Opponent's Report on Dissertation Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague
Opletalova 26, 110 00 Praha 1, Czech Republic
Phone: +420 222 112 330, Fax: +420 222 112 304

Author:	**Petr Gapko**
Advisor:	RNDr. Martin Šmíd, Ph. D.
Title of the Thesis:	Three Essays on Credit Risk Quantification
Type of Defense:	**DEFENSE**
Date of Pre-Defense:	April 16, 2014
Opponent:	**doc. RDR. Jiří Witzany, Ph.D.**

Address the following questions in your report, please:

a) Can you recognize an original contribution of the author?
b) Is the thesis based on relevant references?
c) Is the thesis defendable at your home institution?
d) Do the results of the thesis allow their publication in a respected economic journal?
e) Are there any additional major comments on what should be improved?
f) Were your comments raised at the pre-defense, addressed in the dissertation submitted to the regular defense? (The pre-defense report is enclosed below)
g) What is your overall assessment of the thesis? (a) I recommend the thesis to be defended without major changes; (b) The thesis is not defendable.

(Note: The report should be at least 2 pages long.)

Content of the Report:
The Dissertation Thesis consists of three essays on actual topics in the area of credit risk management and modeling. Each of the essays brings an original contribution and is based on relevant references. All the papers have been published, or allow publication in a respected economic journal.

The first essay proposes a novel model of the distribution of credit losses on a mortgage portfolio. The main difference compared to the Vasicek's or other similar models lies in using of generalized hyperbolic distribution for the factors of the process underlying the rates of default. Parameters of the model are estimated on a US mortgage loans delinquency dataset. The model is used to forecast the loss rates and their quantiles one year ahead and to compare the results with the Basel IRB formula calculation. I see the two following important methodological issues:

a. The assets of i-th borrower are assumed to follow the equation $\log A_{i,t} = \log A_{i,t-1} + Y_t + U_{i,t}$ where Y_t is the systematic factor and $U_{i,t}$ is an idiosyncratic factor. Therefore, given a history of the systematic factors $Y_{\tau\ \tau\le t}$ we need to deal with the cumulative idiosyncratic vector $Z_{i,t} = \sum_{\tau\le t} U_{i,\tau}$. However, the key assumption of the

model is that all $Z_{i,t}$ are identically distributed which appears to be in a contradiction with the cumulative nature of $Z_{i,t}$. In fact, one would rather assume that the factors $U_{i,t}$ are i.d., but not $Z_{i,t}$. I would like to ask the author to explain and defend the proposed approach very well since it is a key component of the model used in the first two essays.

b. The model is applied to delinquency data as if the delinquency rate represented the proportion of defaulted borrowers on a large fixed set of borrowers without any entrances and exits. However, this is not the case. First, defaulted cases are somehow resolved by selling the property and/or by write-offs. Thus, the number of defaulted cases depends on the speed of the workout process which in turn might depend on legislative changes etc. Second, there are new mortgages that enter the dataset while many other mortgages exit due to repayment. All those entrances and exits might mutually "cancel-out", but in any case the observed delinquency rates present only a very rough approximation of the theoretical loss rate. In fact, the author tries to capture this complex dynamics in the third paper. This discrepancy should be in any case commented and explained already in the first essay.

Both aforementioned issues are explained in the final version of the thesis.

The second essay extends the model from the first essay by adding a dynamical LGD model that is possibly interconnected with the default rate model. The theoretical model is estimated on the same dataset, and the estimated model is used to calculate the default rate and LGD quantiles. However, again I am not sure that the used LGD proxy based on the given dataset corresponds well to the theoretical concept. LGD is defined as the expected loss in case of default driven mainly by the collateral value in case of the mortgage loans. But the author uses the proportion of started foreclosures at time *t* as the LGD proxy. I would like the author to explain carefully and defend the claim that those two quantities should be at least approximately equal. A discussion of this problem was added to the last version of the thesis.

The third essay is technically the most advanced. It takes into account entrances and exits through repayments. It proposes a relatively complex model in order to estimate expected and expected default rates and LGDs. The model incorporates default rate and loss rate (LGD) systematic factors that are allowed to depend on macroeconomic factors. The identified macroeconomic factors are the GDP growth and the unemployment rate. The estimated model is used in order to predict the 99.9% loss quantile that is compared with an IRB estimate. The model solves the criticism in terms of the dataset application as pointed with respect to the first two essays. However, the cost is a high complexity of the model and of its estimation. I would recommend paying more attention to models that would be applicable to standard default and loss databases where account-level loss and recovery observations are available. Even though such datasets are not usually public they are normally available to large banks and regulators.

The IRB estimation of the unexpected loss defined as the product of the DR conditional on the 99.9% systematic factor quantile and of the LGD conditional on the 99.9% systematic factor quantile is too conservative making the two compared results artificially too much different. IRB does not require banks to use the 99.9% quantile LGD but a "downturn" LGD which is often interpreted as 90-95% quantile LGD but definitely not 99.9%. I would like the author to explain better his comparison.

The results were recalculated using the 95% loss quantile instead of 99.9% in the version of the thesis presented for defense.

Overall, in spite of the critical remarks, the three essays of the thesis bring original results in the area of credit portfolio loss modeling. The thesis would be defendable at my home university or at similar universities abroad. The remarks and questions stated above can be discussed during the defense or may serve as possible ideas for further research. **Therefore, I recommend the thesis for defense without major changes.**

Date:	
Opponent's Signature:	
Opponent's Affiliation:	doc. RDR. Jiří Witzany, Ph.D. VŠE