Institute of Economic Studies, Faculty of Social Sciences
Charles University in Prague

Bayesian default probability models

Petra Andrlíková

IES Working Paper: 14/2014

Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Bibliographic information:
Andrlíková P. (2014). “Bayesian default probability models” IES Working Paper 14/2014. IES FSV. Charles University.

This paper can be downloaded at: http://ies.fsv.cuni.cz

Bayesian default probability models

Petra Andrlíková[a]

[a]IES, Charles University Prague
E-mail: andrlikova@gmail.com

May 2014

Abstract:
This paper proposes a methodology for default probability estimation for low default portfolios, where the statistical inference may become troublesome. The author suggests using logistic regression models with the Bayesian estimation of parameters. The piecewise logistic regression model and Box-Cox transformation of credit risk score is used to derive the estimates of probability of default, which extends the work by Neagu et al. (2009). The paper shows that the Bayesian models are more accurate in statistical terms, which is evaluated based on Hosmer-Lemeshow goodness of fit test, Hosmer et al. (2013).

Keywords: default probability, bayesian analysis, logistic regression, goodness-of-fit

JEL: C11, C51, C52, G10

1 Introduction

Credit rating shall serve as an evaluation of debtor's creditworthiness and is issued by credit rating agencies (CRAs). The creditworthiness is commonly understood as the ability of a debtor to repay its debt and is therefore closely link to default probability. Nevertheless, CRAs deliberately do not publish their best estimate of default probability leaving the relation between credit rating and the value default probability uncertain. This paper proposes a way how to estimate the conditional probability of default given a certain credit rating in order to reveal correlation between credit rating and default risk based on observed default events.

The main objective of this work is to translate the observed default rates into default probability forecasts for given rating grades in the system. Tasche (2012) refers to this process as the *calibration* process, where the resulting forecasts of default rates in rating grades are called *PD (probability of default)* curve. The default rate forecasts are made using two groups of forecasting techniques. The first group are the variations of logistic regression using the classical or frequentionist type of analysis. The second group consists of Bayesian alternatives of logistic regression models. The forecasting techniques are evaluated based on the goodness-of-fit Hosmer-Lemeshow test, Hosmer *et al.* (2013).

The primary contribution of this paper to current literature is the finding that Bayesian methodology improves the model fit of observed default rates in estimating default probabilities. This paper provides a tool how to estimate default probabilities on rating entities while ensuring its monotonous behaviour. Moreover, the paper proposes a solution to small data sample problem, where any statistical inference becomes troublesome. This is usually due to the problem of non-invertible Hessian matrix, which is solved by introducing Bayesian analysis to derive the values for variance parameters.

This paper is focused on the estimation of point in time, i.e. the most timeliest, of default probabilities for companies rated by S&P in years 2011 and 2012. These years are demonstrated within this paper to provide the most up-to-date information on rating default probabilities.[1] The S&P ratings database is utilized since it represents the largest available set of rated firms in rating history. It is however assumed that similar conclusions would be made with Moody's or other rating agency data sets.

[1]The same analysis was conducted on years 2000 until 2010 and the same conclusion that Bayesian techniques improve model fit was made.

Logit analysis has been the most dominant methodology used in default probability estimation, at least in terms of Journal of Banking and Finance publications, Altman & Saunders (1997). Martin (1977) uses logit analysis to predict bank failures in 1975 and 1976. West (1985) measured the financial condition of financial institutions with logit model and assigned a probability of being a problematic bank. Moreover, Platt & Platt (1991) employed the logit model to test whether industry relative accounting ratios are better in predicting corporate bankruptcy. Lawrence *et al.* (1992) use the logistic regression to predict default probability on mobile home loans and Douglas Smith & Lawrence (1995) utilized logit model to find variables that provide the best prediction of a loan moving into a default state.

This paper follows the work by Neagu *et al.* (2009) that proposes a methodology for translating credit risk scores into probabilities of default using a piecewise logistic regression model and Box-Cox transformation of credit risk score. These two techniques were evaluated as the best considering the expected deviation of the forecasted *PDs* with the observed default rates. Nevertheless, in certain low default portfolio, the estimation process of these two models using the frequetionist analysis may become troublesome due to the over-parametrization problem, which usually takes a form of nonivertible Hessian matrix. In these cases, Gill & King (2004) propose to rethink the model, respecify it, rerun the analysis or get more data. This is however in most cases not feasible and therefore we opt for another option, which is the Bayesian analysis. The problem of a singular Hessian matrix is solved with Bayesian analysis by using algorithms that enable to draw directly from the posterior distribution.

This work concentrates primarily on the statistical problem of the forecast accuracy, which is analogous to Neagu *et al.* (2009). The purpose of this paper is therefore to choose the best value of the goodness-of-fit metric, the Hosmer-Lemeshow test. Jankowitsch *et al.* (2007) show that improving the accuracy of the *PD* estimation leads to significant increases in portfolio returns.

The paper is structured as follows: Section 2 describes the research problem. Section 3 defines the methodology used for *PD* estimation and defines the two groups of forecasting techniques. Classical logistic regression models are confronted with Bayesian logistic regression techniques. Section 4 provides the empirical results and the evaluation of the models. Section 5 concludes.

2 Problem setting

Table 1 provides the observed default rates of corporate companies with S&P rating grades published in the default studies S&P (2013) and S&P (2012). The purpose of this analysis is to estimate the conditional *probability of default* for each rating that would be as close as possible to the empirical default rates from Table 1 for given years and on average does not overestimate or underestimate the level of observed default rates.

Table 1: Observed default rates, 2011, 2012

rating	DR (%)		Count	
	2011	2012	2011	2012
AAA	0.00	0.00	51	24
AA+	0.00	0.00	36	51
AA	0.00	0.00	120	61
AA-	0.00	0.00	207	238
A+	0.00	0.00	357	337
A	0.00	0.00	470	445
A-	0.00	0.00	560	548
BBB+	0.00	0.00	473	523
BBB	0.00	0.00	549	589
BBB-	0.20	0.00	508	525
BB+	0.00	0.00	260	311
BB	0.00	0.00	319	333
BB-	0.00	0.74	403	403
B+	0.39	0.57	509	526
B	1.19	1.39	586	646
B-	3.99	3.34	301	299
CCC/C	15.94	26.62	138	154

Source: S&P (2013), S&P (2012).

Moreover, the *PD* forecasts shall have the property of monotonicity. Monotonicity is required to maintain the economic interpretation of agency ratings, where better rating imply higher creditworthiness and lower default risk. The proposed techniques will therefore aim to find such fitting function that would smoothen the non-monotonous observed default rates and still provide a reasonably good fit to data.

This study proposes several forecast techniques how to estimate the conditional *PDs* while maintaining the above mentioned properties. The forecasting techniques are defined in Section 3 together with the problems one can encounter in the estimation process. The complication in the estimation process may arise with small data samples, which very often holds for many default studies. Therefore a Bayesian analysis is conducted on the given data sample and the techniques are summarized in Section 3.

The combination of small data sample and large number of parameters might lead to a noninvertible *Hessian* matrix, which was also a case in this study. There are several potential sources of singular *Hessian* matrix. First, multicollinearity, which is a statistical phenomenon, when two or more variables in the regression model are highly correlated. In our model, we have only one explaining variable, multicollinearity can be thus excluded. Gill & King (2004) note that receiving a computer-generated *"Hessian not invertible"* message is a common occurrence in applied quantitative research. It can occur even in Monte Carlo simulation experiments while drawing the data from a given statistical model. Unfortunately, there exists no computational trick to make the nonivertible *Hessian* matrix invertible, with the available data set and model.

When we talk about the nonivertible or singular *Hessian* matrix, it is important to mention what it actually implies for our estimates of the model. If the *Hessian* matrix is nonivertible, the variance matrix does not exist. But it does not necessarily mean that estimators that maximize the likelihood function do not contain valuable information. Gill & King (2004) mention that discarding such analysis is not optimal as it would lead to potentially biased procedure.

We have decided to keep the models, even in case the singular *Hessian* matrix is generated, and deal with the singularity problem by introducing the Bayesian estimation procedure to derive the values for variance of parameters. This is in line with proposal by Gill & King (2004). The Bayesian analysis using the logistic regression is summarized in Section 3.

All forecasting techniques are evaluated based on Hosmer-Lemeshow goodness-of-fit test. The results are included in Section 4. The results show that the Bayesian procedure is not only theoretically more correct but it also provides a better fit.

3 Methodology

3.1 Classical econometrics

We are given n observations of the pair (x_i,y_i), i=1, 2, ..., n, where y_i defines the values of independent variable for company i. The independent variable can reach values 1, which denotes a defaulted company i, or 0 which refers to a non-defaulted company i.

In classical or frequentionist econometrics, the general estimation method for model with a dichotomous outcome is *maximum likelihood*. This method maximizes the probability of obtaining the observed data set by maximizing the *likelihood function*. The *maximum likelihood estimators* are the parameters that maximize the *likelihood function*.

Various distributions functions have been proposed for the variables with dichotomous outcome, as discussed in Cox (1989). The logistic distribution is chosen due to the following reasons. First, it is extremely flexible and easy to use. Second, the parameters of the model provide meaningful estimates of effect. Details on the incentives for selecting the logistic distribution are further discussed in Hosmer *et al.* (2013).

Standard logistic regression

The general form of the logistic regression model is defined as

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} \tag{1}$$

where π(x) represents the conditional mean of Y given x, $\pi\left(x\right) = \mathrm{E}\left(Y \mid x\right)$.

We refer to this basic form of the logistic regression as to the simple logistic regression model (SLG).

Piecewise logistic regression

Neagu *et al.* (2009) note that the standard logistic regression tends to overestimate the forecasted PDs at one end and underestimate at the other end. They propose to split the data in two sets and estimate the standard logistic regression separately. We decided to include a breaking point in the logistic regression x_0 that defines the value of the score (or numerical transformation of rating in this case) and helps to better capture the observed default rates that the standard logistic regression model. The piecewise logistic regression model (PWLG) is defined as

$$\pi(x) = \begin{cases} \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} & \text{for } x < x_0 \\ \frac{e^{\beta_0+\beta_1 x_0+\beta_2 x}}{1+e^{\beta_0+\beta_1 x_0+\beta_2 x}} & \text{for } x \geq x_0 \end{cases} \tag{2}$$

Note that x_0 is another parameter that is estimated within this model.

Box Cox logistic regression

The score or rating data observed in real-world applications tends to exhibit a high degree of skewness, as is visible from Table 1. It is therefore recommended to transform the rating or score variable by e.g. the Box-Cox power transformation, see Granger & Newbold (1977) and Neagu *et al.* (2009).

Box-Cox logistic regression model (BCLG) is given as follows.

$$\pi(x) = \frac{e^{\beta_0+\beta_1\left(\frac{x^\lambda-1}{\lambda}\right)}}{1+e^{\beta_0+\beta_1\left(\frac{x^\lambda-1}{\lambda}\right)}} \tag{3}$$

The Box-Cox power transformation is done in other reduce anomalies such as non-additivity, non-normality and heteroscedasticity. A review of the transformation techniques can be found e.g. in Sakia (1992).

3.2 Bayesian econometrics

In the setting of the Bayesian logistic regression model, we follow the approach by Hosmer *et al.* (2013). This approach is further extended by implementing more complex models. The approach by Hosmer *et al.* (2013) defines the standard version of the Bayesian logistic regression only.

The definition of the Bayesian models used remains the same as defined in Section 3.1. The only difference is that we now assume that the parameters of the model are random variables. The individual models are subscribed below.

The Bayesian procedure will be defined on the example of standard logistic regression model. Further variations of the Bayesian logistic regression models are defined in Section 3.2 and 3.2.

Bayesian logistic regression

The Bayesian logistic regression model (BLG) with one predictor is defined as

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} \tag{4}$$

with the *"prior" distribution* of parameters:

$$\beta_0 \sim N(\mu_0, \sigma_0^2) \tag{5}$$

and

$$\beta_1 \sim N(\mu_1, \sigma_1^2). \tag{6}$$

The prior distribution is defined before the analysis of data. We choose the common normal distribution, similarly to Gelman *et al.* (2003). The prior distribution has the following interpretation. The larger the prior variance, σ_0^2 and σ_1^2, the lower precision of the prior distribution is assumed and therefore a lower weight is put on the prior mean values μ_0 and μ_1.

In Bayesian analysis, we are interested in determining the distribution of the parameters given the observed data

$$f(\beta_0, \beta_1 \mid Y). \tag{7}$$

The *likelihood* distribution of the observed data given the parameters is

$$f(Y \mid \beta_0, \beta_1) = \prod_{i=1}^{n} \pi(x_i)^{y_i} \left\{1 - \pi(x_i)^{y_i}\right\}^{1-y_i}. \tag{8}$$

The relationship between the expressions in equations (7) and (8) is defined by Bayes' Theorem. The *"posterior"* distribution of parameters given the data is given by

$$f(\beta_0, \beta_1 \mid Y) = \frac{f(\beta_0, \beta_1) f(Y \mid \beta_0, \beta_1)}{f(Y)}. \tag{9}$$

The denominator of equation (9) is the distribution of the observed data, which can be calculated by integrating the joint density over all parameters, or summing all the probability of data

$$f(Y) = \int f(\beta_0, \beta_1) f(Y \mid \beta_0, \beta_1) d\beta_0, d\beta_1. \tag{10}$$

Formula (9) for the posterior distribution of parameters can be interpreted in the following manner. The posterior distribution is a combination of our "prior" belief about the distribution of parameters and the observed data. It is particularly uneasy to evaluate the expression (9). The increasing computational power made it possible to find the posterior distribution using simulation methods. The simulation method that is used in this paper is the *Markov Chain Monte Carlo (MCMC)* method.

The samples from the *MCMC* simulation depend only on the previous value a create a Markov Chain, Ross (2001). The general form of *MCMC*, the Metropolis Algorithm is used, Metropolis & Ulam (1949) and Metropolis *et al.* (1953). Metropolis Algorithm is described in Hosmer *et al.* (2013).

Bayesian piecewise logistic regression

The model definition for Bayesian piecewise logistic regression (BPWLG) is equivalent to model defined in (2) with the assumption of normal distribution of parameters

$$\beta_0 \sim N(\mu_0, \sigma_0^2), \tag{11}$$

$$\beta_1 \sim N(\mu_1, \sigma_1^2), \tag{12}$$

$$\beta_2 \sim N(\mu_2, \sigma_2^2) \tag{13}$$

and

$$x_0 \sim N(\mu_{x_0}, \sigma_{x_0}^2). \tag{14}$$

Bayesian Box Cox logistic regression

The Bayesian logistic regression model with Box-Cox power transformation, or the Bayesian Box-Cox logistic regression model (BBCLG), follows the definition (3) and the prior distribution of parameters is

$$\beta_0 \sim N(\mu_0, \sigma_0^2), \tag{15}$$

$$\beta_1 \sim N(\mu_1, \sigma_1^2) \tag{16}$$

and

$$\lambda \sim N(\mu_\lambda, \sigma_\lambda^2). \tag{17}$$

4 Data and Empirical results

4.1 Data

Data used in this study are available in S&P (2013) and S&P (2013), which contains global corporate default study with the information of default or non-default events for corporate companies rated by Standard and Poor's credit rating agency in years 2011 and 2012. Data cover US, Europe and emerging market regions and other developed countries. For further information about the particular companies analyzed in this study, refer to S&P (2013) and S&P (2013). Sectors covered within this study include financial institutions, insurance companies and other nonfinancial corporates. The analysis is conducted on the years 2011 and 2012 in order to receive a point-in-time estimate of default probability relevant to the latest observed default rate in credit ratings.

Table 2: Numerical transformation of ratings

rating	num. rating
AA	1
AA-	2
A+	3
A	4
A-	5
BBB+	6
BBB	7
BBB-	8
BB+	9
BB	10
BB-	11
B+	12
B	13
B-	14
CCC/C	15

Table 2 presents the numerical transformation of letter-ratings. It implies equi-distant intervals of resulting default probabilities in log-odds, which might be a restriction of the model but is indeed a property of many internal rating systems.

4.2 Empirical results

Summary of empirical results of all variations of logistic regression analyzed in this study is depicted in Table A.1 in the Appendix. One can note that the resulting default probabilities from SLG and PWLG model are of the same values in all ratings. This is due to the fact that the coefficient β_1 from equation (2) turned out to be nonsignificant. The values of the estimated coefficients is provided in Table A.2 in the Appendix. The results of diagnostic tests used for Bayesian models are available upon request. The convergence test were conducted in order to find the suitable number of simulations. The number of simulations was increased from $100,000$ to $200,000$ when necessary.

4.3 Goodnes-of-fit test

In order to rank the model in terms of their statistical power, Hosmer-Lemeshow test is used, Hosmer *et al.* (2013). Hosmer-Lemeshow test statistics is a measure of goodness-of-fit, in which observations are distributed into ten equal sized groups according to their predicted default probabilities. The the chi-square test statistics is defined as

$$GL_{HL}^2 = \sum_{j=1}^{10} \frac{(O_j - E_j)^2}{E_j\left(1 - \frac{E_j}{n}\right)} \sim \chi_s^2 \tag{18}$$

where
n_j = Number of observations in the j^{th} group,
O_j = Observed number of default cases in the j^{th} group,
E_j = Expected number of default cases in the j^{th} group,
s = Number of degrees of freedom, in our case, $s = 10 - 2 = 8$.

Table 3: Hosmer-Lemeshow test statistics

Model	2011	2012
SLG	0.02123	0.28811
PWLG	0.02124	0.28810
BCLG	0.02257	0.27441
BLG	0.02170	0.30651
BPWLG	0.02054	0.27838
BBCLG	0.01872	0.17753

Source: Author's calculations.

Table 3 summarizes the values of the Hosmer-Lemeshow test statistics. All the estimated models provide a good-fit of the observed default rates in years 2011 and 2012. The values do not only provide information about the goodness-of-fit of the individual models but the test statistics itself can serve as a ranking tool of the statistical power, where the lower the value, the better the fit of the model. We can therefore conclude that in 2011 and 2012 Bayesian Box-Cox logistic regression model is superior to all models analyzed in this paper. In general, Bayesian logistic regression models provide a better fit compared to standard logistic regression models. This is not true in case of the standard form of logistic regression (SLG vs. BLG), where the classical type of logistic regression itself already does a good enough job and the introduction of the Bayesian technique is not necessary. Nevertheless, in case of the piecewise model and Box-Cox transformation, the Bayesian aproach outperfoms the classical logistic regression models rather markedly. In case of the piecewise logistic regression model, the Bayesian approach decreases the value of Hosmer-Lemeshow test by 3% in both analyzed years and in Box-Cox transformed model, the decrease of the test statistics reaches 35% in 2012 and 17% in 2011.

5 Conclusion

In this paper, we have described an approach how to estimate a probability of default given any rating scale available. It enables us to reveal the relation between the credit ratings issued by the CRAs and the value of default probability. We use models on the logistic regression basis, which offer great flexibility in the calibration of the model since this method can be used on the scale of ratings as well as on buckets of score. Moreover, the method satisfies the condition of monotonicity of estimated default probabilities, which is required by commercial banking institutions that follow the current Basel Accords banking regulation.

The standard logistic regression model is confronted with piecewise and Box-Cox logistic regression. The value added of the Bayesian analysis procedures is tested based on the goodness of fit test. The analysis is conducted on 2011 and 2012 default history of companies rated by S&P credit rating agency. It covers the industries that are included in the database published in S&P (2013) and S&P (2012).

The Bayesian method with the combination of the Box-Cox transformation improved the model fit compared to other techniques and was proven to be the superior method to all in both years 2011 and 2012. The impact of the

Bayesian procedure on the goodness-of-fit is quite evident. If we take a look at the winning Bayesian Box-Cox logistic regression model and compare it with the classical Box-Cox logistic regression model, the goodness-of-fit statistics decreases by 17% in 2011 and 35% in 2012.

This paper is mainly relevant for large institutions that follow Basel regulatory Accords and deal with the problem of small data samples. The methodology proposed in this study may be used on existing external rating scale as well as on any internal rating systems implemented by the individual institutions.

Bibliography

ALTMAN, E. I. & A. SAUNDERS (1997): "Credit risk measurement: Developments over the last 20 years." *Journal of banking & finance* **21(11)**: pp. 1721–1742.

COX, D. D. R. (1989): *The analysis of binary data*, volume 32. CRC Press.

DOUGLAS SMITH, L. & E. C. LAWRENCE (1995): "Forecasting losses on a liquidating long-term loan portfolio." *Journal of Banking & Finance* **19(6)**: pp. 959–985.

GELMAN, A., J. B. CARLIN, H. S. STERN, & D. B. RUBIN (2003): *Bayesian data analysis*. CRC press.

GILL, J. & G. KING (2004): "What to do when your hessian is not invertible alternatives to model respecification in nonlinear estimation." *Sociological methods & research* **33(1)**: pp. 54–87.

GRANGER, C. W. J. & P. NEWBOLD (1977): "Forecasting economic time series." *Economic Theory and Mathematical Economics* .

HOSMER, D., S. LEMESHOW, & R. STURDIVANT (2013): *Applied Logistic Regression*. Wiley Series in Probability and Statistics. Wiley.

JANKOWITSCH, R., S. PICHLER, & W. S. SCHWAIGER (2007): "Modelling the economic value of credit rating systems." *Journal of Banking & Finance* **31(1)**: pp. 181–198.

LAWRENCE, E. C., L. D. SMITH, & M. RHOADES (1992): "An analysis of default risk in mobile home credit." *Journal of Banking & Finance* **16(2)**: pp. 299–312.

MARTIN, D. (1977): "Early warning of bank failure: A logit regression approach." *Journal of banking & finance* **1(3)**: pp. 249–276.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, & E. Teller (1953): "Equation of state calculations by fast computing machines." *The journal of chemical physics* **21**: p. 1087.

Metropolis, N. & S. Ulam (1949): "The monte carlo method." *Journal of the American statistical association* **44(247)**: pp. 335–341.

Neagu, R., S. Keenan, & K. Chalermkraivuth (2009): "Internal credit rating systems: Methodology and economic value." *The Journal of Risk Model Validation* **3(2)**: pp. 11–34.

Platt, H. D. & M. B. Platt (1991): "A note on the use of industry-relative ratios in bankruptcy prediction." *Journal of Banking & Finance* **15(6)**: pp. 1183–1194.

Ross, S. M. (2001): "Stochastic processes. 1996."

Sakia, R. (1992): "The box-cox transformation technique: a review." *The statistician* pp. 169–178.

S&P (2012): "Default, transition, and recovery: 2011 annual global corporate default study and rating transitions." Report.

S&P (2013): "Default, transition, and recovery: 2012 annual global corporate default study and rating transitions." Report.

Tasche, D. (2012): "The art of pd curve calibration." *Available at SSRN 2189870* .

West, R. C. (1985): "A factor-analytic approach to bank condition." *Journal of Banking & Finance* **9(2)**: pp. 253–266.

Appendix

Table A.1 Predicted PDs in ratings (%), years 2011 and 2012

rating	DR	SLG	PWLG	BC	BLG	BPWLG	BBC	Year
AAA	0.00	4.70E-08	4.70E-08	3.36E-07	0.00E+00	0.00E+00	0.00E+00	2011
AA+	0.00	1.61E-07	1.61E-07	6.00E-07	0.00E+00	0.00E+00	0.00E+00	
AA	0.00	5.49E-07	5.50E-07	1.22E-06	0.00E+00	0.00E+00	0.00E+00	
AA-	0.00	1.88E-06	1.88E-06	2.74E-06	4.66E-06	1.21E-05	0.00E+00	
A+	0.00	6.42E-06	6.42E-06	6.66E-06	5.60E-06	1.12E-05	0.00E+00	
A	0.00	2.20E-05	2.20E-05	1.74E-05	2.98E-05	4.04E-05	8.51E-06	
A-	0.00	7.51E-05	7.51E-05	4.85E-05	1.04E-04	1.07E-04	3.93E-05	
BBB+	0.00	2.57E-04	2.57E-04	1.43E-04	3.54E-04	3.94E-04	1.18E-04	
BBB	0.00	8.77E-04	8.78E-04	4.46E-04	1.07E-03	1.26E-03	4.66E-04	
BBB-	0.20	3.00E-03	3.00E-03	1.46E-03	3.38E-03	3.79E-03	1.76E-03	
BB+	0.00	0.01	0.01	0.00	0.01	0.01	0.01	
BB	0.00	0.04	0.04	0.02	0.04	0.04	0.02	
BB-	0.00	0.12	0.12	0.07	0.12	0.12	0.09	
B+	0.39	0.41	0.41	0.26	0.41	0.41	0.32	
B	1.19	1.38	1.38	1.03	1.38	1.36	1.20	
B-	3.99	4.57	4.57	4.18	4.56	4.53	4.45	
CCC/C	15.94	14.07	14.07	15.88	14.07	14.21	15.28	
AAA	0.00	1.13E-09	1.13E-09	3.04E-05	0.00E+00	0.00E+00	4.17E-05	2012
AA+	0.00	4.62E-09	4.62E-09	3.44E-05	0.00E+00	0.00E+00	1.96E-05	
AA	0.00	7.72E-08	7.72E-08	5.68E-05	0.00E+00	0.00E+00	4.92E-05	
AA-	0.00	3.15E-07	3.15E-07	8.28E-05	4.20E-06	0.00E+00	9.24E-05	
A+	0.00	1.29E-06	1.29E-06	1.32E-04	0.00E+00	1.48E-06	1.66E-04	
A	0.00	5.27E-06	5.27E-06	2.28E-04	7.87E-06	1.24E-05	2.79E-04	
A-	0.00	2.15E-05	2.15E-05	4.30E-04	3.10E-05	3.01E-05	6.13E-04	
BBB+	0.00	8.79E-05	8.79E-05	8.86E-04	1.21E-04	1.34E-04	1.26E-03	
BBB	0.00	3.59E-04	3.59E-04	1.99E-03	3.94E-04	5.03E-04	2.89E-03	
BBB-	0.00	1.47E-03	1.47E-03	4.86E-03	1.60E-03	1.72E-03	7.18E-03	
BB+	0.00	0.01	0.01	0.01	0.01	0.01	0.02	
BB	0.00	0.02	0.02	0.04	0.02	0.03	0.05	
BB-	0.74	0.10	0.10	0.12	0.10	0.10	0.16	
B+	0.57	0.41	0.41	0.41	0.40	0.41	0.50	
B	1.39	1.64	1.64	1.55	1.61	1.63	1.74	
B-	3.34	6.39	6.39	6.10	6.35	6.35	6.27	
CCC/C	26.62	21.82	21.82	22.64	22.11	21.88	21.39	

Source: Author's estimations

Table A.2 Estimated coefficients of the classical logistic regression model, years 2011 and 2012

	2011		2012	
	Estimate	Std Error	Estimate	Std Error
SLG Model				
β_0	-22.707	6.391	-26.612	5.990
β_1	1.229	0.025	1.408	0.020
PWLG Model				
β_0	-22.707	2.528	-26.613	2.447
β_1	1.000	.	1.000	.
β_2	1.229	0.157	1.408	0.142
x_0	0.000	.	0.000	.
BCLG Model				
β_0	-13.966	.	-15.005	1.366
β_1	5.739	.	0.082	.
λ	0.485	.	2.018	0.046

Source: Author's estimations

Table A.3 Posterior distribution of Bayesian estimates, years 2011 and 2012

	N	Mean	StdDev	P25	P50	P75	Year
BLG Model							
β_0	200000	-22.7488	1.5081	-23.7674	-22.7306	-21.7179	2011
β_1	200000	1.231	0.0939	1.1671	1.2302	1.2946	
β_0	200000	-9.1503	0.4693	-9.4605	-9.1492	-8.848	2012
β_1	200000	0.3561	0.0311	0.336	0.3561	0.3767	
BPWLG Model							
β_0	200000	-22.5815	2.3078	-24.137	-22.6418	-21.1915	2011
β_1	200000	-0.2418	2.1266	-1.5421	-0.1554	1.0924	
β_2	200000	1.2469	0.1438	1.1454	1.242	1.3404	
x_0	200000	0.4063	2.7747	-1.159	0.1139	1.5422	
β_0	200000	-26.7169	1.8585	-27.9204	-26.7442	-25.5142	2012
β_1	200000	1.2334	2.2023	-0.1781	0.765	2.4742	
β_2	200000	1.417	0.1246	1.3337	1.4147	1.4968	
x_0	200000	-0.3663	1.9525	-1.2683	-0.2436	0.643	
BBC Model							
β_0	100000	-22.9168	0.8774	-23.4069	-22.8163	-22.2625	2011
β_1	100000	1.2653	0.1562	1.1521	1.2693	1.3702	
β_0	100000	-15.1339	0.4664	-15.5169	-15.0736	-14.766	2012
β_1	100000	0.1256	0.0436	0.0934	0.1199	0.1531	

Source: Author's estimations

IES Working Paper Series

2014

1. Diana Zigraiova, Petr Jakubik: *Systemic Event Prediction by Early Warning System*
2. Pavla Blahova, Karel Janda, Ladislav Kristoufek: *The Perspectives for Genetically Modified Cellulosic Ethanol in the Czech Republic*
3. Oxana Babecká Kucharčuková: *The impact of exchange rate volatility on trade: Evidence for the Czech Republic*
4. Oliver Polyák: *The Impact of Euro Adoption on Export Performance: Comparison of the Czech Republic and Slovakia*
5. Lucie Davidová, Vladimír Benáček: *Determinants of Austrian International Trade: Analysis Based on the Gravity Model*
6. Julie Chytilová, Václav Korbel: *Individual and Group Cheating Behavior: A Field Experiment with Adolescents*
7. Petr Janský: *Consumer Demand System Estimation and Value Added Tax Reforms in the Czech Republic*
8. Pavel Hait, Petr Janský: *Inflation Differentials among Czech Households*
9. Hana Džmuráňová, Petr Teplý: *Risk management of savings accounts*
10. Tomáš Fiala, Tomáš Havránek: *Ailing Mothers, Healthy Daughters? Contagion in the Central European Banking Sector*
11. Petra Andrlíková: *Is Barrier version of Merton model more realistic? Evidence from Europe.*
12. Martin Pospíšil, Jiří Schwarz: *Bankruptcy, Investment, and Financial Constraints: Evidence from a Post-Transition Economy*
13. Matt Hunt, Goran Serdarević, Tom Ovington, Clive Kenny: *Evidence for a ladder of investment in Central and Eastern European countries*
14. Petra Andrlíková: *Bayesian default probability models*

All papers can be downloaded at: http://ies.fsv.cuni.cz

Univerzita Karlova v Praze, Fakulta sociálních věd
Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26
E-mail : ies@fsv.cuni.cz

http://ies.fsv.cuni.cz