Referee Report on Martin Vojtek's CERGE-EI Dissertation
"Essays on Interest Rates and Credit Risk"
Martin Čihák

I have read this draft dissertation with interest. Overall, I think that it has a potential to be a strong thesis. *In my view, this draft warrants a dissertation defense (subject to the comments and suggestions below), and eventual award of a doctorate title.*

General

On the technical level, I found this a well executed thesis. The author shows a good command of the modern banking and finance literature. The theoretical models and the empirical estimations testify to a high degree of competence, and provide valuable insights. At the same time, I think that the thesis can still be strengthened, mostly in terms of motivation and interpretation of the results, but also in terms of the data being used and some of the methodologies employed.

As a general point, the author could devote more space to carefully motivating his work. Starting with the first paragraph of the Preface as an example, what is so special about the situation in countries that joined the EU in May 2004? Why is this interesting to a broader audience?

A related general comment is that the author aims to say something about the "new member states" that joined the EU in 2004, but a large portion of the data being used is from the preceding period, and therefore relatively outdated. Extending the data to more recent periods, in particular to the recent episode of financial turbulence (since the summer of 2007), would allow the author to perform a much richer analysis.

Another comment that appears in one form or another in each of the chapters is the measurement of "goodness of fit": it is not always clear how the author decides that certain estimates are good/consistent/optimal. This has implications for some of the conclusions, as discussed in more detail below.

Chapter 1

- The author could clarify the contribution of this paper to the literature, acknowledging that some research on this topic (the term structure of interest rates in emerging markets) has already been done. In that context, the author may need to revisit his strong statement that "no literature about the calibration of the models of term structure of interest rates to transition markets is available" (page 13, and in similar words, also page 10). Indeed, numerous papers exist on modeling term

structure of interest rates in these countries. To provide just a few examples for the Czech Republic, see the articles in *Czech Journal of Economics and Finance* on this subject by Michal Slavik (2001), which builds on previous research by Viktor Kotlan (1999) and I. Matalik (1994). Similarly, there are working papers and articles analyzing yield curves in the other Visagrad countries. The author should acknowledge this earlier work, identifying clearly what is the value added of the calculations shown in this paper compared to the previous literature.

- A methodological contribution of the paper is that instead of calibrating the model to match the price of some frequently traded derivatives, such as swaptions or caps and floors, it uses an alternative estimation that does not require trading in such derivatives (page 10). To establish the reliability of this alternative approach, it would be useful to show its results for markets for which the estimate with traded derivatives is possible (e.g., those in advanced European countries or in the United States). For advanced markets, does the alternative estimation provide the same or very similar results to the "standard" approach?

- How is the goodness of fit analyzed in the paper? It seems that this is largely based on examining charts. For example, for the Czech Republic, the text refers to a chart of the estimated instantaneous conditional volatility (page 25 and Fig 1.2), and states that the image is "fairly consistent with the data." Based on this, the author concludes that "it is clear that the calibration procedure is able to reveal the true market development." This appears rather arbitrary.

- In the conclusion, the author makes a link between the goodness of fit of the model and the degree of development of the market, suggesting that the problems that the estimation has encountered in Hungary and Slovakia are due to the lower level of development of the market in those two countries compared to Poland and Czech Republic. However, this link is not analyzed explicitly. Moreover, the author suggests other explanations for these problems (for Slovakia, for example, he suggests that the problem may be in the sharp change in the interest rate level), possibly undercutting this "developmental" explanation. It would be useful to at least demonstrate (e.g., in a table or chart with cross-country comparisons of a few developmental indicators) that the markets in Hungary and Slovakia clearly lag behind those in Poland and Czech Republic in terms of basic indicators of development.

- The analysis would be more persuasive if it were based on more recent data, after the Visegrad countries became EU members. It is not clear why the data series stop on November 28, 2003. Not only this choice seems arbitrary, it is also almost 7 years old now. It means that the study covers only the period prior to these countries joining the European Union. Why not extending the time series more to the present, especially considering that the paper deals with high-frequency financial market data that are available with short lags?

- Adding more recent data could also help to address another concern, which is the shortness of the time span analyzed. For example, for Hungary, the time span is only about a year. It is not clear how robust are the results derived from such a short period.

- It would be interesting to show out-of-sample predictions from the model, for the later periods (especially the period of the recent financial turbulence could be extremely interesting for showing the robustness of the model).

- The paper is built on the assumption that default does not exist (page 14). This seems a rather strong assumption, not only given the experience from the current global crisis, but also given the emphasis that the author puts on default risk in Chapter 3.

- The paper could contain a discussion of the reliability of the various IBOR rates. These are usually measured as "the best rates among best banks" and based on a survey. During the recent global financial crisis, major questions have been raised about reliability of LIBOR, even for very advanced, liquid markets. How representative is the WIBOR and similar rates, in markets that are much less developed and much less liquid?

- The author suggests that the analysis has "life-important" implications for market regulators. However, it is unclear what exactly these implications are. It would be useful to discuss more specifically what the findings imply for regulatory/supervisory policy in Slovakia and Hungary on one hand compared to the Czech Republic and Poland.

Chapter 2

- The paper, while containing some useful analysis, would benefit from a clearer motivation. What is the purpose of analyzing consistency between estimates of euro area entry dates derived from market prices and those from market participant surveys? If they were found inconsistent, would it mean that market participants' assessments are irrational? What would be the policy or other implications of such an inconsistency?

- Instead of checking internal consistency of market expectations, it would perhaps be more interesting to analyze the substantive factors (fiscal performance, structural reforms, monetary policy, etc.) influencing the market perceptions of euro area entry. This analysis seems to be left out from the paper.

- The paper would benefit from a clearer analytical framework. The paper focuses on examining the consistency between market's views on euro area accession and surveys of market analysts, and the author concludes that these two are consistent.

However, it is unclear what the benchmark is for deciding whether they are consistent or not. For example, the survey suggested that Poland would enter the euro area by 2015 (presumably with a 100 percent probability or close to that), but the author finds the probability to be only 83 percent. That seems to be an important difference, and possibly an "inconsistency" between the two methods. (Certainly there is an inconsistency between the survey, which saw euro area entry prior to 2015 very likely even in Poland and Hungary, and the quoted paper by Cincibuch and Hornikova, who see it happening only after 2015. Why are these findings assessed as "consistent," rather than "inconsistent"?).

- Against this background, I am not sure how one can arrive at the conclusion that "the method used in the paper fitted market surveys very well" (page 68).

- A potentially important drawback of the paper is that most of the analysis relies on data that are more than 3 years old (the latest yield curve observation being from April 2006). Things have changed substantially due to the impact of the global financial crisis (e.g., Hungary requested an IMF/EU financial package). The author mentions the change in euro area accession strategies in 2009, but the core of the analysis is based on relatively old data.

- The statement that "the Czech Republic is in the best position to fulfill the Maastricht criteria" (page 68) may need to be toned down in view of recent developments (recent estimates from the Czech Ministry of Finance projected the full-year fiscal gap at 4.5 percent of GDP, well over the relevant Maastricht criterion).

- It may be problematic that the model in this chapter abstracts from credit risk (page 56). The importance of credit risk has been highlighted during the recent global financial turbulence (and it is also problematic given the emphasis on credit risk in chapter 3 of this thesis).

Chapter 3

- I am concerned about the representativeness of the data used in the paper. The data are based on a single bank operating in a rather specific market ("small and medium-sized retail clientele in the area of real property purchase and reconstruction") in a small country (the Czech Republic) in the period before EU accession. How transferable are these results to banking markets in the new EU member countries? (Possibly not much, because otherwise banks could copy these models and use them, rather then each trying to develop their own, as described in footnote 5 on pp. 81-82).

- The paper may need to be more specific in explaining in what sense the developed specification is "optimal" (page 77, abstract, first line; and also page 95, bottom).

- Relatedly, the paper needs to include an explicit discussion of Type I and Type II errors in the estimates. Presumably the estimates are trying to minimize some function of these types of errors. But what is the relative weight allocated to these two types of errors in the estimates, and why? The paper seems to bypass this issue by focusing on maximization of c, the space under the ROC curve. However, a bank's risk manager (or banking supervisor) may put very different relative weights on Type I and Type II errors, for instance because he or she is much more concerned about missing defaults than about misclassifying good loans. In other words, it may not be sufficient to look only at the parameter c (which compresses all the information contained in the ROC curve into one number), and one may need to look at the full curve. Framing this discussion in terms of Type I and Type II errors may be useful.

- Has the author examined the introduction of macroeconomic variables? In economies that have gone through major developments, these may well be significant in determining counterparties' ability to repay. The ongoing global financial crisis has illustrated rather dramatically the impact of macroeconomic developments on individual defaults.

- Given that the estimates relate to a bank operating in an emerging economy undergoing rapid changes, it would seem important to test for structural breaks in the series.

- Page 83: It will be important to provide a clear economic motivation for including the various variables, to avoid the perception of data mining.

- It is not clear to what extent the model takes into account collateral. A riskier but well collateralized loan may be more profitable for a bank that a somewhat less risky loan without a collateral. It seems particularly relevant given that the bank in question focuses on loans to "retail clientele in the area of real property purchase and reconstruction" (page 82). (This is relevant also for the comparison between retail lending and corporate credit, mentioned on page 78.) The estimates include a variable "own resources", but it is not clear to what extent this variable includes the quality of the underlying collateral.

- Relatedly, what is important for a bank is not only the probability of default (PD) on a loan (which is what this paper focuses on), but also the loss given default (LDG) of a loan. For example, a low-PD loan may be more of a concern if it is associated with a very high LGD. The estimates in this paper seem to assume the same LGD for all the loans, which may not be realistic.

- To what extent do the estimates take into account the bank's interest rate setting policy? Banks tend to offer higher interest rates on loans to clients that are perceived as more risky. This in turn makes such loans more difficult to repay, increasing the

probability of default, other things being equal. It might be useful to control for the lending rate setting policy, to distinguish the underlying relationship between rating and riskiness from the impact of the differentiated lending rates.
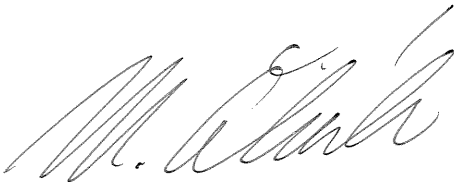
- The text emphasizes the distinction between credit risk measurement in retail and non-retail lending (page 79). Indeed, loan size is smaller, and financial information is more scarce in retail lending, but otherwise, is the methodology so fundamentally different? Also in non-retail lending, one would use parametric and non-parametric techniques along similar lines as discussed here.

- The abstract claims that the model is based on "data from a new EU economy" (page 77, abstract), but the data being used are only up to 2004, i.e. before the Czech Republic even entered the EU (and also before the really rapid growth of retail credit started).

- The motivation of the paper is somewhat out of synch with the contents. The analysis is based on data from the period of 1999-2004 (section 3.2), but the motivation of the paper (section 3.1.2, "objective") is based on information from the 2006-7 period. It would better to use the same data throughout the paper. Also on substance, credit developments in these two periods were quite different, with retail credit growing much more slowly in the early period.

- Page 81, top: The comparison with the "older" EU countries could be formulated more cautiously, given that household sectors in at least some of the older EU countries clearly took on too much debt. Also, it would be useful to quote the source of this information, and provide an "as of" date (as regards the average for the older 15 member states, it is important to use a weighted average, or exclude outliers, such as Luxembourg, which tend to skew the simple average upwards artificially).

- Why should socio-demographic variables be (ex-ante) more important in new EU member countries than elsewhere? (Page 81, first full para.)

- The rate of default on the loans in the sample is surprisingly high, leading to questions about how the dataset was created. About 50 percent of the clients defaulted on the loans, in a period when nonperforming loans in aggregate were in the 5-10 percent range and declining steadily. Either this bank was doing an exceptionally poor job selecting clients, or the sample is preselected in a way that increases the default rate.

- If the sample was pre-selected, it may be more appropriate to use a conditional logit regression technique, which permits outcome probabilities to depend on choice-specific characteristics. This methodology is more appropriate than a standard logit model if the sample is constructed using a matching method. Conditional logit models

are widely used in biostatistics and epidemiology when experiments are conducted using matched case-control studies with $1:k_n$ matching, where k denotes the number of matches and n is the total number of groups. One example from the banking literature is Chen, M. A. (2004) Executive option repricing, incentives, and retention. Journal of Finance, Vol. 59, pp. 1167-1199. Another, more recent, example, is Cihak, M., A. Maechler, K. Schaeck, and S. Stolz, 2009, "Who Disciplines Bank Managers?" IMF Working Paper (forthcoming; available upon request).

- There seems to be some inconsistency in how the comparison of Model 3 (i.e. the one without the "Own Resources" variable) with the other models is assessed. In the conclusion, it is described as "marginally worse" (page 96), while in the body of the text it is emphasized that Model 3 actually has a higher value of the c coefficient.

Editorial

- The reference to IMF Working Paper No. 05/151 is incorrect, both in the text and in the list of references. The authors are Hilbers, Ötker, Pazarbasioglu, and Johnsen (pages 80 and 98).

- In Chapter 3, the notation is somewhat confusing, as c is used to denote the thresholds in the CART analysis as well as the "goodness of fit" statistics (the space under the ROC curve).

- The whole text would benefit from re-reading by a native speaker.

- There are numerous small editorial errors that do not make the text incomprehensible, but are nonetheless distracting. Just to give a few examples, on page 9: "such a markets" should be replaced by "such markets"; on page 7 the last sentence in the last full paragraph seems incomplete'; on page 21, the last sentence in the top paragraph has several typos; on page 25, the last line has a typo; on page 66, in the last paragraph, "Slovakia should be" needs to be replaced by "Slovakia was likely to be".

Martin Čihák
Washington, DC, June 2009