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

Conservative Stress Testing: The Role of Regular Verification

Adam Geršl
Jakub Seidler

IES Working Paper: 12/2010

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:
Geršl, A., Seidler, J. (2010). " Conservative Stress Testing: The Role of Regular Verification" IES Working Paper 12/2010. IES FSV. Charles University.

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

Conservative Stress Testing: The Role of Regular Verification

Adam Geršl*
Jakub Seidler#

*Czech National Bank and
IES, Charles University Prague
E-mail: adam.gersl@cnb.cz

Czech National Bank and
IES, Charles University Prague
E-mail: jakub.seidler@cnb.cz

July 2010

Abstract:
This paper focuses on how to calibrate models used to stress test the most important risks in the banking system. Based on the results of a verification of the Czech National Bank's stress testing methodology, the paper argues that stress tests should be calibrated conservatively and slightly overestimate the risks. However, to ensure that the stress test framework is conservative enough over time, a verification, i.e. comparison of the actual values of key banking sector variables – in particular the capital adequacy ratio – with predictions generated by the stress-testing models should become a standard part of the stress-testing framework.

Keywords: stress testing; credit risk; bank capital

JEL: E44; E47; G21

Acknowledgements:
Authors thank Jan Frait, Michal Hlaváček, Tomáš Holub, Kateřina Šmídková, Karel Gabrhel and Petr Staněk for useful suggestions. Both authors acknowledge the support by the Grant Agency of the Czech Republic (GACR 403/10/1235). Adam Geršl also acknowledges the support by the Czech Ministry of Education (Grant MSMT 0021620841). The findings, interpretations and conclusions expressed in this paper are entirely those of the authors and do not represent the views of any of the above-mentioned institutions.

1. Introduction

Stress tests are used by central banks, regulators and commercial financial institutions as a means of testing the resilience of institutions or the entire sector to adverse changes in the economic environment. The global financial crisis uncovered the deficiencies of the stress-testing methodologies used in many countries. Before the crisis, many tests were wrongly indicating that the sector would remain stable even in the event of sizeable shocks (Haldane, 2009). These deficiencies related not only to the configuration of the adverse scenarios used, which had initially seemed implausibly strong but were often exceeded in reality, but also to the shock combination assumed, which had not been adequately anticipated in the scenarios (Čihák et al., 2009; Breuer et al., 2009). A role was also played by deficiencies in model calibration and in the assumed behaviour of banks and markets, and by the absence of testing of liquidity risk alongside traditional financial risks (in particular credit risk and interest rate risk).

Consequently, the assumptions and parameters used in stress tests are gradually being re-examined so that the tests can better analyse the impacts of strong shocks to the financial system. In defence of stress testing, however, it should be mentioned that this is a relatively new tool[1] and hence it still requires ongoing methodological development and refinement.[2]

This paper focuses on how to calibrate models used to stress test the most important risks in the banking system. We argue that stress tests should be calibrated conservatively and slightly overestimate the risks. However, to ensure that the stress test framework is conservative enough over time, a process of verification, i.e. comparison of the actual values of key banking sector variables with predictions generated by the stress-testing models should become a standard part of the stress-testing framework. Direct verification of adverse scenarios is in majority of cases (i.e. non-crisis periods) not possible. Thus, the verification should be performed on baseline scenarios. However, the whole stress-testing model should be calibrated conservatively in order to take into account the uncertainty related to the possible changes in estimated relationships in the case of adverse economic development. Hence, ex-post comparison between reality and predictions generated by baseline scenarios should indicate systematic risk overestimation.

To illustrate our point we present the results of the verification of the Czech National Bank's (CNB) stress testing framework. The CNB has been performing bank stress tests since 2003 and has significantly expanded its methodology over the past few years. The most recent major update was done in mid-2009 and involved and introduction of dynamic features in the system (see section 2). On this occasion, a verification of the overall stress-testing methodology was conducted in the context of the aforementioned international debate on the reliability of the predictions of the impacts of shocks to the banking sector. The aims were to

[1] Tools based on various types of financial soundness indicators have traditionally been used to assess the resilience of financial institutions (Geršl and Heřmánek, 2007).

[2] The formal obligation of commercial banks to conduct stress tests on their own portfolios was only introduced by Basel II (for banks using advanced methods for calculating capital requirements), which was implemented in the EU in 2006–2007. However, nowadays, there is a set of guidelines by CEBS related to stress testing in commercial banks (see CEBS 2009).

demonstrate whether the stress test assumptions were correctly configured and to identify any deficiencies in those assumptions.

The analysis reveals that the current CNB stress-testing system generally errs on the right – i.e. pessimistic – side and slightly overestimates the risks. This leads on average to estimates of key financial soundness indicators (in particular capital adequacy) that are lower (more conservative) than the actual values. Some verification results were used to further develop the stress tests.

To our knowledge, there is no other study that would systematically and transparently present the verification of someone's stress testing methodology. With this paper we would like to make a contribution to the debate on how to develop and calibrate reliable stress testing frameworks.

The paper is structured as follows. Section 2 briefly describes the CNB's stress-testing methodology as of end-2009 that was subsequently verified. Section 3 summarises the verification methodology and presents summary conclusions of the verification for capital adequacy (including its two main constituents, i.e. regulatory capital and risk-weighted assets, RWA) and some other key banking sector variables used in the stress tests. This section also contains a summary of the main improvements introduced following the verification and a brief description of the next steps planned for the development of the banking sector stress tests. The conclusion summarises the verification results and proposes a medium-term plan for further developing the tests.

2. Current banking sector stress-testing methodology of the CNB

The original banking sector stress-testing methodology applied at the CNB was based on the IMF methodology used for FSAP missions (Čihák, 2005; Čihák and Heřmánek, 2005). The CNB later switched from testing historical ad-hoc scenarios defined by a combination of shocks (e.g. a 20% rise in non-performing loans, a 15% exchange rate depreciation) to using consistent macroeconomic scenarios generated by the CNB's prediction model and related credit risk and credit growth sub-models (Čihák, Heřmánek and Hlaváček, 2007; Jakubík and Schmieder, 2008; Jakubík and Heřmánek, 2008). This framework was used for the previous Financial Stability Report 2008/2009 (CNB, 2009).

In the second half of 2009, the CNB significantly updated the banking sector stress-testing methodology in three respects. First, the tests were "dynamised", in the sense of switching to quarterly modelling of shocks and their impacts on banks' portfolios. This change was described in a box in the CNB Financial Stability Report 2008/2009 (CNB, 2009, pp. 63–64). Second, in the credit risk area there was a changeover to "Basel II terminology", i.e. to capturing the credit risk of several separate portfolios using the standard parameters PD, LGD and EAD and relating risk-weighted assets to those parameters using procedures specified in the IRB approach to calculating capital requirements.[3] The final major innovation was the extension of the shock impact horizon from one to two years (or eight subsequent quarters).

Alternative macroeconomic scenarios

Alternative macroeconomic scenarios still serve as the starting point for stress testing in the updated methodological framework. The scenarios are designed using the CNB's official prediction model supplemented with an estimate of the evolution of some additional variables, which are not directly generated by the model. "Stress scenarios" are constructed based on the identification of risks to the Czech economy in the near future. To compare the stress outcome

[3] PD – probability of default; LGD – loss given default; EAD – exposure at default; IRB – internal ratings based.

with the most probable outcome, the stress tests use a baseline scenario, i.e. the current official macroeconomic prediction of the CNB.

The predictions for GDP growth, inflation and other macroeconomic variables enter credit risk and credit growth models. They were developed to capture changes in banks' credit portfolios and credit risk. The stress tests work explicitly with the four main loan portfolio segments by debtor and/or credit type (non-financial corporations, loans to households for house purchase, consumer credit and other loans), to which the sub-models are also adjusted. The credit risk models are used to predict PD for the individual loan segments, whereas the credit growth models are used to estimate the growth in bank portfolios in relation to the macroeconomic situation and (after certain adjustments) to estimate the evolution of risk-weighted assets.

In the stress tests, the prediction for macroeconomic and financial variables for individual quarters is reflected directly in the prediction for the main balance-sheet and flow indicators of banks. The tests are dynamic, i.e. for each item of assets, liabilities, income and expenditure there is an initial (the last actually known) stock, to which the impact of the shock in one quarter is added/deducted, and this final stock is then used as the initial stock for the following quarter. This logic is repeated in all eight quarters for which the prediction is being prepared. The consistency between stocks and flows is thus ensured.

Credit risk

Credit risk testing is the most important area of stress testing. This testing is based on the use of PD for each of the four main segments of the loan portfolio. The second credit risk parameter is LGD, which is currently determined by expert judgement, with different amounts being set for different scenarios and different credit segments in line with the regulatory rules, commercial bank practices, the approaches applied by some rating agencies (Moody's, 2009) and existing estimates based on market data (Seidler and Jakubík, 2009). The third parameter is EAD, which is determined as the volume of the non-default part of the portfolio (i.e. excluding non-performing loans).

An increase in PD and LGD has two main effects on individual banks.

First, the expected loan losses (in CZK millions), against which banks will create new provisions of an equal amount and record them on the expenses side of the profit and loss statement as impairment losses, are calculated as the product of PD, LGD and EAD for each credit segment and quarter.[4] Total assets are then symmetrically reduced by the amount of these expenses.

The product of PD and the volume of the non-default portfolio forms the volume of new non-performing loans (NPLs) for each quarter. This allows us to generate the volume of total NPLs in the following eight quarters for each bank, and subsequently for the banking sector as a whole, according to the following equation:

$$NPL_{t+1} = NPL_t + \sum_{i=1}^{4} PD_{t+1,i} NP_{ti} - aNPL_t \qquad (1)$$

where *NPL* are non-performing loans, *PD* is the probability of default, *NP* is the non-default portfolio in the four segments defined above and *a* is an NPL outflow parameter (i.e. write-offs or sales of existing NPLs, i.e. the default part of the portfolio). Parameter *a* is set by

[4] According the relevant CNB decree and IFRS, banks are not required immediately to create provisions exactly equal to expected losses, but rather they must create provisions equal to realised losses, i.e. for new NPLs. However, if the loans are gradually reclassified during the quarter into the NPL (i.e. default) category to the extent predicted by PD, banks will ultimately create these provisions in the originally estimated amount.

expert judgement at 15% for all segments, i.e. 15% of NPLs are written off/sold each quarter and subsequently disappear from the total volume of NPLs and (gross) assets of the bank. This calibration was chosen on the basis of discussions with commercial banks and estimates conducted as part of the verification, which are described in more detail at the end of the next section.

The credit growth model leads to an estimate of the gross volume of loans in individual segments. Using relation (1) for NPL modelling, this allows us to determine for each bank, and subsequently for the banking sector as a whole, the NPL/total loans ratio, a standard indicator of the banking sector's health.

Second, in the case of banks applying the Basel II IRB approach to the calculation of capital requirements for credit risk, the capital requirements (or risk-weighted assets, RWA[5]) for credit risk are a function of PD, LGD and EAD. Given that the largest banks in the Czech Republic apply this approach, this relation is applied to all banks for the sake of simplicity. Given a constant non-default portfolio volume, i.e. EAD, an increase in PD and LGD thus generally results in an increase in RWA and therefore a decrease in capital adequacy.[6]

Interest rate and currency risk

The macroeconomic scenarios contain a prediction of the evolution of the simplified koruna and euro yield curves (rates with 3M, 1Y and 5Y maturities). A change in interest rates has a direct effect on bank balance sheets in two main items, namely interest profit and the value of bond holdings.[7] A rise in short-term rates thus reduces the interest rate profit of those banks which have an excess of short-term liabilities over short-term assets. However, the calculation is adjusted by expert judgement to take account of the business policies of commercial banks, which respond relatively little to market interest rate changes on the deposit side.

The prediction for long-term interest rates is used to estimate profits/losses from the revaluation of bond holdings (except for bonds held to maturity and bonds with a variable coupon dependent on interest rates). The calculation is based on the estimated duration of the bond portfolios, which is calculated by expert judgement on the basis of a more detailed knowledge of the maturity structure. Account is also taken of bond portfolio hedging using IRS (interest rate swaps), which for some banks lessens the impact of interest rate changes.

The quarter-on-quarter change in the CZK/EUR exchange rate is applied to the net open foreign currency position (including off-balance-sheet items), generating either a loss or a profit depending on the sign of the net open position and the direction of the exchange rate change.[8]

Interbank contagion risk

Interbank contagion risk is modelled in two selected periods (in the fourth and eighth quarters). The test uses data on interbank exposures, with the capital adequacy of individual banks being used to determine their probability of default (PD).[9] As interbank exposures are mostly unsecured, LGD is assumed to be 100%. The expected losses due to interbank

[5] Risk-weighted assets = capital requirements (in CZK millions)×12.5.

[6] This channel of the impact of increased PD and/or LGD on banks is one of the main sources of the much-criticised procyclicality of Basel II (see Geršl and Jakubík 2010).

[7] At the same time, however, interest rate changes have an indirect effect on credit risk via their effect on the PD estimate.

[8] For example, a positive open foreign currency position and appreciation of the koruna leads to losses.

[9] The PD values in relation to capital adequacy ratios (CAR) are set by expert judgement as follows: PD = 100% for negative CAR; PD = 25% for CAR between 0% and 5%; PD = 15% for CAR between 5% and 8%; PD = 5% for CAR between 8% and 10%; PD = 0.5% for CAR greater than 10%.

exposures are calculated for each bank according to the formula PD×LGD×EAD, where EAD is the net interbank exposure. If these losses are relatively high and will lead to a reduction in the bank's capital adequacy and thus an increase in its PD, there follows another iteration of the transmission of the negative effects to other banks through an increase in the expected losses. These iterations are performed until this "domino effect" of interbank contagion stops, i.e. until the rise in PD induced in one bank or group of banks does not lead to a rise in the PD of other banks.

Profit, regulatory capital and capital adequacy

The stress test assumes that banks will continue to generate revenues even in the stress period, particularly net interest income (interest profit) and net fee income. For these purposes, an analytical item of the profit and loss account called "adjusted operating profit" has been constructed. This consists of interest profit (+), fee profit (+), administrative expenses (−) and some other (non-shock) items.[10] The volume of adjusted operating profit was initially determined by expert judgment for the individual scenarios. A model estimate of this item was introduced only in mid-2010 (CNB 2010).

Regulatory capital is modelled in accordance with the applicable CNB regulations. Each bank enters the first predicted quarter with initial capital equal to that recorded in the last known quarter. If a bank generates a profit in the first predicted quarter (i.e. its adjusted operating profit is higher than its losses due to the shocks), its regulatory capital remains at the same level (is not increased). If, however, it generates a loss, its regulatory capital is reduced by the amount of that loss. The impacts of the shocks are thus reflected in a reduction of capital only if they exceed adjusted operating profit and the bank generates a loss.

It is assumed that those banks which generate a profit for the entire financial year will decide on profit distribution and dividend payments in the second quarter of the following year. Here we assume that each bank, when increasing its capital from retained earnings of the previous financial year, will try to get to its initial capital adequacy ratio if its previous year's profits are sufficient.[11] Depending on the change in RWA, several scenarios are thus possible:

(a) the bank distributes the entire profit and does not strengthen its regulatory capital (in the event of unchanged RWA);

(b) the bank uses part of its profit to strengthen its capital and distributes the remainder (in the event of an increase in RWA; however, the entire retained earnings of the previous year will not be needed to reach the initial level of capital);

(c) the bank uses the entire profit to strengthen its capital (in the event of a relatively sizeable increase in RWA); depending on the size of the increase in RWA, however, it may not reach the original capital adequacy ratio;

(d) the bank pays dividends that exceed the profit generated (in the event of a decrease in RWA) and thereby also distributes part of retained earnings of previous years.

Total capital adequacy is then calculated for the individual quarters as the ratio of regulatory capital to total RWA. The portion of RWA relating to credit risk is modelled on the basis of

[10] In previous Financial Stability Reports this adjusted operating profit was called "net income". Adjusted operating profit is broadly equivalent to the item "pre-provision profit", i.e. operating profit gross of losses on non-performing loans, but differs in that it does not include the impacts of other (interest rate and exchange rate) shocks, whereas pre-provision profit does.

[11] This assumption may not be very realistic at certain times, as banks may decide to pay higher dividends and reduce their capital adequacy ratio below the initial level.

the credit risk parameters (see above), while the other components of RWA (or of the capital requirements for other risks) for the individual quarters are determined by expert judgement.

3. Verification of the stress tests

The objective of the verification is to examine to what extent the assumptions and sub-models used in the stress testing framework are in line with reality. A problematic aspect of the verification is that the tests use stress – i.e. unlikely – scenarios, which may not occur in reality. Hence, we cannot subsequently compare predictions based on adverse scenarios with reality. For this reason, only the scenario that represents the most likely evolution of the economic environment, i.e. the no-stress baseline scenario of the CNB forecast, could be used for the verification.[12]

The prediction using the baseline (i.e. likely) scenario should indicate slightly higher risks than those that occur in reality. This is because the whole system should have a "conservative" buffer to offset the uncertainty associated with estimating losses given adverse economic developments, when relations (for example between GDP growth and risk parameters such as PD) estimated by standard econometric techniques on data from mainly calm periods can change suddenly for the worse. This requirement implies that stress test prediction errors should be evaluated differently from the errors of standard macroeconomic predictions, where deviations in either direction are regarded as "equally bad". In verifications using baseline scenarios, it is appropriate to apply an asymmetric view in the stress tests and tolerate prediction errors towards modest overestimation of the risks.

The verification was conducted on quarterly data in the period 2004 Q4–2009 Q2, i.e. for 19 periods in all. The actual values of key variables for the banking sector as a whole are compared with the predictions generated by the current stress-testing methodology for the individual quarters using the relevant baseline scenario of the forecast. As the updated stress-testing methodology allows us to create a prediction for the next eight quarters, it was necessary to choose a prediction horizon. The results presented in this paper are based on a one-year prediction.[13] The predictions for past quarters were therefore created subsequently using the updated stress-testing methodology in order to verify that methodology and do not match the values published in CNB Financial Stability Reports.

Two statistics based on the mean prediction errors were used to verify the selected variables: the mean absolute error (MAE) defined by equation (2):

$$\text{(2)} \qquad \frac{1}{n}\sum_{t=1}^{n}\left|P_t - A_t\right|$$

and the mean error in direction (MED) defined as:

$$\text{(3)} \qquad \frac{1}{n}\sum_{t=1}^{n}\frac{P_t - A_t}{\left|A_t\right|},$$

[12] The first attempt to verify the stress tests using the baseline forecast scenario was made back in 2007 (Hlaváček et al., 2007), when the capital adequacy ratio and NPL growth predictions generated by the 2006 stress-testing methodology were compared with their real counterparts.

[13] This means, for example, that the actual outcome in 2007 Q4 was compared with the prediction for that quarter made one year earlier, i.e. on bank portfolios as of 2006 Q4 using the January 2007 baseline scenario. Internally, however, the verification was performed for all prediction horizons and the results are qualitatively similar.

where P_t denotes the value of the prediction of the estimated variable for the given quarter, A_t denotes the actual value and t represents the quarter for which the prediction is being made.[14]

MAE serves for simple presentation of the mean prediction error in the units in which the given variable is expressed, while MED expresses whether the given variable was overestimated or underestimated on average and thus gives the degree of "conservativism".

The prediction error of the capital adequacy ratio and other key banking sector variables can be split into two main factors. The first is the potential prediction error caused by inaccuracy in the estimates of the macroeconomic variables entering the stress-testing mechanism (interest rates and the exchange rate), and the second concerns the assumptions and sub-models used in the stress test itself (e.g. the assumptions about how the bank raises its regulatory capital, what interest and non-interest yields it achieves and how sensitive it is to interest rate risk). The macroeconomic prediction error can be eliminated in the verification by using the actual (ex post) values of macroeconomic variables. The residual error is then due to inaccuracies in the assumptions and sub-models of the stress-testing framework and the intentional conservative buffer.

The most important output variable of the tests is the estimate of the capital adequacy ratio (CAR). The mean absolute deviation (MAE) for CAR equates to roughly 1.6 p.p. of the capital adequacy ratio (see Table 1). This means, for example, that the test predicts CAR of 11.4% instead of 13%.

Table 1: Deviation of capital adequacy ratio estimate
Estimate for 1-year horizon

Mean absolute error (MAE)	2004–2009	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009
Prediction – stress test	1,6	1,0	0,8	1,6	2,1	1,9
Prediction – known macro	1,5	0,9	0,6	1,1	2,0	2,5
Mean error in direction (MED) in %						
Prediction – stress test	-10,8	-1,7	-6,5	-13,1	-17,2	-15,3
Prediction – known macro	-8,8	1,9	-1,3	-7,1	-16,3	-20,0

This prediction error equates to roughly 1.8 standard deviations. In the individual shorter periods this error gradually shrinks to 0.8 p.p. (i.e. 1 standard deviation) but then grows again slightly from 2007 onwards. Only a small part of the error is due to errors in the macroeconomic forecast, as the MAE statistic decreases only modestly with knowledge of actual macroeconomic developments.

The negative MED statistic of -10.8% shows that the real values were higher on average in the period as a whole and the stress tests thus tended to generate overvalued CAR estimates (see Table 1). This fact is also demonstrated by Chart 1, which reveals that a lower-than-actual CAR is predicted from the end of 2006 onwards. The resulting CAR was thus underestimated for most periods, in line with the conservative design of the tests. This conclusion remains valid even when the predictions are adjusted for the error in the prediction of macroeconomic variables.

[14] As part of the verification we also computed other prediction error statistics, e.g. the mean percentage error, the mean weighted percentage error, the mean quadratic error and the mean percentage quadratic error. The verification results using these statistics, however, did not differ significantly from the results using MAE and MED, which are easier to interpret.

Chart 1: Verification of CAR estimate
(CAR in %; estimate for 1-year horizon)

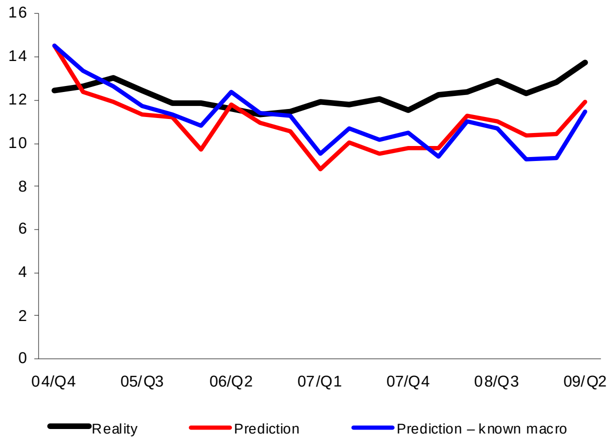
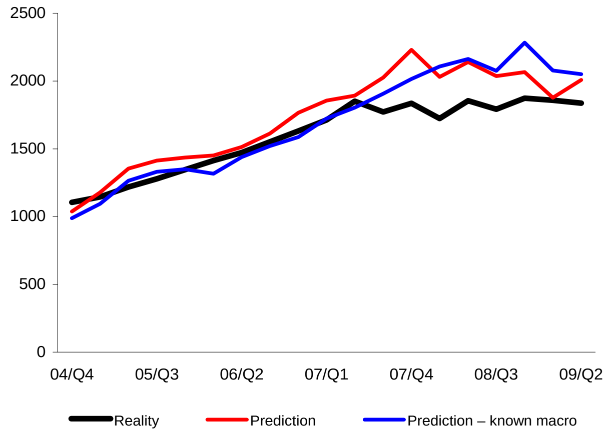


Chart 2: Verification of RWA estimate
(RWA in CZK billions; estimate for 1-year horizon)



The estimate of a lower-than-actual CAR is due to inaccuracy in the estimate of both RWA and regulatory capital. With few exceptions the stress test overestimated RWA (see Chart 2) and simultaneously tended to underestimate regulatory capital (see Chart 3). The decomposition of the error in the CAR estimate into the part caused by inaccurate prediction of RWA and the part caused by inaccurate prediction of regulatory capital shows that the contributions of the two items to the error are balanced on average. The overestimation of risk-weighted assets has two sources: first, the credit growth model tends to predict higher credit volumes than the ex-post turnout. While on a first sight an underestimation of credit growth seems to be the conservative calibration, the opposite is true at least from the point of view of risk-weighted assets. Second, the framework uses the estimates of PDs and LGDs as a base of risk weights (IRB approach) which are also overestimated.

Chart 3: Verification of regulatory capital estimate
(in CZK billions; estimate for 1-year horizon)

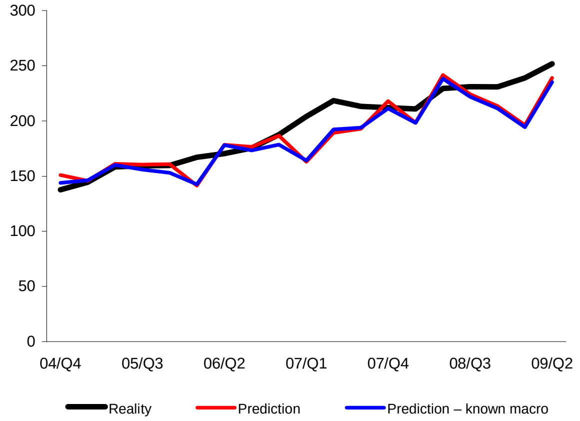
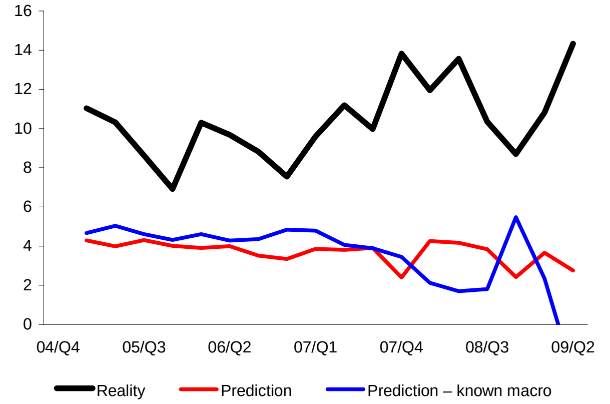
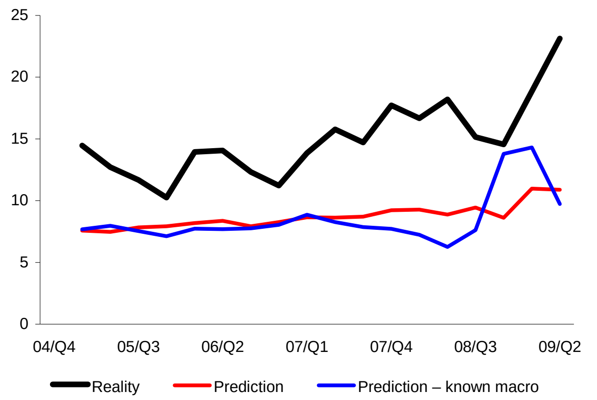


Chart 4: Verification of profits estimate
(quarterly figures, in CZK billions; estimate for 1-year horizon)



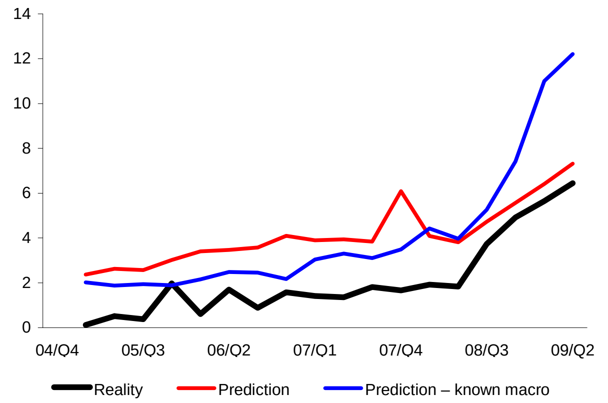
Regulatory capital is regularly increased out of after-tax profits, so the estimate of profits is an important parameter for the evolution of capital. Profits are calculated as the difference between adjusted operating profit and losses due to the individual shocks tested (see section 2). The verification of this variable revealed that the stress test systematically underestimates after-tax profit (Chart 4). This is due to two factors. First, the test systematically underestimates adjusted operating profit directly through the assumption about its level (for the baseline it was assumed that adjusted operating profit will be 90% of the average for the previous two years). This is also in line with the more conservative approach to risk assessment (Chart 5). The second cause is that the stress test tends to overestimate the impact of the main risk tested, i.e. credit risk, in the form of higher-than-actual PD and related higher provisioning for NPLs (recorded in the "losses from impairment" category), partly also due to a too conservative expert estimates of LGD (Chart 6).

Chart 5: Verification of pre-provision profit estimate
(quarterly figures, in CZK billions; estimate for 1-year horizon)



Note: Pre-provision profit equals adjusted operating profit + impact of market profits/losses (interest rate and FX risk).

Chart 6: Verification of credit losses estimate
(quarterly figures, in CZK billions; estimate for 1-year horizon)



The NPL ratio is a closely monitored financial stability indicator. We therefore present detailed verification results for this variable as well. A comparison of the actual NPL ratios with their predicted values reveals overshooting of the estimates, especially since the end of 2007, for both non-financial corporations (see Chart 7) and households (see Chart 8).

Chart 7: Verification of NPL ratio – corporations
(in %; estimate for 1-year horizon)

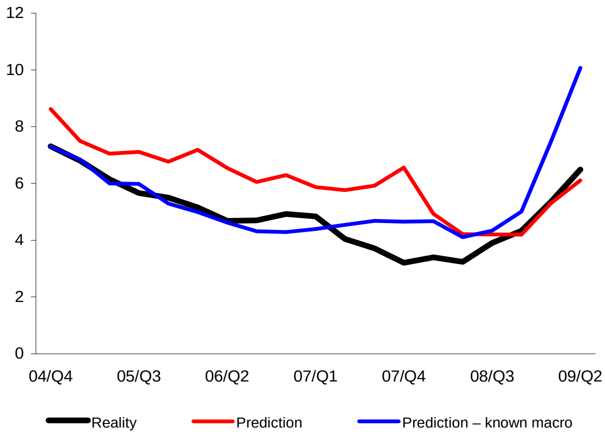


Chart 8: Verification of NPL ratio – households
(in %; estimate for 1-year horizon)

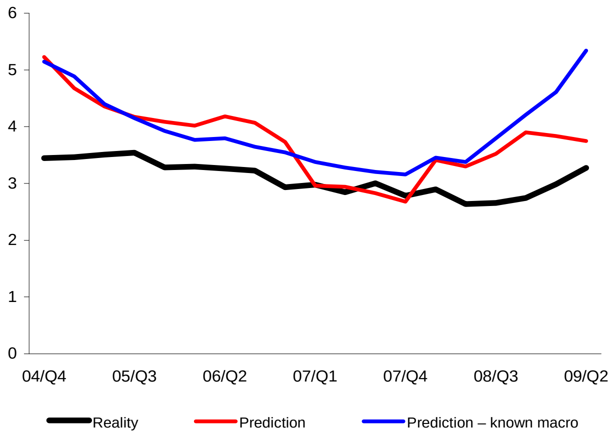


Table 2 shows that MAE was around 1.3 p.p. for non-financial corporations and 0.7 for households. While the NPL estimates for corporations improve significantly with knowledge of the macroeconomic environment, the opposite is true for households in some periods. In overall comparison, however, the household NPL estimate is more accurate.

Table 2: Deviation of NPL ratio estimate for corporations and households
Estimate for 1-year horizon

NPL ratio – corporations

Mean absolute error (MAE)	2004–2009	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009
Prediction – stress test	1.3	1.1	1.4	1.9	1.4	0.6
Prediction – known macro	0.8	0.1	0.2	0.6	0.8	1.5

Mean error in direction (MED) in %	2004–2009	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009
Prediction – stress test	27.8	18.3	26.2	45.5	38.5	12.1
Prediction – known macro	12.3	-0.1	-3.2	6.1	20.6	31.0

NPL ratio – households

Mean absolute error (MAE)	2004–2009	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009
Prediction – stress test	0.7	1.1	0.8	0.5	0.4	0.8
Prediction – known macro	0.9	1.1	0.7	0.4	0.7	1.3

Mean error in direction (MED) in %	2004–2009	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009
Prediction – stress test	21.6	30.7	25.6	12.1	13.9	26.7
Prediction – known macro	27.7	30.5	21.0	14.2	24.1	43.5

The overestimation of the NPL ratio is due both to the aforementioned conservative calibration of the PD risk parameter and, to some extent, to underestimation of outflow parameter a from equation (1). To determined the optimum value of a, numerical minimisation of the MAE error statistic was performed in various time intervals of 2004–2009. The optimum outflow a for the entire period under review was 20% on average. Owing to the deliberate overestimation of the potential risks this parameter was conservatively set at 15% in the tests.[15]

Despite the relatively positive message of the verification results, further gradual refinement of the predictions is desirable. The main problem in the credit risk area is with the sub-models and assumptions used, as they excessively overestimate the impact of credit risk in the form of losses on impaired loans. While the direction towards overestimation is correct, the degree of overestimation should be held in a reasonable range.[16] At the same time, the conservative prediction of adjusted operating income (and, as a result, overall profits) seems to be too far from the ex-post reality, so adjustments in this area are also needed.

Following the verification, the CNB has started to modify the stress testing framework in order to bring estimates closer to the reality, while still preserving a certain degree of conservativism (CNB 2010). In the credit risk area, this involved a recalibration of the credit risk models, linking the parameter LGD to the macroeconomic environment and better prediction of risk-weighted assets. In the profits area, a new bank income model linking adjusted operating profit to developments in the macroeconomic environment was developed (CNB 2010, Box 7).

When recalibrating and adjusting the overall stress testing framework, there are four main ways to preserve a conservative buffer in estimates. The CNB has been using all of them in combination. First, if a parameter is set expertly (such as adjusted operating profits or the LGDs in the verified version of the framework), it should be set conservatively. Second, if a parameter is estimated via a model, a more conservative definition of the estimated parameter (i.e. the dependent variable) could be used. The CNB has for example used a PD that was

[15] The sensitivity of the NPL ratio estimate to change in a reveals that an increase in a of 5 p.p. (i.e. from the 15% used to the optimum value of 20%) – i.e. a faster outflow of NPLs from banks' balance sheets – causes on average a decline in the NPL ratio of one-tenth (e.g. from 10% to 9%).

[16] The results of verification of other key variables (not reported here, but available from authors upon request) indicated that next to a large overestimation of credit losses, market losses (FX and bond revaluations) are also to some extent overestimated.

based on 30+ days in arrears definition of default rate which is generally higher than the standard Basel 90+ days definition.[17] Third, if the parameter (dependent variable) is correctly calculated, the buffer can be achieved by changing (to the worse) some of the coefficients (elasticities) estimated within an econometric model that is usually using data over a calm period. And finally, while the model itself is estimated by a traditional econometric method, predictions could be adjusted by one standard deviation of the volatility of dependent variable (in the conservative direction). The last way has been used in the new modelling of adjusted operating income of banks in order to intentionally underestimate the income capacity of banks in the stress periods.

The further development of the stress tests should be based on regular verification. This should become an integral part of the banking sector stress-testing framework to enable ongoing assessment of whether the assumptions are realistic and a conservative buffer is being maintained in the risk predictions.[18]

4. Conclusion

This paper focused on how to calibrate parameters used in stress tests. It argued that the parameters should be calibrated conservatively and should slightly overestimate risks in order to take into account the uncertainty related to the possible changes in estimated elasticities in the case of adverse economic development. This means that the ex-post comparison between reality and predictions generated by baseline scenarios should indicate systematic risk overestimation.

We used the case study of the CNB's banking sector stress-testing methodology and presented the results of a verification of that methodology. Such verification is a tool that should be used regularly as a guide for refining the assumptions and models used. The results of the verification, conducted at the end of 2009, reveal that the CNB stress tests err on the right – i.e. pessimistic – side and slightly overestimate the risks. This leads on average to capital adequacy estimates that are lower (more conservative) than the actual values. This is consistent with the design of the stress tests, which should be built on conservative assumptions. However, account should be taken of the fact that the level of conservatism, i.e. the degree of overestimation of the risks, in the methodology can only be fully assessed after the effects of the current recession disappear.

The verification results also indicated areas where further refinement of the stress tests is desirable. The main such areas are credit risk (more accurate estimates of PD and LGD), modelling of bank income in relation to the macroeconomic scenario, better estimation of risk-weighted assets, and certain enhancements in calculating the impacts of market risks. These areas were already to some extent tackled in the newest version of the CNB's stress testing framework as presented in the FSR 2009/2010 (CNB 2010).

As to further development of banking sector stress-testing framework as applied by the central banks, there remain two main medium-term challenges which were not discussed in detail in the paper. The first challenge is to incorporate the feedback effect of a weakened banking sector on the economy in the form of a radical decline in the supply of loans – known as deleveraging – and the related impact on the economy. A first attempt of incorporating a feedback effect for the CNB stress testing framework has been presented in Geršl and Jakubík

[17] Given the results of the verification as to the large overestimation of credit losses, the CNB changed to the standard 90+ definition of default rate from June 2010 (CNB 2010). However, it still includes some conservative margin.

[18] Regular verification – i.e. retrospective assessment of prediction performance – is also routinely performed as part of the creation of predictions for monetary policy purposes – see for example CNB (2008).

(2010). The second challenge is to integrate credit, market and balance-sheet/funding liquidity risks in one overall framework, ideally in parallel with the interbank contagion test.

References

BREUER, T., JANDACKA, M., RHEINBERGER, K., SUMMER, M. (2009): *How to Find Plausible, Severe and Useful Stress Scenarios*, International Journal of Central Banking, Vol. 5, pp. 205–224.

CEBS (2009): *Guidelines for Stress Testing*, Committee of European Banking Supervisors, 2009.

CNB (2010*): Financial Stability Report 2009/2010*, Czech National Bank, 2010.

ČIHÁK, M. (2005): *Stress Testing of Banking Systems*, Czech Journal of Economics and Finance, Vol. 55, pp. 418–440.

ČIHÁK, M., HEŘMÁNEK, J. (2005): *Stress Testing the Czech Banking System: Where Are We? Where Are We Going?* Research and Policy Notes 2005/02, Czech National Bank, Research Department.

ČIHÁK, M., HEŘMÁNEK, J., HLAVÁČEK, M. (2007): *New Approaches to Stress Testing the Czech Banking Sector*, Czech Journal of Economics and Finance, Vol. 57, pp. 41–59.

ČIHÁK, M., TIEMAN, A., ONG, L. (2009): *Stress Tests: What Have We Learned from the Global Financial Crisis?* MCM Seminar, International Monetary Fund, June 2009.

CNB (2008): *Evaluation of the Fulfilment of the CNB's Inflation Targets 1998–2007*, Czech National Bank, 2008.

CNB (2009): *Financial Stability Report 2008/2009*, Czech National Bank, 2009.

GERŠL, A., HEŘMÁNEK, J. (2008): *Indicators of financial system stability: towards an aggregate financial stability indicator?* Prague Economic Papers 2/2008, pp. 127-142.

GERŠL, A., JAKUBÍK, P. (2010): *Procyclicality of the Financial System and Simulation of the Feedback Effect,* Financial Stability Report 2009/2010, Czech National Bank, pp. 110-119.

HALDANE, A. G. (2009): *Why Banks Failed the Stress Test*, Speech given at the Marcus Evans Conference on Stress-Testing, 9–10 February 2009.

JAKUBÍK, P., HEŘMÁNEK, J. (2008): *Stress Testing of the Czech Banking Sector*, Prague Economic Papers, Vol. 3, pp. 195–212.

JAKUBÍK, P., SCHMIEDER, C. (2008*): Stress Testing Credit Risk: Is the Czech Republic Different from Germany?* CNB Working Paper 9/2008.

MOODY'S (2009): *Approach to Estimating Czech Banks' Credit Losses*, Moody's Global Banking, July 2009.

SEIDLER, J., JAKUBÍK, P. (2009): *Estimation of Expected LGD*, Financial Stability Report 2008/2009, Czech National Bank.

IES Working Paper Series

2009

1. František Turnovec : *Fairness and Squareness: Fair Decision Making Rules in the EU Council?*
2. Radovan Chalupka : *Improving Risk Adjustment in the Czech Republic*
3. Jan Průša : *The Most Efficient Czech SME Sectors: An Application of Robust Data Envelopment Analysis*
4. Kamila Fialová, Martina Mysíková : *Labor Market Participation: The Impact of Social Benefits in the Czech Republic*
5. Kateřina Pavloková : *Time to death and health expenditure of the Czech health care system*
6. Kamila Fialová, Martina Mysíková : *Minimum Wage: Labour Market Consequences in the Czech Republic*
7. Tomáš Havránek : *Subsidy Competition for FDI: Fierce or Weak?*
8. Ondřej Schneider : *Reforming Pensions in Europe: Economic Fundamentals and Political Factors*
9. Jiří Witzany : *Loss, Default, and Loss Given Default Modeling*
10. Michal Bauer, Julie Chytilová : *Do children make women more patient? Experimental evidence from Indian villages*
11. Roman Horváth : *Interest Margins Determinants of Czech Banks*
12. Lenka Šťastná : *Spatial Interdependence of Local Public Expenditures: Selected Evidence from the Czech Republic*
13. František Turnovec : *Efficiency of Fairness in Voting Systems*
14. Martin Gregor, Dalibor Roháč : *The Optimal State Aid Control: No Control*
15. Ondřej Glazar, Wadim Strielkowski : *Turkey and the European Union: possible incidence of the EU accession on migration flows*
16. Michaela Vlasáková Baruníková : *Option Pricing: The empirical tests of the Black-Scholes pricing formula and the feed-forward networks*
17. Eva Ryšavá, Elisa Galeotti : *Determinants of FDI in Czech Manufacturing Industries between 2000-2006*
18. Martin Gregor, Lenka Šťastná : *Mobile criminals, immobile crime: the efficiency of decentralized crime deterrence*
19. František Turnovec : *How much of Federalism in the European Union*
20. Tomáš Havránek : *Rose Effect and the Euro: The Magic is Gone*
21. Jiří Witzany : *Estimating LGD Correlation*
22. Linnéa Lundberg, Jiri Novak, Maria Vikman : *Ethical vs. Non-Ethical – Is There a Difference? Analyzing Performance of Ethical and Non-Ethical Investment Funds*
23. Jozef Barunik, Lukas Vacha : *Wavelet Analysis of Central European Stock Market Behaviour During the Crisis*
24. Michaela Krčílková, Jan Zápal : *OCA cubed: Mundell in 3D*
25. Jan Průša : *A General Framework to Evaluate Economic Efficiency with an Application to British SME*
26. Ladislav Kristoufek : *Classical and modified rescaled range analysis: Sampling properties under heavy tails*
27. Natálie Švarcová, Petr Švarc : *Diffusion Processes on Complex Networks*

28. Goran Serdarević, Petr Teplý : Efficiency of EU Merger Control in the 1990-2008 Period
29. Jiri Novak, Dalibor Petr : *Empirical Risk Factors in Realized Stock Returns*
30. Karel Janda, Eva Michalíková, Věra Potácelová : *Vyplácí se podporovat exportní úvěry?*
31. Karel Janda, Jakub Mikolášek, Martin Netuka : *The Estimation of Complete Almost Ideal Demand System from Czech Household Budget Survey Data*
32. Karel Janda, Barbora Svárovská : *Investing into Microfinance Investment Funds*

2010

1. Petra Benešová, Petr Teplý : *Main Flaws of The Collateralized Debt Obligation's Valuation Before And During The 2008/2009 Global Turmoil*
2. Jiří Witzany, Michal Rychnovský, Pavel Charamza : *Survival Analysis in LGD Modeling*
3. Ladislav Kristoufek : *Long-range dependence in returns and volatility of Central European Stock Indices*
4. Jozef Barunik, Lukas Vacha, Miloslav Vosvrda : *Tail Behavior of the Central European Stock Markets during the Financial Crisis*
5. Onřej Lopušník : *Různá pojetí endogenity peněz v postkeynesovské ekonomii: Reinterpretace do obecnější teorie*
6. Jozef Barunik, Lukas Vacha : *Monte Carlo-Based Tail Exponent Estimator*
7. Karel Báťa : *Equity Home Bias in the Czech Republic*
8. Petra Kolouchová : *Cost of Equity Estimation Techniques Used by Valuation Experts*
9. Michael Princ : *Relationship between Czech and European developed stock markets: DCC MVGARCH analysis*
10. Juraj Kopecsni : *Improving Service Performance in Banking using Quality Adjusted Data Envelopment Analysis*
11. Jana Chvalkovská, Jiří Skuhrovec : *Measuring transparency in public spending: Case of Czech Public e-Procurement Information System*
12. Adam Geršl, Jakub Seidler : *Conservative Stress Testing: The Role of Regular Verification*

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