

where b the installment. The corresponding percentage loss of the creditor then equals to

$$G_t^{\tau,i} = \frac{D_t^{\tau,i}\max(0, h_t - P_t^{\tau,i})}{h_t^\tau}$$

where $D_t^{\tau,i}$ is the zero-one variable indicating the default and h_t^τ is the principal outstanding at t.

The overall default rate, charge-off rate, respectively, are then defined as

$$Q_t = \frac{\sum_{1\le\tau\le t}\sum_{1\le i\le N^\tau} D_t^{\tau,i}}{\sum_{1\le\tau\le t} N_t^\tau}, \qquad t > 1,$$

$$G_t = \frac{\sum_{1\le\tau\le t}\sum_{1\le i\le N^\tau} G_t^{\tau,i}}{\sum_{1\le\tau\le t} N_t^\tau}, \qquad t > 1,$$

respectively, where for each t, N_t^τ is the number of debts from the τ-th generation which did not default until t.

For more rigorous description of the model, see (Šmíd, 2015) or (Gapko & Šmíd, 2012).

3. Dependency among portfolios

In order to predict the future loss in a multi-portfolio world, the dependency among the portfolios has to be considered. In our case, we consider two portfolios, both consisting of real estate loans (mortgages). One is a portfolio of residential mortgages and the latter commercial mortgages. The interconnectedness between these two portfolios might arise from several root causes:

- Real estate development, which is commercial by its nature, but translates into residential mortgages once the development is completed
- Entrepreneurial mortgages, which might be included in the commercial as well as in the residential segment
- Interconnectedness of the commercial and residential real estate markets
- Interconnectedness between companies' and individuals' credit performance (residual dependency, which is not explained by macroeconomic conditions)

As from the previous section we know that default rates and the percentage loss on a portfolio can be translated to underlying risk factors. Thus interconnectedness between the portfolios can be explained by the interconnectedness of the underlying factors. Therefore, once the factors are extracted, we then examine the dependency structure among the factors (common and individual for both portfolios).

4. Data and Empirical Estimation

4.1 Data description

The dataset used consists of four time series, namely residential and commercial mortgage delinquency rates, which are proportions of loans more than 30 days past due (30+) on the total balance, and residential and commercial mortgage charge off rates, which are proportions of charged off loans (net of recoveries) on the average total balance. The dataset was downloaded from the United States Federal Reserve System and thus includes the US nationwide statistics. The time period covered ranges from 1991 to 2014 in a quarterly granularity. Thus the number of observations used was 96.

Table 4.1 and Figure 4.1 summarize descriptive statistics and show the development of the input data. The 30+ delinquency rates were used as proxy metrics for default rates and the charge-off rates represent real losses from the unpaid balance. From the Figure 4.1 it is obvious that the time series are correlated. Also, the recent economic crisis, which started in the US in late 2007 and impacted the US mortgage and real estate markets excessively is visible, as all time series rocketed up to multiples of their preceding values between 2007 and 2010.

Statistic	**30+ delinquency rate commercial**	**Charge-off rate commercial**	**30+ delinquency rate residential**	**Charge-off rate residential**
Mean value	0.040722	0.009785	0.041660	0.005099
Median	0.026900	0.003433	0.023100	0.0017
Minimum	0.010200	0.000914	0.013900	0.0007
Maximum	0.12060	0.036297	0.11270	0.0277
Standard deviation	0.031489	0.010653	0.032455	0.0068593
Variance	0.77327	1.0888	0.77903	1.3452
Skewness	1.0418	1.0852	1.1517	1.7144
Curtosis	-0.11868	-0.33250	-0.38216	1.6548
5% percentile	0.010670	0.001345	0.015785	0.0008
95% percentile	0.11055	0.031084	0.10596	0.02177

Table 4.1: Descriptive statistics of input data

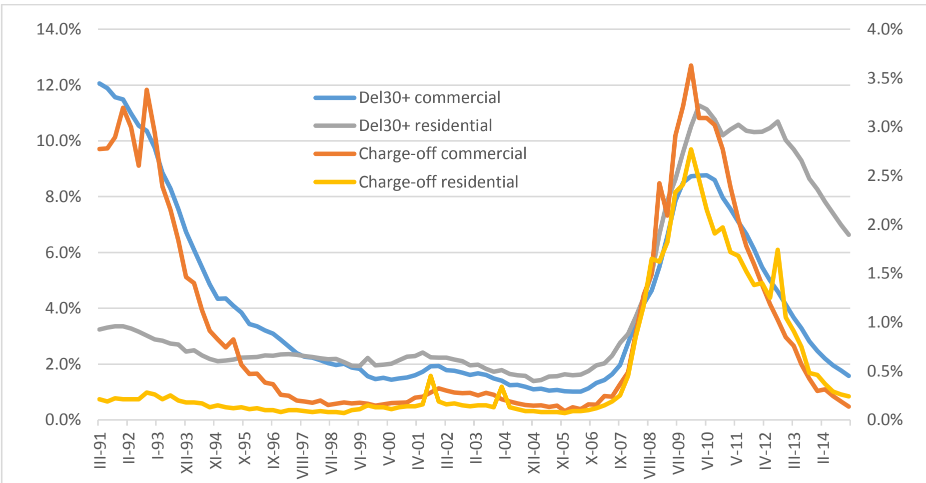

Figure 4.1: Development of the 30+ delinquency rates (left axis) and charge-off rates (right axis)

4.2 Underlying factors extraction

As the individual factors disappear according to the law of large numbers in our model, the only drivers of the default rate and the loss given default remain their respective common factors Moreover, according to (Šmíd, 2015) theorem 3, there exists a one-to one mapping between $(Y_1, I_1, Y_2, I_2, \dots Y_t, I_t)$ and $(Q_1, G_1, Q_2, G_2, \dots Q_t, G_t)$ for each $t \geq 1$. Therefore, the default rates and the charge-off rates can be uniquely translated into the underlying factors. We applied this transformation to our input data with the following parameters:

- The (sigma1)… standard deviation of the initial wealth was set to 0.5
- The (rho)… standard deviation of the loss given default was set to 0.12 according to (Gapko & Šmíd, 2012a)
- The (phi)… autocorrelation factor of the default rate individual factor was set to 0.8, which corresponds to findings in (Hochguertel & Ohlsson, 2011)
- The (psi)… autocorrelation factor of the loss given default individual factor was set to 0.1
- The mortgage interest rate was set to zero for simplicity

The resulting time series of the extracted common factors Y (default rate) and I (loss given default) for both commercial and residential mortgage portfolios are illustrated on the Figure 4.2.

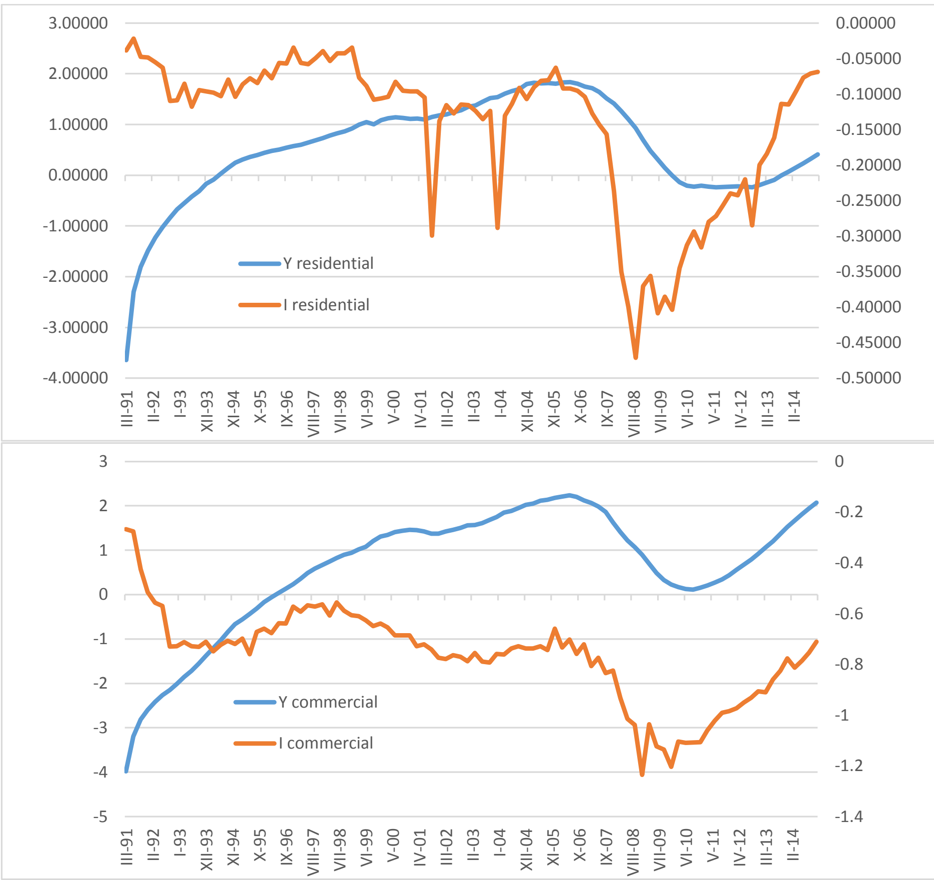

Figure 4.2: The development of the extracted common factors Y (default rate, left axis) and I (loss given default, right axis)

4.3 Estimation of factors and portfolios interconnectedness and external influences

The interconnectedness of the four common factors, namely Y residential (Yr), I residential (Ir), Y commercial (Yc) and I commercial (Ic), was studied. Common sense advises that the default rates (or, alternatively, PDs) and the charge-off rates (or, alternatively, LGDs[2]) should be positively correlated, i.e. in the times of economic expansions, when both the financial situation of debtors and real estate (thus collateral) prices improve, one would expect decreasing default rate and also decreasing LGD. Also, the expectation should be vice versa for the periods of economic downturns.

[2] LGD can be easily obtained by dividing the charge-off rate by the default rate

To examine the mutual relationship between the factors Yr, Ir, Yc and Ic, we used cointegration analysis, starting with a test for their cointegration. We ran the Johansen cointegration test, which showed that the cointegration of the rank 1 is present.

Further, we estimated the interconnectedness of Yr, Ir, Yc and Ic and their dependence on external environment by the VECM model, with Yr, Ir, Yc and Ic being endogenous variables and a set of macroeconomic indicators as exogenous variables. We considered GDP, house price index (HPI), consumer prices (inflation), FED base interest rate, industrial production, unemployment and personal income as representatives of the external environment. The Table 4.2 summarizes the estimation results.

Variable	**Factor**			
Dependent	*Yr*	*Ir*	*Yc*	*Ic*
Constant	0.217 ***	-0.089 **	0.103 ***	-0.084 **
Yr_{t-1}	0.167 ***	-0.097	-0.222 ***	-0.287 ***
Ir_{t-1}	-0.158 **	-0.339 ***	-0.148 **	0.246 **
Yc_{t-1}	0.186 ***	0.342 ***	0.852 ***	0.410 ***
Ic_{t-1}	-0.039	-0.086	0.016	-0.445 ***
GDP YoY	0.550 **	0.380	1.001 ***	0.648 *
CPI inflation YoY	0.110	-1.286 ***	-0.210	-0.402
FED int. rate	-0.018 ***	0.003	-0.013 ***	-0.0002
Correction term	-0.009 ***	0.004 ***	-0.004 ***	0.003 **

Table 4.2: VECM estimation results

Commented [H14]: Tady davate do cointegration equation I lagged depenend variable? To je ale hodne nestandardni, ne? Kointegrace ma zachytit dlouhodoby vztah a promenna pochopitelne nemuze deviate od sebe sama.

The predictive power of the estimate, measured by R-square, is high in the case of Y (over 90%), but quite small in the case of I (between 20% and 30%). The results clearly show how strong the interconnectedness among all factors is and also proves that Y and I factors are dependent on the external environment.

From the estimation we were able to construct future predictions of Yr, Ir, Yc and Ic at a given state of the world, expressed by macroeconomic conditions. We constructed two sets of predictions. First, when there is no change in the external environment, i.e. the macroeconomic variables are constructed in such a way that there is no quarter-over-quarter change. Second prediction was based on the actual values of macroeconomic indicators (GDP, inflation and interest rates) in Q1 and Q2 2015. The former prediction allows us to see the development of Yr, Ir, Yc and Ic ceteris paribus whereas the latter can be compared to the real development and thus backtest the model. The Yr, Ir, Yc and Ic developments with the predictions are shown on Figures 4.3 and 4.4.

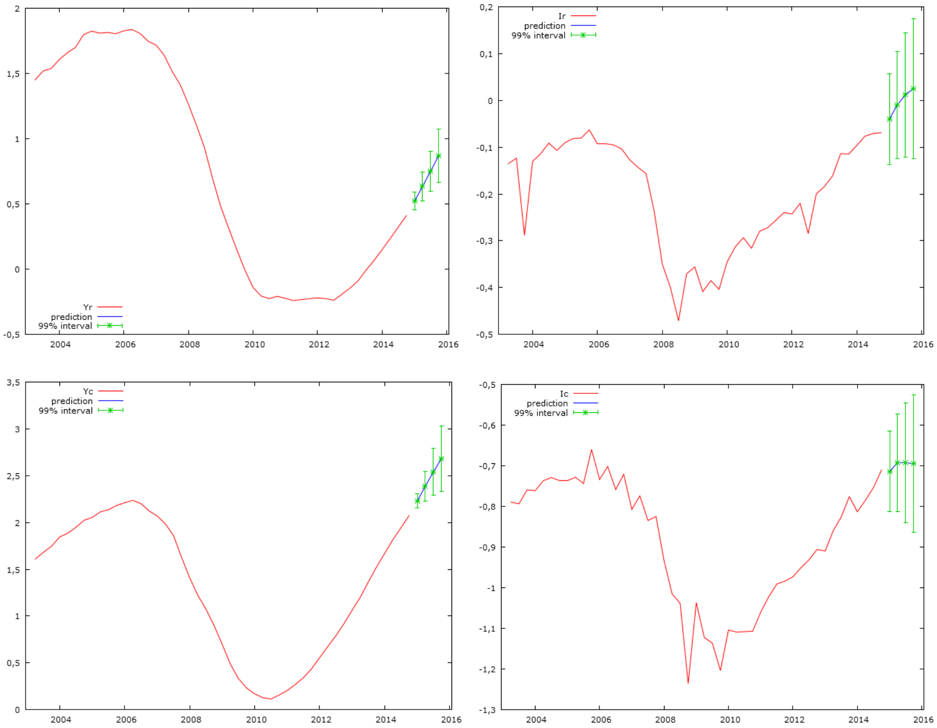

Figure 4.3: The prediction of Yr, Ir, Yc and Ic given unchanged macroeconomic environment

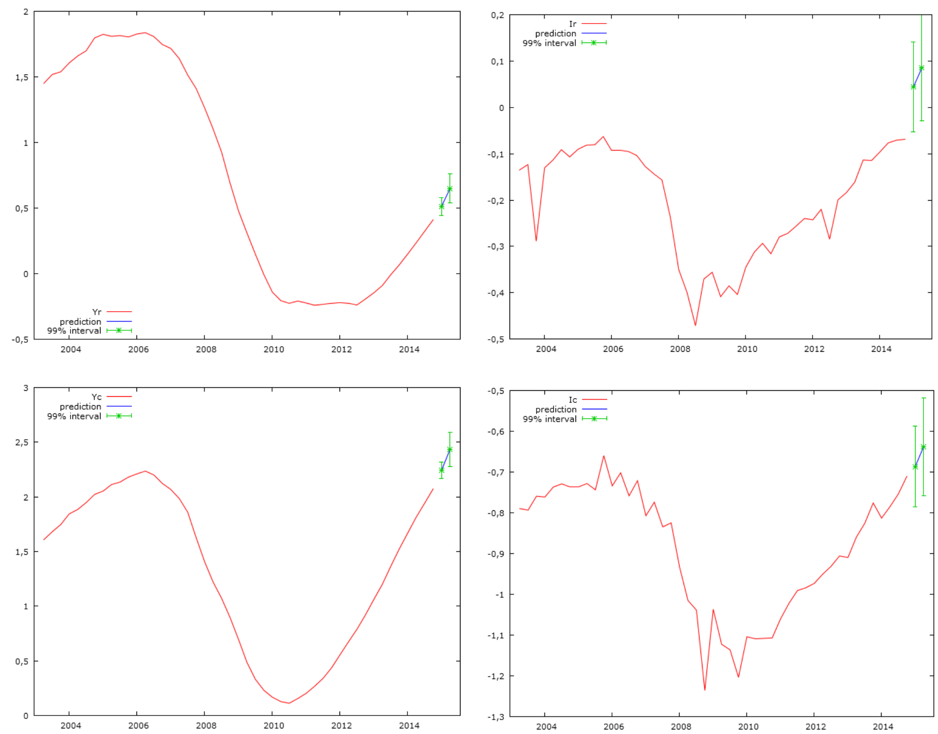

Figure 4.4: The prediction of Yr, Ir, Yc and Ic given real macroeconomic development in Q1 and Q2 2015

In the next step we translated the predicted values of Y and I back to the 30+ delinquency rate (Q) and charge-off rate (G). Figures 4.5 and 4.6 show the datasets of Q residential (Qr), G residential (Gr), and Q commercial (Qc) and G commercial (Gc), respectively, with the predictions for Q1 and Q2 2015 given the real macroeconomic development.

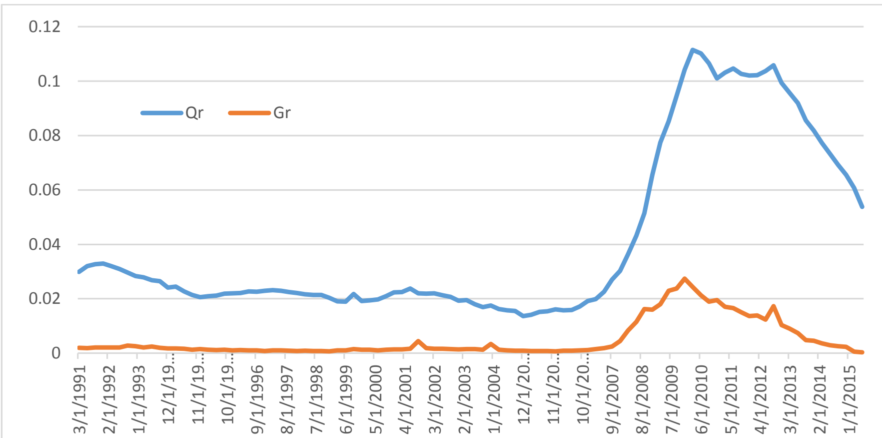

Figure 4.5: Model prediction of the residential 30+ delinquency rate (Qr) and charge-off rate (Gr)

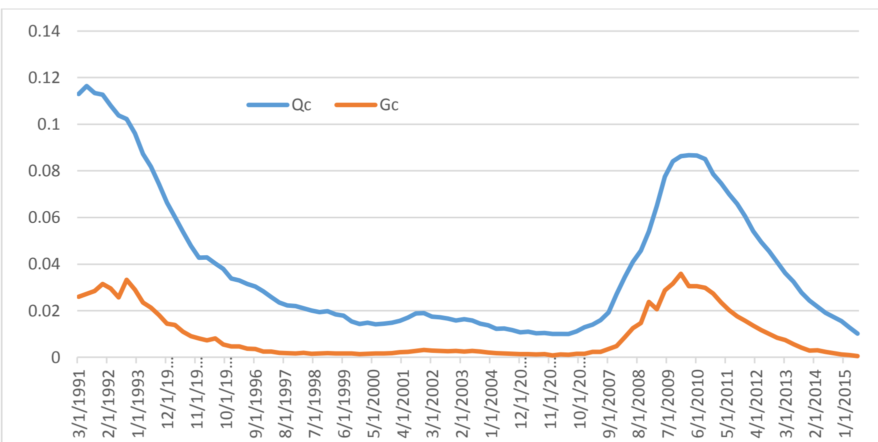

Figure 4.6: Model prediction of the commercial 30+ delinquency rate (Qc) and charge-off rate (Gc)

5. Conclusion

We constructed a multi-period multigenerational multi-portfolio dynamic model of credit losses, which is dependent on the external environment. We estimated the presented model on a large U.S. national portfolio of residential and commercial mortgage loans. The empirical analysis showed that there exists a clear and estimable relationship between the credit risk and the macroeconomic environment. Additionally, we proved that the default rate on the portfolio and the loss given default are not independent, as well as there exists interconnectedness between portfolios. Thus a reasonable model of credit risk has to incorporate the interconnectedness between defaults (represented e.g. by a probability of default) and losses (or, in other words, loss given default) and among risk factors of different portfolios. Finally we demonstrated the possibility of prediction of the credit risk.

The proposed model describes the interconnectedness of the credit risk in the loan book consisting of multiple portfolios with the macroeconomic environment and predicts the credit risk. Additionally, the model, thanks to an inclusion of macroeconomic variables, is capable of estimating potential future development of the portfolio based on different macroeconomic assumptions and therefore can be used as a tool for macroeconomic stress testing. Given all the possible applications, the model can be used as a model of economic capital within a financial institution.

References

Frontczak, R., & Rostek, S. (2015). Modeling loss given default with stochastic collateral. *Economic Modelling*, 44 (2015), pp.162-170.

Frye, J. (2000). Collateral Damage. *Risk*.

Gapko, P., & Šmíd, M. (2012). Dynamic Multi-Factor Credit Risk Model with Fat-Tailed Factors. *Czech Journal of Economics and Finance*, 62(2): 125-140.

Gapko, P., & Šmíd, M. (2012a). Modeling a Distribution of Mortgage Credit Losses. *Ekonomicky casopis*, 1005-1023.

Hamerle, A., Dartsch, A., Jobst, R., & Plank, K. (2011). Integrating macroeconomic risk factors into. *The Journal of Risk Model Validation, vol.5/2*, pp. 3-24.

Hochguertel, S., & Ohlsson, H. (2011). Wealth Mobility and Dynamics Over Entire Individual Working Life Cycles. *ECB Working Paper Series no. 1301*.

Jimenez, G., & Mencia, J. (2009). Modelling the distribution of credit losses with observable and latent factors. *Journal of Empirical Finance*, 235-253.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journla of Finance 29*, Chapter 12.

Pesaran, M. H., Schuermann, T., Treutler, B.-J., & Weiner, S. M. (2003). Macroeconomic Dynamics and Credit Risk: A Global Perspective. *CESifo Working Paper No. 995*.

Pykhtin, M. V. (2003). Unexpected Recovery Risk. *Risk*, 74-78.

Sommar, P. A., & Shahnazarian, H. (2009, September). Interdependencies between Expected Default Frequency and the Macro Economy. *International Journal of Central Banking*, pp. 83-110.

Šmíd, M. (2015). Model of risk and losses of a multigeneration mortgage portfolio. *10th International Scientific Conference Financial management of firms and financial institutions.* Available at http://ssrn.com.

Vasicek, O. A. (1987). *Probability of Loss on Loan Portfolio.* KMV.

Virolainen, K. (2004). Macro Stress Testing with a Macroeconomic Credit Risk Model for Finland. *Bank of Finland Discussion Paper No. 18/2004*.

Witzany, J. (2011). A Two-Factor Model for PD and LGD Correlation. *Bulletin of the Czech Econometric Society*.