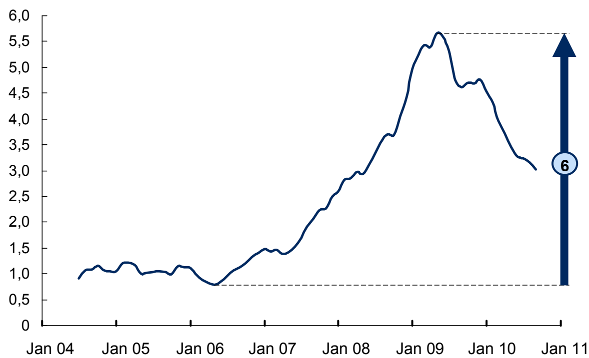
# Discussion to the paper by Martin and František Řezáč

Richard PODPIERA – ČSOB (Ceskoslovenska obchodni banka, a.s.)
(richard.podpiera@seznam.cz)*

This paper addresses an important area, because credit underwriting decisions – and therefore credit scoring models in the case of retail clients – are at the heart of financial intermediation. Credit scoring models are widely used by financial institutions to discriminate, ex ante, between "good" and "bad" retail clients, based on an assessment of their creditworthiness. As such, they help banks and other financial intermediaries to fulfill their mission in the economy – not only to grant credit, but also to collect the funds back so that depositor resources are safeguarded.

Periodical testing of credit scoring model quality is indeed in order. Credit scoring models should be assessed frequently, as the retail market tends to be rather dynamic in emerging markets, in both consumer and mortgage lending. The recent mortgage crisis in the US has shown that the same may be true for developed markets as well. For illustration, the mortgage default rate in the US increased almost six-fold between early 2006 and 2009 (*Figure 1*). Besides such external shocks, changes in risk appetite, product ranges or customer segment focus may demand re-calibration of a bank's models as well.

**Figure 1 First Mortgage Defaults (US, percent)**




*Source:* S&P/Experian Consumer Credit Default Indices, available at standardandpoors.com

This paper is an important and helpful contribution, especially for the practitioner who needs to develop credit scoring models and test their performance. It reviews the methods used to measure the quality of credit scoring models based on both distribution functions and density functions, extends existing formulas, and illustrates their properties in a simulation as well as on real financial data. The real example is indeed very helpful in that it helps bring to life the measures described in the theory section and shows that improved decision making, based on good credit

* The views expressed here do not necessarily represent those of my current or previous employers.

scoring models, can be very substantial in financial terms. Rather than repeating all the points here, we would to highlight one interesting insight – while the most frequently used measures of model quality (the Gini index or Kolmogorov-Smirnov statistics) measure global model performance, local model performance near the cut-off point, which separates the approved and rejected clients, is the key performance indicator. The paper proposes a way to examine the local properties of a given model.

Real life brings multiple complications, however. First, even high-quality statistical models can go only so far. Coming back to *Figure 1*, one would not expect many credit models to have predicted the large jump in defaults on US mortgages. It will be always difficult for statistical models, which are intrinsically backward-looking, to predict discontinuities, especially when their signs may be well hidden in the data, sometimes even on purpose to allow higher prices of related securities in the market. Thus, frequent assessment of credit scoring quality can hardly be a substitute for economic intuition and rational economic thinking, even though a quick reaction to emerging credit problems or shifts in behavior could put a bank ahead of the competition and limit the inflow of new bad clients (the old bad clients will, of course, stay).

Second, both banks and researchers are often missing – as is the case in this paper – information on rejected clients, which may limit the optimization to too narrow a set of customers. Many other studies and, indeed, many banks are faced with limited information about rejected clients. Similarly as in this study, the focus is on optimization between "good" and "bad" clients, but all of these were actually granted credit and we are ignoring all of the rejected clients. It may not be an issue for the illustrative calculations in the paper, but in some cases – especially when banks are too conservative and try to minimize risk rather than manage it – banks may end up optimizing only on a set of good and very good clients. Moreover, even if the bank were to collect and store all the application information about the clients it has rejected, it will never be able to find out whether they would have been good or bad. In such cases, broadening the sample through a managed experiment, e.g., switching off the credit scoring models for a short period of time, could be the only way to extend the customer set. It would be interesting to explore the optimal way to do this, maximizing the information gain with limited credit costs due to lending to bad clients.

Third, decision-making in most banks is more complicated than credit scoring models are capable of encompassing. Banks often decide based on a combination of an application score (based on information from the credit application), a behavior score (based on information about existing client behavior), and K.O. rules (e.g., a current negative record in a credit bureau). The interaction of different credit scores, which may not give consistent recommendations, and K.O. rules, which automatically reject a set of clients (making it all but impossible to test their usefulness without a well-designed test strategy), is not trivial and may be an interesting area for further exploration.