F11[1] Frankfurt a. Main, 28 July 2015

Macro stress tests – Technical Documentation

1. Stresstesting framework: an overview

The Bundesbank regularly carries out macro stress tests for banks in Germany. Macro stress tests are an instrument to assess banks' vulnerabilities against macroeconomic shocks in contrast to micro stress tests which focus on institution-specific risks. For the calculation of the stress effects, the Bundesbank uses data collected for supervisory purposes and other data. The process of conducting macro stress tests can then be divided into four steps:

1. Designing the macro scenarios (a baseline and one or more stress scenarios)
2. Linking macro scenarios to the banks' financial assets and liabilities
3. Aggregating direct or first-round stress effects
4. Assessing second-round effects

This documentation deals with the steps 2 to 4. Taking the macro scenarios as given, we start in Chapter 2 by presenting the link between the macro scenarios and the banks' financial risk exposures. In Chapter 3, we describe the algorithm for aggregating the direct results from the stress test. Chapter 4 is devoted to the second-round effects.

2. Satellite Models

A key question of any macro stresstest is how macroeconomic scenarios are linked to the banks' risk exposures (balance sheets and profit and loss accounts). In the case of 'bottom-up'-stress tests, the central bank or supervisor provides the general macroeconomic scenarios and the bank calculates their potential losses with their internal risk models. In the case of 'top-down'-stress tests the calculations are carried out by the central bank or by the supervisory authorities based on available profit and loss and balance sheet data. This documentation describes the Bundesbank's top-down approach to stress testing. The models that are subsequently presented are so-called satellite models. No feedback effects are modelled to the real economy. Table 1 gives a brief overview of the models for the different risk categories. As the business models of the large and small banks differ substantially, we apply different satellite models for the two types of banks.

[1] With parts from F 14.

	small banks	large banks
Interest income	*panel regression (Section 2.1)*	*panel regression (Section 2.1)*
Interest expense	*panel regression (Section 2.1)*	*panel regression (Section 2.1)*
Net-fee income	*panel regression (Section 2.1)*	*panel regression (Section 2.1)*
Trading income	---	*quantile analysis (Section 2.3)*
Credit risk (including real estate)	*sectoral regression (Section 2.2)* *model for mortgage losses (Section 2.4)*	*panel regression (Section 2.1)* *model for mortgage losses (Section 2.4)*

Table 1: components of bank income and econometric models

2.1 Panel regression

2.1.1 Model and estimation

To model banks' earnings empirically, we employ the following panel data model, which has received much attention in the related literature[2]:

$$y_{it} = \alpha + \phi y_{it-1} + \beta' x_{it} + \gamma' z_t + u_{it} \,,$$
$$u_{it} = \mu_i + \epsilon_{it} \,.$$

Here, index $i = 1, 2, \dots, N$ refers to the cross-section units (banks) and $t = 2, 3, \dots, T$ provides the time index (at quarterly or yearly frequency). The dependent variable y_{it} denotes a specific income component of a bank, such as its interest income (relative to total assets). This component is viewed as a dynamic adjustment process, such that its lag determines the level of the income component today. In addition, bank-specific variables x_{it}, usually based on balance sheet information, help explain the income of a given bank.

The macroeconomic macro variables z_t play a major role in this approach. These variables describe the relevant macroeconomic scenario that affects the whole set of banks.
The remaining determinants of bank income are denoted u_{it}, which has two parts. The bank-specific (fixed) effect μ_i represents unobserved characteristics that may differ between banks, but are constant over time. Finally, ϵ_{it} is an idiosyncratic error term.

The regression analysis involves two steps. First, using the observed variables y_{it}, x_{it} and z_t, the parameters α, ϕ, β and γ are estimated. The model is estimated using the GMM estimator proposed by Arellano and Bond (1991). Details of the estimation procedure are relegated to Appendix 3. Second, forecasts of the bank's income components are obtained for future periods, which is explained next.

[2] See, among others, Covas et al. (2014).

2.1.2 Prediction and forecast uncertainty

We now turn to prediction in this model, assuming that the model has been estimated for the sample indexed by the time periods $t = 2, 3, \ldots, T$. The forecast of the income component in time $T + h$ is then given by

$$\hat{y}_{i,T+h} = \hat{\alpha} + \hat{\mu}_i + \hat{\phi}\hat{y}_{i,T+h-1} + \hat{\beta}' x_{i,T+h} + \hat{\gamma}' z_{T+h} ,$$

for $h = 1,2, \ldots, H$, where $\hat{\alpha}, \hat{\phi}, \hat{\beta}$ and $\hat{\gamma}$ are the estimated parameters resulting from the GMM approach, including the macroeconomic elasticicities $\hat{\gamma}$, and $\hat{y}_{i,T} = y_{iT}$. Here, H is a given forecast horizon (e.g. 3 years).

In our specific application, the bank-specific regressors typically include balance sheet information, which is considered to be sticky in the short run. Hence we use the latest available observations for these variables as proxies for the realisations in the future. The time paths of macroeconomic variables are obtained in a separate model, which is not discussed in this document. These paths deliver the baseline and stress scenarios of macroeconomic factors.

Recall that the GMM estimator is derived in a transformed model, where all variables are considered in first difference (see Appendix 3 for details). Hence parameters attached to variables that are constant over time cannot be estimated, including in particular the intercept α. To produce meaningful forecasts, it is important to approximate the level of the dependent variable of interest, and thus a suitable estimator for the intercept is needed. Therefore, we first determine

$$\tilde{y}_{it} = \hat{\phi} y_{i,t-1} + \hat{\beta}' x_{it} + \hat{\gamma}' z_t$$

and examine the estimator

$$\hat{\alpha} = \frac{1}{NT} \sum_i \sum_t (y_{it} - \tilde{y}_{it})$$

It is easy to see that

$$\hat{\alpha} = \alpha + \frac{1}{N} \sum_i \mu_i + R_{NT} ,$$

where the remainder term R_{NT} is given by

$$R_{NT} = (\phi - \hat{\phi})\bar{y}_{NT} + (\beta - \hat{\beta})' \bar{x}_{NT} + (\gamma - \hat{\gamma})' \bar{z}_T + \bar{\epsilon}_{NT}$$

with $\bar{\epsilon}_{NT} = (NT)^{-1} \sum_i \sum_t \epsilon_{it}$ and analogous expressions apply for the other terms in R_{NT}. Since the GMM estimator is consistent, as the sample size (i. e. the cross-section dimension) increases, the remainder can be neglected. That is, for sufficiently large N, we can treat

$R_{NT} \approx 0$. Moreover, in the FE model, $N^{-1} \sum_i \mu_i \approx 0$ for sufficiently large N and so $\hat{\alpha}$ is a suitable estimator of the intercept.

Using this estimator, the bank- specific effects can be estimated as well. Let

$$\hat{\mu}_i = \frac{1}{T}\sum_t \left(y_{it} - (\hat{\alpha} + \tilde{y}_{it})\right)$$

such that

$$\hat{\mu}_i = \mu_i + \frac{1}{T}\sum_i \epsilon_{it} + R_T \,,$$

with

$$R_T = (\alpha - \hat{\alpha}) + (\phi - \hat{\phi})\bar{y}_T + (\beta - \hat{\beta})'\bar{x}_T + (\gamma - \hat{\gamma})'\bar{z}_T \,,$$

where now $\bar{x}_T = T^{-1} \sum_t x_{it}$. Again, due to the consistency of the GMM estimator, R_T is asymptotically negligible. Provided a suitable law of large numbers applies to the sample average $\bar{\epsilon}_T$, we have $\hat{\mu}_i \approx \mu_i$. This approximation is meaningful if the time dimension is sufficiently large such that the sample average $\bar{\epsilon}_T$ is close to its mean zero. In many classical panels, the time dimension is small (say $T = 5$) and the fixed effect μ_i cannot be estimated consistently. In our case, the time dimension comprises $T = 19$ periods for the small banks and $T = 24$ periods for the large banks, which we consider just large enough to justify $\hat{\mu}_i$ as an appropriate estimator of the bank-specific effect. Using the estimate of the intercept and the bank-specific effect, we can construct forecasts of the income component as described above.

The models described above explain only a fraction of future changes in the income components (see Section 3.3 where we deal with bank-idiosyncratic effects) and the estimated coefficients are subject to estimation errors.[3] In the following, we attempt to assess the resulting forecasting error. We do so by introducing a simplified model, then transferring the results to our models described above. Consider the following (univariate, i.e. without bank dimension) autoregressive process

$$y_{t+1} = \alpha + \beta y_t + \gamma' x_{t+1} + u_{t+1}.$$

[3] In addition, there exist other sources of uncertainty, for instance model misspecifications or structural breaks, which can be quite substantial especially if the forecast horizon is long. However, in the context of this documentation, we do not deal with these other sources.

We are interested in constructing forecasts of y_{t+j} for $j = 1,2,\dots,H$ at time t, where H is a given forecast horizon. To do so, it is important to introduce the relevant information set available at time t. Suppose that the current information set involves $y_t, y_{t-1}, \dots, u_t, u_{t-1}, \dots$ as well as all current, past and *future* x_t (hence treating these regressors as deterministic). Then by recursive substitution, the optimal forecast under quadratic loss is given by the conditional mean, i.e.

$$f^*_{t,t+j} = \alpha\left(1+\beta+\ \dots + \beta^{j-1}\right) + \gamma'\left(x_{t+j} + \beta x_{t+j-1} + \ \dots + \beta^{j-1}x_{t+1}\right) + \beta^j y_t$$

where $f^*_{t,t+j}$ denotes the (mean-squared error) optimal forecast of y_{t+j} made at time t, resulting in a forecast error of the form

$$e^*_{t,t+j} = u_{t+j} + \beta u_{t+j-1} + \beta^2 u_{t+j-2} + \ \dots + \beta^{j-1}u_{t+1}$$

Provided that the sequence of idiosyncratic errors u_{t+1} is normally distributed with variance σ^2 and serially uncorrelated, the 95% confidence interval for the j - step ahead forecast is giv en by

$$\left(f^*_{t,t+j} - 1.96\,\sigma\left(\frac{1-\beta^{2j}}{1-\beta^2}\right)^{0.5},\ f^*_{t,t+j} + 1.96\,\sigma\left(\frac{1-\beta^{2j}}{1-\beta^2}\right)^{0.5}\right)$$

Note that the width of the forecast interval is increasing in the forecast steps *j,* that however there exists an upper limit, which corresponds to a multiple of the unconditional standard deviation of the process.

In the above time series model, this confidence interval can be estimated by employing the OLS estimators of the parameters β and σ. Please note that this confidence interval only takes into account the uncertainty due to the bank-idiosyncratic innovations, not accounting for the estimation errors in the coefficients. We neglect the estimation errors concerning the coefficients for two reasons: First, the estimation errors are said to be small relative to the uncertainty due to the innovations and we have reason to believe that this holds especially true for panel regressions.[4] Second, if we treated the estimated coefficients as random variables, the forecast variance would be a function of expectations of random variables raised to the power of four and more, making it hard to determine confidence intervals analytically.

We transfer the result from above to our models as follows. Instead of the *N* banks in our sample, we look at one representative bank, namely the bank whose forecast of the income component under consideration corresponds to the median of the banks in the sample. Please note that we look at the forecast interval of one single (yet hypothetical) bank, not at the forecast interval of the median.[5]

As mentionend above, it is important to note that the resulting confidence intervals can only give a vague sense of the forecast uncertainty due to at least four limitations. First, the estimation error associated with the estimated coefficient is not incorporated in this analysis. Second, the above analytical expression is based on normality of the idiosyncratic shocks.

[4] For a textbook discussion, see Wooldridge (2012), page 660.

[5] The forecast interval of the median (and mean) is very small as bank-idiosyncratic effects are cancelling out in the large cross-section of banks (See Section 3.3).

Third, due to the assumption of a static balance sheet, the bank-specific regressors are treated as known in the future and the uncertainty in their realization is ignored. Fourth, mis-specifications and the risk of future structural breaks are ignored.

2.1.3 In-sample estimation results

We now turn to applying the above procedures to the German banking sector. For the small banks, the income components interest income, interest expense and net fee income are examined by means of the panel regression approach described above. For loan loss provisions, a different approach is applied (cf. section 2.2). The income components are all expressed relative to total assets in percentage terms and are the dependent variables in the regression models. The data that is used for the small banks is based on yearly balance sheet information covering the period 1995 to 2013.

In Appendix 4, the estimation results are presented in detail. Here, we focus on the implied long-run macroeconomic elasticities. These elasticities are computed as the sum of the contemporaneous and lagged macroceonomic coefficients divided by one minus the coefficient of the lagged dependent variable, see Maddala et al. (1997) for details.

	Interest income	**Interest expense**	**Net fee income**
3M EURIBOR	0.19	0.42	
10 J Bundesanleihe	0.51	0.14	
GDP growth			0.01

Table 2: long – run macroeconomic elasticities for small banks

The elasticities reflect changes in bank income due to changes in the macroeconomic environment. For instance, a one percentage point change in the 10-year Bund yield is associated with an increase of interest income (relative to total assets) of about 0.5 percentage points in the long run on average. Similar conclusions apply to the other components. As expected, long-term interest rates have a stronger effect on interest income than short-term interest rates. This relationship reverses for interest expenses. For net fee income, only GDP growth turned out to be significant, whereas interest rates were insignificant.

For the large banks, the loan loss provisions (relative to total loans) are examined in addition to the income categories studied for the small banks. Also, net interest income is examined rather interest income and expense separately. Instead of the yearly data employed for the small banks, quarterly data is used for the 12 largest banks (on a bank holding company level), covering Q1 2008 to Q4 2013. The magnitude of these elasticities is relatively small, implying that these income components are relatively insensitive to changes in the macroeconomic factors considered in the model.

	Net loan loss provisions	Net interest income	Net fee income
3M EURIBOR		-0.004	
10 J Bund.		0.001	
GDP growth	-0.042		0.002

Table 3: long – run macroeconomic elasticities for large banks.

In addition, in appendix 5 the estimation results are reported for an enriched model with additional macroeconomic factors (see appendix 3 for details) For brevity, the results are shown for interest income and interest expense for the small banks, and net interest income for the large banks. It turns out that the results for the small banks are quite robust in the sense that the including the additonal factors does not change the estimates of the observed macroeconomic factors much. The estimates of the lagged dependent variable shrink a little, while the in-sample fit as measured by the adjusted R^2 (in the fixed-effects regressions) improves only slightly. Overall, the more parsimonious model without the additional factors is chosen to perform the stresstest.

With these models, the bank income components are predicted. Appendix 6 presents the forecasted income components for the small banks along with the confidence bands as explained in Subsection 2.1.2.[6] Similarly, in Appendix 7, the resulting operating results for the small and large banks are depicted.

2.1.4 Pseudo out-of-sample forecast exercise

To evaluate the models discussed above in terms of their forecast accuracy, a pseudo out-of-sample forecast exercise is performed. To this end, the sample is divided into an estimation sample and a forecast sample. The estimation sample is used to obtain initial estimates of the model parameters and to produce a one-step ahead forecast of the relevant bank income component. The average squared deviation from the actual observation is then

$$e_{T_e+1} = \frac{1}{N} \sum_i^N \left(y_{i,T_e+1} - \hat{y}_{i,T_e+1}\right)^2 , \qquad (1)$$

where T_e denotes the initial number of time periods in the estimation sample, y_{i,T_e+1} is the true value of the bank income component of bank i in time $T_e + 1$, and $\hat{y}_{i,T_e+1}$ is a suitable forecast. The estimation sample is then extended to incorporate the next time period in the panel, the models are reestimated and another one-step ahead forecast is produced. This

[6] For brevity, only the results for the small banks are shown.

process continues iteratively until the penultimate period in the sample. The time average of these forecast errors is given by

$$e = \frac{1}{T_f} \sum_{j=1}^{T_f} e_{T_e+j} \,, \qquad (2)$$

where T_f is the number of time periods in the forecast sample, such that $T = T_e + T_f$. In the empirical application, we have $T_e = 10, T_f = 9$ for the small banks and $T_e = T_f = 12$ for the large banks. Two forecasts are compared in this exercise. The first one is the forecast generated by the dynamic panel model as explained in Subsection 2.1.2. The second approach uses the most recent observation of the dependent variable as the one-step ahead forecast.

To compare these two forecasts, we report the ratio of the overall forecast error e of the two approaches. Here, a ratio below one indicates that the dynamic panel provides a more accurate forecast than the competing approach. Table 4 presents the results. It turns out that dynamic panel provides more accurate forecasts in all but one case. In future work, the net fee income for the small banks is examined more closely to improve the forecasting performance in this model.

In addition, as a time series of forecast errors becomes available, we can depict the evolution of the forecast errors over time, see Appendix 8. Regarding the interest income and expense of the small banks, it can be seen that the dynamic panel produces more accurate forecast in almost all periods. In particular, the simple forecasting approach performs relatively poorly in the crisis period 2008 and 2009. In case of net fee income for the small banks, the panel model is outperformed by the simple forecasting approach. Turning to the large banks, the patterns become less clear, but the simple approach displays relatively large forecast errors in the second and third quarter in 2012. Overall, we consider the forecasting performance of the dynamic panel model as satisfactory, although more work is needed to relate the

Small banks	**Interest income**	**Interest expense**	**Net fee income**
MSFE ratio	0.87	0.67	1.76
Large Banks	**Loan loss prov.**	**Net interest inc.**	**Net fee income**
MSFE ratio	0.91	0.48	0.85

Table 4: out - of - sample forecast exercise

Note: Entries are the ratios of the overall forecasting error in (2) of the forecast by the dynamic panel model (see Subsection 2.1.2) relative to the forecast of the competing approach in which the most recent observation of the dependent variable is used as the one-step ahead forecast. Thus, a ratio below one indicates that the panel forecast provides a more accurate forecast that the competitor.

proposed approach to alternative methods and to improve the forecast performance of the existing models.

2.2 Sectoral regressions

For small banks, the domestic credit portfolios makes up the bulk of their loans to the real economy; loans to the foreign real economy amount to less than two per cent on average. The Bundesbank's borrowers' statistics offers a detailed breakdown of the loans to the German real economy, where the domestic credit portfolio (at bank level) is broken down into 24 industries and three types of loans to private households (mortgage loans are excluded in our analysis because we deal with them separately, see Section 2.4). To make use of this very detailed data, we proceed as follows: In the first step we estimate the relationship between the aggregate write-down rates for the 27 industries/loan types and the GDP-growth and the unemployment rate. These two factors turned out to have a high explanatory power for the write-down rates. Lags of the explanatory variables are chosen such as to maximize the coefficient of determinaion $R^{2.}$ In the second step, we use predicted values of the write-down rates to obtain forecasts for the credit losses in the individual bank credit portfolios:

$$credit\_losses_{t,i} = \sum_{j=1}^{27} w_{t,i,j} \cdot Q_{t,j}$$

where $w_{t,i,j}$ is the share of industry *j* in bank *i*'s credit portfolio and $Q_{t,j}$ is the predicted write- down rate of industry *j*.

In Appendix 9, we give the estimated sensitivities of GDP-growth and the unemployment rate and the corresponding lags (0 - 4 quarters) which showed the best (in-sample) fit.

2.3 Quantile analysis

As shown above, we determine most of the banks' income components, such as interest income and credit losses, by ways of regression analyses, either by panel regression or by sectoral regression. This approach is applicable if there exists a statistically significant relationship between the macro variables and the income components.

For the trading income, however, this is not the case. We are not able to explain a bank's trading income by, say, interest rates, GDP growth, exchange rates and stock market returns. One reason for this could be that banks, in particular large banks, usually do not follow simple trading strategies, say a buy-and-hold strategy to maximize their trading returns. More complex strategies involve frequent buying and selling of assets as well as the use of derivatives to hedge positions against market movements. However, it would be inappropriate to exclude the trading income from macro stress tests as it makes up a large fraction of total income for the larger banks. To obtain a stress effect on trading income, we assume a monotonic relationship between the stock market return and the returns from trading. We procced as follows: First, we derive the empirical distribution of a bank's standardized trading income where we standardize a bank's quarterly trading income with the respective trading book RWAs, i.e. we determine the empirical distribution of $TI_{i,t}$ /$RWA_{i,t}$ using data of 12 large banks and 20 quarters).[7] Second, concerning the stock market returns, we determine the quantiles of the CDAX returns provided in the macro scenarios relative to its historical distribution.[8]. Third, we assume that the quantiles of the stock return distribution in the different macro-scenarios correspond to the quantiles of the standardized trading income distribution. To obtain a bank's trading income, we multiply the bank's trading book RWA with the respective quantiles in the different macro-scenarios:

$$TI_{t,i} = RWA_{t,i}^{TB} \cdot F_{TI\_RWA}\left(G_{DAX}^{-1}(r_{DAX,t})\right)$$

where $TI_{t,i}$ denote the trading income, $r_{DAX,t}$ the stock market return and $F_{TI\_RWA}(\cdot)$ and $G(\cdot)$ the cumulative distribution functions of the standardized trading income and the stock market return, respectively.

2.4 Real estate

2.4.1 Dataset

The dataset on residential real estate lending is built up from various sources: The core of our dataset is derived from the borrower statistics ("Kreditnehmerstatistik", KNS) which gives data on the volume of outstanding mortgages and the MFI interest rate statistics (MIR) with data on volumes of new mortgage lending. The borrower statistics includes bank-by-bank data for all German banks while the MIR-statistics includes individual bank data for a representative sample of 240 German banks. Data for house prices is available from 2004 to 2014 from the Association of German Pfandbriefbanken (vdp) and bulwiengesa AG. In addition, we make use of a one-off Bundesbank / BaFin survey on mortgage lending to households. The survey has been conducted among 116 banks within 24 selected towns and cities, especially those which have witnessed particularly strong rises in housing prices. It covers the years from 2009 to 2013 and contains data on new and existing mortgage loans, distributions of German sustainable Loan-to-Values ("Beleihungsausläufe"), initial amortization quotas and default probabilities on an annual basis.

As the one-off survey includes disaggregated bank-level data for the 24 selected towns and cities, we decided to build our dataset around three different geographical aggregates. Since house price dynamics might differ between these aggregates we use different house price indicators.

Metropolitan area: The aggregate consists of Berlin, Frankfurt am Main, Munich, Hamburg, Stuttgart, Cologne as well as Düsseldorf. To depict the price dynamic in this aggregate we use data from bulwiengesa AG on house prices in the seven largest German cities.

Urban area, covered by the survey: The aggregate consists of Aachen, Augsburg, Bonn, Bielefeld, Bremen, Dresden, Erlangen, Hannover, Heidelberg, Münster, Leipzig, Essen, Lübeck, Magdeburg, Mannheim, Nuremberg, Wiesbaden. To reflect the price dynamic in this aggre-

[7] See Bundesbank (2013), p.59.

[8] To account for the serially correlated stock market returns in the macro scenarios (in contrast to the serially nearly indepedent empirical stock returns), we use year-to-year stock returns (instead of quarterly returns).

gate we use price data from the bulwiengesa AG for 127 German cities which are not a metropolitan city.

Other Domestic: This aggregate encompasses housing data for all other regions. To depict the price dynamic in this region we use the price indicator provided by the vdp that reflects the broadest measure of housing prices in Germany.

Regarding data cleansing and updating of data gaps, we use the following adjustment algorithm for the metropolitan and urban areas:

1. We only use those survey values where the value is based on at least 50 percent of the total reported mortgage volume within any bank-year combination. Otherwise we use average survey values for the respective geographical region in any given year.
2. For the time period 2004 and 2008, for which we do not have any data from the survey, we assume that the data is identical to the earliest distributions reported under the survey. Furthermore, we assume that the distributions in 2014 are the same as in 2013. The later assumption seems warranted given the slow to non-observable dynamics in the distributional data. While we acknowledge that the assumption of identical distributions before 2008 is somewhat harder to justify given the lack of representative data and the inherently different housing price dynamics before the crisis, non-representative surveys by the Association of German Pfandbriefbanken (vdp) give no hints about large changes in LTV distributions for the respective time period. More importantly, older mortgage vintages contribute significantly less to potential losses in a downturn such that the importance of a precise LTV distribution declines somewhat for older vintages (see also the loss estimation methods in the next section). This is particularly true in an environment of raising house prices which Germany is experiencing since the beginning of the financial crisis.

Given the backdating and updating of the survey data, we use the following adjustment algorithm:

1. Data on default probabilities, sustainable LTV distribution and the redemption quotas comes from the survey only. For banks that participated in the survey we assume that the distributions of these data are the same for the aggregates “Domestic” and “Urban”. For banks that did not participate in the survey we assume that the distributions are equal to the average of all banks within the aggregate “Urban” in a given year.

2. Regarding data on new mortgage volumes we use the survey data directly, if the observation belongs to the aggregate of a metropolitan or survey urban area. Alternatively, if the observation belongs to the aggregate “other domestic” we use the mortgage data from the MIR-statistics less the volumes reported in the survey. Since the MIR-statistics is sample based, for all remaining institutions we approximate new mortgage volumes by using the borrowers statistics and the average relation between new to existing loans based on the MIR-sample.

2.4.2 Residential Real Estate: Loss estimation

For each simulated year *T*, we estimate the amount of residential real estate related provisions ($RREP$) as the product of the outstanding notional of residential mortgages ($notional$) and the time- and bank-specific provision rate ($Provision\ rate$)

$$\boldsymbol{RREP_{j,T} = Provision\ rate_{j,T} * notional_{j,T}}$$

For each bank *j*, the provision rate in year *T* is given by the averaged product of default probability (PD) and the loss-given default (LGD). Both variables are modelled separately for each mortgage vintage *t* and various “Beleihungsauslauf” ($BelA_K$) categories *K*. Moreover, calculations are differentiated by the area of loan origination (*s=1* for loans in seven metropolitan areas, *s=2* for loans in the survey urban aggregate, *s=3* for all other loans). The provision rate at the bank level is calculated as a weighted average using relative outstanding mortgages volumes as analytical weights ($w_j^{t,T,s,K}$):

$$\boldsymbol{Provision\ rate_{j,T} = \sum_{t=2004}^{T} \sum_{s=1}^{3} w_j^{t,T,s,K} * PD_j^{t,T,s,K} * LGD^{t,T,s,K}}$$

The $\boldsymbol{LGD}$ is directly related to the recovery value in the case of default ($\boldsymbol{LGD^{t,T,s,K} = 1 - Rec^{t,T,s,K}}$) which in turn is approximated by

$$\boldsymbol{Rec_j^{t,T,s,K} = min\left\{1\ ;\ \frac{(1 + \Delta p_{t,T,s}) * (1 - \Delta f_t)}{(1 - \Delta B) * \overline{BelA_K} * \left[1 - amortized\ share_j^{t,T}\right]}\right\}}$$

with $\Delta \boldsymbol{p_{t,T,s}}$ being the cumulative percentage price increase between *t* and *T* in area category *s*, $\Delta \boldsymbol{f_t}$ the scenario-dependent price discount on the market value in case of a default, $\boldsymbol{\overline{BelA_K}}$ the average “Beleihungsauslauf” of category *K* (see also Table A13 in Appendix 12) and $\Delta \boldsymbol{B}$ being an estimate of the empirical discount on the market value for the residential real-estate to get the German Mortgage Lending Value (MLV, “Beleihungswert”) of the respective property. Finally, the amortized share is calculated linearly using bank-specific data on initial amortization rates and an assumed 30-year time horizon.

Default probabilities ($\boldsymbol{PD}$s) are based on initial average bank estimates for the full-year 2014 plus a stress add-on which mirrors the cumulative change of the aggregate unemployment $\Delta \boldsymbol{u}$ rate since the end of 2014, albeit with a time lag of one year. The time lag of one year is a reasonable assumption as, most likely, a borrower does not default at the instant of becoming unemployed but when dissaving all her assets. Finally, in order to take into account improved loan servicing over time, default probabilities are set to zero after seven years without a default[9]:

$$\boldsymbol{PD_j^{t,T,s,K} = PD_j^{2014,s,K} + \Delta u^{T-1} \qquad if\ T - t \leq 7}$$
$$\boldsymbol{PD_j^{t,T,s,K} = 0 \qquad else}$$

The estimate of the German MLV discount ($\Delta \boldsymbol{B}$, see Table A14 in Appendix 12) is based on regulatory requirements stemming from the PfandBG and BelWertV as well as discussion with private sector experts. Afterwards, the last free parameter ($\Delta \boldsymbol{f_t} = \boldsymbol{0.10}$ in baseline scenario, see Table A14 in Appendix 12) is calibrated in order to match the residential real es-

[9] Based on informal information exchange with real-estate experts from the Association of German Pfandbriefbanken (vdp).

tate related provision rate (0.04% of the outstanding notional in 2014) at the aggregate level reflecting the currently strong demand for German residential real estate. The higher discount rate in the adverse scenario ($\Delta \boldsymbol{f_t} = \mathbf{0.40}$ in adverse scenario) is motivated by the presumed strong decline in housing demand and is in line with past banking supervision experiences.

3. First-Round Effects

3.1 Default algorithm

Given the forecasts of banks' earnings components as described in the previous chapter, the net operating result can be determined. To this end, the operating costs and the other operating result need to be taken into account. As these are relatively stable over time when considered as a fraction of total assets, we compute the median operating cost (relative to total assets) between 2008 and 2013 and set the operating cost for the forecast horizon equal to this value in each period. We adopt the same approach for the other operating result. With these choices, the net operating profit can be determined for each period over the forecast horizon:

$$\hat{P}_{js} = \widehat{NI}_{js} + \widehat{NF}_{js} + \widehat{NLP}_{js} + \widehat{NLP}_{js}^{\ H} + \widehat{NT}_{js} + \widehat{NOR}_{js} - \hat{C}_{js} \,,$$

where $j = 1,2,\dots,N$ and $s = T+1, T+2, \dots, T+H$, with $\hat{P}_{jt}$ being the predicted net operating profit for bank j at time s, and $\widehat{NI}$, $\widehat{NF}, \widehat{NLP}$, $\widehat{NT}, \widehat{NLP}^H$, $\widehat{NOR}$ and $\hat{C}$ are the predictions of the net interest income, the net fee income (see Section 2.1, respectively), the net loan loss provisions for the credit portfolio (excluding real estate, see Section 2.2), the net trading income (see Section 2.3), the net loan loss provisions for the real estate credit portfolio (see Section 2.4), the (net) other operating result, and the operating costs, respectively, where the subscripts are left out for simplicity. Note that for the small banks, the net trading income is assumed to be zero in all periods over the forecast horizon.

For a given stock of tier 1 capital in the last sample period, the predicted tier 1 capital in the first forecast period can be computed. To this end, we distinguish two cases. In the first case, it is assumed that a positive operating profit is distributed fully among the owners of the bank. Thus, $\hat{e}_{j,T+1} = e_{jT}$ if $\hat{P}_{j,T+1}$ is positive, where e_{jT} is the observed tier 1 capital in the last sample period T for bank j, and $\hat{e}_{j,T+1}$ is the predicted tier 1 capital in the following period. If the operating profit is negative, however, the corresponding tier 1 capital is reduced by this amount due to the loss absorbing function of equity. Analogously, tier 1 capital is determined iteratively for the remaining periods in the forecasting period. In the second case, any operating profit in the first forecasting period is fully retained and so $\hat{e}_{j,T+1} = e_{jT} + 0.7\hat{P}_{j,T+1}$, if $\hat{P}_{j,T+1}$ is positive, where a tax rate of 30% is assumed. A negative profit in the first forecasting period reduces tier 1 capital as before. This process continues iteratively until the last forecasting period is reached. Each of these two schemes produces a series of predicted tier 1 capital.

Next, in each case, the ratio of the predicted tier 1 ratio is determined as the predicted tier 1 capital relative to the predicted risk weighted assets. The calculation of risk weighted assets (predicted values) is discussed in the following section. Starting with the first forecast period, if the tier 1 ratio of a given bank drops below the regulatory lower bound of 6%, the bank defaults and is removed from the sample. In the next period, the remaining banks are considered and are assumed to default if the current tier 1 ratio falls below the 6% threshold. This process continues iteratively.

This default algorithm is thus relatively straightforward. It should be noted, however, that this approach makes some simplifying assumptions. In particular, the banks' balance sheets are assumed to be static, i.e. the bank management cannot react to the stress by, say, deleveraging.

3.2 RWA-adjustment

The economic environment has not only an impact on the banks' profits and losses and thereby on their capital. The risk weighted assets (RWA) are also affected. For the banks that use internal rating based models (so-called IRBA-banks) for their credit risk, we adjust the risk weighted assets (RWA) as follows:

$$RWA_{t,i} = RWA_{0,i}^{credit} \cdot \frac{RW(pd_t)}{RW(pd_0)} + RWA_{0,i}^{other}$$

where pd_t is the cross-sectional median probability of default for customer loans in quarter t, estimated as the loss rate in the banks' credit portfolios, predicteded by the models described in the Sections 2.1 and 2.2, devided by the loss given default (LGD). The LGD is assumed to be fix and equal for each credit portfolio amounting 45%. $RW(\cdot)$ is the function for the risk weights according to Basle II given in Appendix 10. The RWA are divided into RWA for credit risk and other RWA. The adjustment is made only for the RWA for credit risk.

3.3 Dealing with banks' idiosyncratic risk

The models that link the macroeconomic stress to the banks' balance sheets often have only limited explanatory power as measured by the coefficient of determination (R^2). This is especially relevant for models regarding credit risk.[10] The remaining unexplained part is due to idiosyncratic noise. When aggregating the banks' results in a linear manner, for instance when calculating arithmetic averages, this limited explanatory power is not problematic, because the banks' idiosyncratic noise in the cross section of the banks cancels out. By contrast, aggregating numbers in a non-linear way is more problematic. This is the case, for instance, when the number of defaults or the aggregate capital shortfall is to be determined.

[10] Memmel et al. (2015) find for the write-down rate of credit portfolios of German banks a coefficient of determination of less than 10%, whereas the coefficient of determination for the net interest income is found to be above 40% by Memmel and Schertler (2013).

This problem is resolved by adding noise to the results, so that the modelled variation of the bank results corresponds to the actual variation. This noise does not add further information to the result of a single bank, but it makes the aggregation more meaningful as the example shows below.

We assume the following relationship between the macroeconomic variable x and the earnings y_i of bank *i*

$$y_i = \alpha + \beta \cdot x + \varepsilon_i \tag{3}$$

Often, the bank's earnings $y_i^{st,m}$ under stress are modeled as follows (index "m" for model):[11]

$$y_i^{st,m} = \alpha + \beta \cdot x^{st} \tag{4}$$

where x^{st} is the value of the macroeconomic variable in case of stress. The actual value (index "act") for the earnings of this bank, however, amounts to

$$y_i^{st,act} = \alpha + \beta \cdot x^{st} + \varepsilon_i$$

Let R^2 be the coeffienct of determination in (3), let *N* be the number of banks and σ_y^2 be the variance of the earnings, then the variance of the difference between the average earnings in the model and the average of the actual earnings in case of stress is

$$\operatorname{var}\left(\frac{\sum y_i^{st,m}}{N} - \frac{\sum y_i^{st,act}}{N}\right) = \frac{\sigma_y^2\left(1-R^2\right)}{N} .$$

If the number of banks is large, the difference between the average earnings from the model and the average actual earnings vanishes, even if the explanatory power is small,

$$\overline{y}^{st,m} \xrightarrow{p} \overline{y}^{st,act}$$

If instead one looks at a non-linear function of the earnings, for instance the default of a bank, then this difference between the averages is not vanishing even if the number of banks is large. A default of bank *i* is given if its earnings are below a certain threshold:

$$a_i = \begin{cases} 1 & \text{in case } y_i < c \\ 0 & \text{otherwise} \end{cases}$$

For reasons of simplification, we assume that x and ε_i are jointly normally distributed (with mean zero). Under this assumption, we obtain for the probability of a default of bank *i*

$$\Pr\left(y_i < c\right) \equiv E\left(a_i\right) = \Phi\left(\frac{c-\alpha}{\sigma_y}\right) ,$$

$$E\left(a_i^{st,act}\right) := E\left(a_i^{act} \mid x = x^{st}\right) = \Phi\left(\frac{c-\alpha-\beta \cdot x^{st}}{\sigma_y \cdot \sqrt{1-R^2}}\right)$$

and

$$E\left(a_i^{st,m}\right) = 0 \quad \text{(or – in case of severe stress – 1)}$$

Accordingly, after the realisation of the macro variable $x = x^{st}$:

[11] See Aikman (2009).

$$\overline{a}^{st,act} \xrightarrow{p} \Phi\left(\frac{c-\alpha-\beta\cdot x^{st}}{\sigma_y\cdot\sqrt{1-R^2}}\right)$$

and

$$\overline{a}^{st,m} \xrightarrow{p} 0 \quad \text{(or – in case of severe stress – 1)}$$

The solution above suggests that (4) is replaced by the following advancement (index: “m1”):

$$y_i^{st,m1} = \alpha + \beta\cdot x^{st} - \nu_i \tag{5}$$

where ν_i is a noise term with variance $(1-R^2)\cdot\sigma_y^2$. It can be seen as the bank-specific part of the write-down rate of the credit portfolio. Applying Equation (5), the model averages of the earnings and the defaults converge to the actual values.

The assumption of normality above is mode for the ease of exposition only. In reality, the empirical data of credit write-down rates can be better described by the exponential distribution. This holds especially true for the extreme parts (tails) of the distribution. In the following table the quantiles of the yearly write-down rate of large banks in Germany are displayed, and in addition, the corresponding values of the exponential distribution with the same standard deviation as the empirical data.

Quantiles of the write-downrate	Values from the empirical distribution	Values from the exponential distribution
90%	1.89%	2.30%
95%	2.84%	2.99%
97%	3.67%	3.50%
99%	4.75%	4.60%

Table 5: Quantiles of the banks’ write-down rates in the credit portfolio

Accordingly, the following specification seems reasonable:

$$\nu_i + \frac{1}{\lambda} := \varepsilon_i \sim E(\lambda) \tag{6}$$

with

$$\lambda = \frac{1}{\sigma_y\cdot\sqrt{1-R^2}}$$ [12]

In the following we describe how the capital shortfall (gap) is determined when ideosyncratic risk is accounted for. We start with the following notation:

EK_i^v : Regulatory capital of bank *i* before the stress test

ΔEK_i : Change in the bank’s capital that is due to the systematic part of the stress

[12] $\lambda = 116.40$ is calibrated so that the standard deviation of ν_i equals the estimated standard deviation of the loss rate in the banks’ credit portfolios (0.99892%) times the coefficient of determination of nationwide banks concerning credit risk (26.04%) in Memmel et al. (2015).

RWA_i^n : Risk weighted assets (after the stress)

F_i : Customers loans

c : Regulatory minimum capital ratio

ν_i : Idiosyncratic write-down rate on custumor loans (see Equation (6)).

A_i : Indicator variable that takes one the value of one in event that the regulatory captial requirements are not fullfilled.

$$A_i = \begin{cases} 1 \quad in\,case\; EK_i^v + \Delta EK_i - c \cdot RWA_i^n < \nu_i \cdot F_i \\ \quad 0 \quad otherwise \end{cases}$$

GAP_i : Capital gap, i.e. the euro amount that is necessay so that the bank attains the minum capital requirements.

Then it is easy to show the following

$$GAP_i = \begin{cases} c \cdot RWA_i^n - EK_i^v - \Delta EK_i + \nu_i \cdot F_i \quad in\,case\; A_i = 1 \\ \quad 0 \quad otherwise \end{cases} \tag{7}$$

Then one can show by virtue of Equation (8) in Appendix 1:

$$\Pr\left(A_i = 1\right) = e^{-\lambda u_i}$$

with

$$u_i = \max\left(0, \frac{EK_i^n + \Delta EK_i - c \cdot RWA_i^n}{F_i} + \frac{1}{\lambda}\right)$$

Finally, after some manipulations of Equation (7) in combination with Equations (9) and (10) in Appendix 1, we derive:

$$E\left(GAP_i\right) = e^{-\lambda u_i} \cdot \left(c \cdot RWA_i^n - EK_i^v - \Delta EK_i + F_i \cdot u_i\right)$$

4. Second-round effects

In the previous chapters we described how to derive estimates for the direct effects of a change in macroeconomic conditions on the banks' profit and losses and balance sheet positions. Second-round effects occur if a failure of one bank leads to losses at other "connected" banks, for instance at banks that have direct exposures to the failed banks. The literature on second round effects due to direct mutual exposures distinguishes between two approaches in investigating these effects:

1. The cascade algorithm: The cascade or round-by-round algorithm works as follows: one or more banks default for an exogeneous reason, for instance due to losses in the first round of a macro stress test. Banks that have credit exposure to these banks suffer losses and their equity is reduced accordingly. If the losses exceed their equity

or their capital ratio falls below a certain threshold, these banks also default, thereby transmitting the original shock. The process continues until no new bank defaults occur. Recently, Fink et al. (2014) refined this algorithm taking into account the effects from the banking regulation according to Basle II.

2. Endogenous loss distribution (See Eisenberg and Noe, (2001)): Also starting with one or more exogenous defaults, this algorithm distributes the losses among the banks in the interbank market such that a new equilibrium exists. In contrast to the cascade algorithm, the amount of necessary write-downs is endogenously determined.

In our study, we apply the cascade algorithm, the so-called round-by-round algorithm, mainly for empirical reasons. Namely, one observes that the distribution of the loss given default (LGD) for interbank exposure is markedly u-shaped (much probability mass for small and large losses given default, but litte mass for medium LGDs) and only loosely dependent on bank characteristics (See Memmel et al., (2012)). This suggests modelling the loss given default (LGD) of interbank exposures as a beta-distributed random variable,[13] which can be more easily done in the cascade algorithm than in the other algorithm because in the other algorithm the LGD is endogenously determined whereas the cascade algorithm allows for exogenous LGDs.

We proceed as follows: Using data from the German credit register, we obtain the bilateral exposures among the banks in Germany. We restrict the analysis to the large banks, because the smaller banks tend to be much less connected via the interbank market (Savings banks and credit cooperatives are mainly connected to their respective central institutions). First, we determine the direct effects on the banks' profit and losses as well as its RWA and tier 1 capital. Then we add a noise term as outlined in Section 3.3. For all failed banks, i.e. banks with tier 1 capital ratios below 6 %, we determine the losses of connected banks. Here we assume a random LGD drawn from a beta distribution. The noise term and the LGD are drawn from independent random variables for each bank and each bank-to-bank relationship, respectively. This process is repeated until no further bank fails. The whole algorithm is repeated in 100,000 simulation runs. If the combined stress effect of the first-round effect and the bank-specific effect leads to banks with capital ratio below the threshold, the cascade algorithm starts, possibly leading to additional defaults and reductions in the capital ratios. The additional losses cannot be attributed to single banks, but only to the system as a whole, whereas the first-round effects can be attributed to the individual banks.

[13] This is done in Memmel et al. (2012) where we also take the parameters of the beta-distribution from (See Appendix 11).

Appendix 1: Exponential distribution

Let the random variable ε be exponentially distributed with parameter λ, then its density and cumulative density functions are $\left(x \geq 0\right)$:

$$f_{\varepsilon}(x) = \lambda \cdot e^{-\lambda x}$$

and

$$F_{\varepsilon}(x) = 1 - e^{-\lambda x} \tag{8}$$

.

The expectaion and the variance are $1/\lambda$ und $1/\lambda^2$, respectively, what can be derived from the following equations:

$$\int x\lambda \cdot e^{-\lambda x} dx = -\left(\lambda x + 1\right) \cdot e^{-\lambda x} \cdot \frac{1}{\lambda} + C \tag{9}$$

$$\int x^2 \lambda \cdot e^{-\lambda x} dx = -\left(\lambda^2 x^2 + 2\lambda x + 2\right) \cdot e^{-\lambda x} \cdot \frac{1}{\lambda^2} + C \tag{10}$$

Appendix 2: Data

Here, we give a very brief overview about the data sources that we use.

Source	Variable(s)
EGV and Sonderdatenkatalog	*Interest income and expense, net fee income*
Deutsche Bundesbank	*3M EURIBOR, 10-year Bund, GDP growth*
EJB	*Total assets, Loans, book equity*
Sonderdatenkatalog	*Loan loss provisions, Tier 1 captial*
E UEB	*Risk-weighted assets (credit)*

Table A1: Data sources for the small banks

Source	Variable(s)
Bankscope	*Total loans*
Deutsche Bundesbank	*3M EURIBOR, 10-year Bund*
Destatis	*GDP growth*
Deutsche Bundesbank (Statistics)	*Total assets, net interest income, net fee income, loan loss provisions, trade result, book equity, Tier 1 capital*

Table A2: Data sources for the large banks

Appendix 3: Details on GMM estimation in dynamic panel models

In the fixed effects model introduced in Section 2.1, the bank-specific effect is allowed to be correlated with some or all of the regressors included in the model.[14] By design, in the dynamic version of the fixed effects model, the lagged dependent variable y_{it-1} is correlated with the disturbance term u_{it} , as both variables depend on the bank-specific effect μ_i. Correlation between the regressors (the lagged dependent variable and possibly other explanatory variables in the model) invalidates the ordinary least squares (OLS) estimator, as consistency no longer holds. The standard alternative procedure is the fixed effects (FE) estimator. Here, all variables are considered in deviation from their time averages,

$$\tilde{y}_{it} = y_{it} - \frac{1}{T-1}\sum_{s=2}^{T} y_{is}$$

This FE estimation eliminates the bank – specific effect and thus allows estimation of the parameters in the model. The FE estimator results as the OLS estimator in a regression model in which all variables are transformed in the above manner.

It turns out, however, that the FE estimator in the dynamic model is biased, even in large samples.[15] Here, the theoretical framework assumes the time dimension to be fixed (and small), while the cross-section dimension is large (say $N = 1000, T = 5$). The bias of the FE estimator would be negligible if the time dimension was also large (say $T = 100$). In many panel datasets, however, the number of time periods is limited, and the FE estimator is inconsistent.

An instrumental variables estimator, also referred to as a GMM estimator, provides a consistent alternative in this model. The GMM estimator applies to the model in which all variables are considered in first difference (relative to the previous period):

$$\Delta y_{it} = \phi \Delta y_{it-1} + \beta' \Delta x_{it} + \gamma' \Delta z_t + \Delta \epsilon_{it} ,$$

where now $\Delta y_{it} = y_{it} - y_{it-1}$ and the remaining variables are defined analogously. Note that the bank – specific fixed effect has been eliminated from the model. The above model is referred to as the *transformed model* in the following discussion.

In the transformed model, the lagged dependent variable in first difference $\Delta y_{it-1} = y_{it-1} - y_{it-2}$ is correlated with the transformed idiosyncratic error $\Delta\epsilon_{it} = \epsilon_{it} - \epsilon_{it-1}$, so OLS in the transformed model is not consistent. By using suitable instruments which are correlated with the lagged dependent variable in first difference but are not correlated with the transformed error term, a GMM estimation procedure applies.

[14] In contrast, in the random effects model, the bank – specific effect would be assumed to be independently distributed from regressors in the model.

[15] See Nickell (1981).

To this end, the following assumption is made:

$$E\ [\ \epsilon_{it} \mid y_i^{t-1}, x_i^T, z^T\ , \mu_i\ ] = 0\ .$$

Here, $y_i^{t-1} = (y_{i1}, y_{i2}, \ldots, y_{it-1})'$, $x_i^T = (x_{i1}, x_{i2}, \ldots, x_{it-1}, x_{it}, x_{it+1}, \ldots, x_{iT})'$, and $z^T = (z_1, z_2, \ldots, z_T)'$, and $E[\cdot \mid \cdot]$ is the conditional expecation. To interpret this assumption, notice that

$$E[y_{it} \mid y_i^{t-1}, x_i^T, z^T, \mu_i] = \alpha + \phi\, y_{it-1} + \beta' x_{it} + \gamma' z_t + \mu_i$$

.

Hence the assumption implies that the dynamic model is correctly specified in the sense that after taking the dynamics of the income process, the (observed and unobserved) bank – specific and macroeconomic factors into account, no systematic error is made in describing the average (or expected) income component of the bank. In addtion, the assumption also says that the idiosyncratic error displays no serial correlation, as $E[\epsilon_{it}\epsilon_{it-s}] = E[\epsilon_{it-s}\, E\ [\ \epsilon_{it} \mid y_i^{t-1}, x_i^T, z^T\ , \mu_i\ ]] = 0$ for $s \geq 1$. Note that the regressors x_{it} and z_t are assumed to be strictly exogenous. Hence feedback from the dependent variable (a component of a bank's income) to the regressors is excluded. The absence of feedback effects is plausible for the macroeconomic regressors, as a single bank's income does not affect future macroeconomic variables (such as GDP growth or the 10 – year interest rates) . For the bank-specific regressors x_{it} , however, strict exogeneity could be considered as a strong assumption. Strict exogeneity can be relaxed by assuming that bank -specific regressors are predetermined,

$$E\ [\ \epsilon_{it} \mid y_i^{t-1}, x_i^t, z^T\ , \mu_i\ ] = 0\ .$$

In this case current idiosyncratic errors are uncorrelated with contemporaneous and past bank-specific regressors, $E[\epsilon_{it} x_{is}] = E[x_{is}\, E\ [\ \epsilon_{it} \mid y_i^{t-1}, x_i^t, z^T\ , \mu_i\ ]] = 0$ for $s \leq t$. Feedback from shocks to bank income to future bank-specific regressors is allowed, however. In the following analysis, the assumption of strict exogeneity is made. In a supplementary analysis (not shown), the estimated coefficients of the macroeconomic factors are found to be fairly robust with respect to this assumption, and estimation under strict exogeneity is computationally less costly. Hence, the simplifying assumption of strict exogeneity is considered to be reasonable.

In this setup, suitable instruments can be chosen to perform estimation in the transformed model. For instance,

$$E\ [y_{it-2}\ \Delta\epsilon_{it}] = E[y_{it-2}\ \epsilon_{it}] - E[y_{it-2}\ \epsilon_{it-1}]\ ,$$

such that under strict exogeneity,

$$E[y_{it-2}\,\epsilon_{it}] = \; E\big[E[y_{it-2}\,\epsilon_{it}|y_i^{t-1}\,, x_i^T, z^T, \mu_i]\,\big] = \; E\big[y_{it-2}\,E[\epsilon_{it}|y_i^{t-1}\,, x_i^T, z^T, \mu_i]\,\big] = \; 0$$

Analogously, we obtain $[y_{it-2}\,\epsilon_{it-1}] = 0$, such that there is no correlation between y_{it-2} and $\Delta\epsilon_{it}$, while y_{it-2} and $\Delta y_{it-1} = y_{it-1} - y_{it-2}$ are correlated by construction. Hence past levels of the dependent variable can be employed to instrument the lagged dependent variable in first difference in the transformed model. We refer to the set of equations

$$E[y_{it-s}\Delta\epsilon_{it}\,] = 0\,, \qquad t \;=\; 3,4\,,\dots.T;\;\; s = 2,3,\dots,t-1$$

as *moment conditions*. Note here that eliminating the fixed effect by taking first differences instead of deviations from means as in the fixed effects transformation is crucial. The fixed effects transformation yields an error term that depends on past (and future) errors (through the deviation from the time average), which invalidates an instrumental variable approach using past levels of the dependent variable.

The GMM estimator can be formulated explicitly by adopting matrix notation. Let

$$y_i = \phi y_{i,(-1)} + X_i\beta + Z\gamma + \; u_i$$

with $y_i = (y_{i2},\dots,y_{iT})'$, $y_{i,(-1)} = (y_{i1},\dots,y_{iT-1})'$, $X_i = [x_{i1},\dots,x_{iT}]'$ (a $T \times K_1$ matrix in which K_1 is the number of bank – specific explanatory variables included in the model) and $Z = [z_1,\dots,z_T]'$ (a $T \times K_2$ matrix, which lists the observations of the K_2 macroeconomic factors). The whole panel can be written compactly as

$$y = W\delta + u.$$

Here $y = (y_1,\dots,y_N)'$, $u = (u_1,\dots,u_N)'$ and $W = (W_1',\dots,W_N'\;)'$ in which $W_i = [y_{i,(-1)},\; X_i,\; Z]$. The transformed model can be written as

$$Dy = DW\delta + D\epsilon,$$

where $D = I_N \otimes D_T$ and D_T is the $(T-2)\times T$ matrix, which selects the appropriate first differences,

$$D_T = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & -1 & 1 \end{pmatrix}.$$

For a given matrix of instruments, denoted Π, the GMM estimator results as

$$\hat{\delta} = \left(W'D'\Pi\,\widehat{\Omega}^{-1}\,\Pi'DW\right)^{-1}\left(W'D'\Pi\,\widehat{\Omega}^{-1}\Pi'Dy\right),$$

in which $\widehat{\Omega}$ is a suitable estimator of the variance- covariance matrix of the transformed disturbance term $\Pi'D\epsilon$.

Types of instruments

We now discuss the selection of the instruments for the GMM estimator. To this end, it is useful to arrange the moment conditions also in matrix notation. Let the $(T-2)\times(T-1)(T-2)/2$ matrix Π_{1i} be given by

$$\Pi_{1i} = \begin{pmatrix} y_{i1} & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & y_{i1} & y_{i2} & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 0 & y_{i1} & y_{i2} & y_{i3} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & y_{i1} & y_{i2} & \cdots & y_{iT-2} \end{pmatrix}$$

für $i = 1, 2, \dots, N$ and $\Pi_1 = (\Pi'_{11}, \Pi'_{12}, \dots, \Pi'_{1N})'$. The matrix Π_{1i} makes it possible to formulate the whole set of $(T-2)(T-1)/2$ moment conditions in a compact fashion, such that

$$E[\Pi_{1i}'\,(D_T\epsilon_i)] = 0,$$

for each unit in the cross section. Arellano and Bond (1991) proposed this approach to obtain the asymptotically efficient GMM estimator. The estimator is efficient in the sense that *all* available moment conditions are exploited in the estimation procedure.

It turns out, however, that the GMM estimator involves a bias-variance trade-off. The more moment conditions are used to estimate the underlying parameters of the model, the more precise these estimates will be. As the number of moment conditions increases relative to the cross-section dimension (N), however, the GMM estimator is biased toward the FE estimator, see Roodman (2009). Hence the robustness of the procedure is checked by varying the number of instruments that are employed.

Therefore, we consider two alternative ways to select the instrument matrix Π_i. In the first case, only the latest available lag of the level of the dependent variable is used as an instrument. These $(T-2)$ conditions give rise to the $(T-2)\times(T-2)$ matrix Π_{2i},

$$\Pi_{2i} = \begin{pmatrix} y_{i1} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & y_{i2} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & y_{i3} & 0 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & y_{iT-2} \end{pmatrix}$$

The whole set of this instrument selection is denoted as $\Pi_2 = (\Pi'_{21}, \Pi'_{22}, \ldots, \Pi'_{2\mathrm{N}})'$. Moreover, a linear combination (in this case a simple sum) of the moment conditions can be used by means of the $(T-2) \times (T-2)$ matrix

$$\Pi_{3i} = \begin{pmatrix} y_{i1} & 0 & 0 & \cdots & \cdots & 0 \\ y_{i2} & y_{i1} & 0 & \cdots & \cdots & 0 \\ y_{i3} & y_{i2} & y_{i1} & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ y_{iT-2} & y_{iT-3} & y_{iT-4} & \cdots & \cdots & y_{i1} \end{pmatrix}$$

which also entails $(T-2)$ conditions. Again, the set of these instruments is denoted as $\Pi_3 = (\Pi'_{31}, \Pi'_{32}, \ldots, \Pi'_{3\mathrm{N}})'$. In contrast to Π_2 , here the number of effectively used conditions is reduced to $(T-2)$ without ignoring any moment conditions. In either case, $E[\Pi_{2\mathrm{i}}{}' (D_T \epsilon_i)] = E[\Pi_{3\mathrm{i}}{}' (D_T \epsilon_i)] = 0$.

The relevant estimators which result under the different designs of the instrument matrices Π_1, Π_2 and Π_3 , are referred to as *GMM1*, *GMM2* and *GMM3*.

One step versus two step GMM

The above GMM estimator

$$\hat{\delta} = \left(W'D'\Pi\,\widehat{\Omega}^{-1}\,\Pi'DW\right)^{-1}\left(W'D'\Pi\,\widehat{\Omega}^{-1}\Pi'Dy\right)$$

requires a suitable estimator $\widehat{\Omega}$ of the variance-covariance matrix of the transformed disturbance term. Following Roodman (2009) this estimator can be formulated as $(\Pi' H \Pi)$, where H is an estimator of the variance-covariance matrix of the error $D\epsilon$ and Π is a matrix of instruments as described above. If $= I$, where I is the $N(T-2) \times N(T-2)$ identity matrix, the disturbance is assumed to be homoskedastic. The resulting consistent one-step GMM estimator is denoted as $\hat{\delta}_1$. The one-step estimator can be used to construct an estimator of the variance-covariance marix of $D\epsilon$ that allows for more complex and empirically relevant forms of the matrix, including the case of heteroskedasticity. This estimator is denoted as $H(\hat{\delta}_1)$, such that the efficient two-step GMM estimator can be given explicitly as

$$\hat{\delta}_2 = \left(W'D'\Pi\left(\Pi' H(\hat{\delta}_1)\Pi\right)^{-1}\Pi'DW\right)^{-1}\left(W'D'\Pi\left(\Pi' H(\hat{\delta}_1)\Pi\right)^{-1}\Pi'Dy\right)$$

Additional efficiency gains can be obtained by iterating this procedure.

As Arellano and Bond (1991) pointed out, the standard errors implied by the two-step estimator can be biased. Windmeijer (2005) devised a small-sample correction of the two-step standard errors, which is used in the following analysis.[16]

Time-specific effects

In the above panel model, unobserved time (fixed) effects can be incorporated in addition to the unobserved bank-specific effects μ_i :

$$u_{it} = \mu_i + \lambda_t + \epsilon_{it} .$$

Here λ_t represents all economic factors that affect all banks at a given point in time, but are not observed directly. In panels with a sufficiently small time dimension, such time fixed effects can be included in the analysis by adding dummy variables for the respective time periods in the sample to the list of regressors.

It turns out, however, that in our specific application, including time-specific fixed effects *and* macroeconomic variables makes estimation of macroeconomic elasticities as explained above impossible. The reason is that the parameters associated with the macroeconimic factors z_t are not separately identifiable from parameters of the time indicator variables. More generally, this observation holds true for all variables that do not vary across cross – section units. Since obtaining macroeconomic elasticities explicitly is fundamental for our stresstesting framework, we refrain from including time fixed effects in the model. In future work, we could extend the current methodology to allow for time fixed effects in the model, possibly analogous to Kripfganz and Schwarz (2013).

Additional macroeconomic factors

In the empirical application, a range of macroeconomic factors is experimented with, including short – term and long – term interest rates (3M EURIBOR and 10 year Bundesanleihe, respectively), GDP growth and the unemployment rate. The relevant factors selected by hypothesis tests to arrive at a parsimonious specification. This small set of observed macroeconomic factors as described in z_t may not fully capture all relevant macroeconomic effects on bank income, however. To robustify the analysis, additional macroeconomic factors are extraced from a large panel of macroeconomic time series ranging from Q1 1991 - Q4 2013. These factors are selected and estimated in a static factor model, see Bai and Ng (2002), for example. The number of factors chosen by the information criterion ICP2 is four. These factors are included in the empirical analysis to obtain the model

$$y_{it} = \alpha + \phi y_{it-1} + \beta' x_{it} + \gamma' z_t + \delta' \hat{f}_t + u_{it},$$

[16] See Roodman (2009), page 97: „ … the reported standard errors, with his [NB: Windmeijer's] correction, are quite accurate, so that two – step estimation with corrected standard errors seems modestly superior to cluster – robust one – step estimation".

where $\hat{f}_t$ denote the estimated factors from the factor model estimated by the principal component estimator, see Bai and Ng (2002). See Subsection 2.1.3 for further discussion.

Appendix 4: Results of the panel regressions

The following tables present the estimation results for the dynamic panel data models. In each table, the FE and relevant GMM estimators are shown. The preferred model is highlighted and is the efficient GMM estimator (GMM1) for the small banks. In all cases, the estimated macroeconomic coefficients are fairly robust with respect to the number of instruments (see GMM2 and GMM3), however. For the large banks, the relatively small cross-section dimension requires more experimentation with the number of instruments. Here, smaller numbers of instruments are selected to avoid overfitting bias. In case of net interest income for the large banks, there is little evidence for a dynamic model and the static FE specification is preferred. In the specifications presented here, the test of Arellano and Bond (1991) for serial correlation at lags one to three is performed and does not reject the absence of serial correlation at lags two and three. The results are not shown for brevity.

	FE	**GMM1**	GMM2	GMM3
Lagged interest income (in % of total assets)	0.261	**0.338**	0.721	0.503
	(0.043)	**(0.048)**	(0.051)	(0.062)
	***	***	***	***
3M EURIBOR (in %)	0.144	**0.127**	0.118	0.137
	(0.008)	**(0.005)**	(0.008)	(0.009)
	***	***	***	***
10 – year Bundesanleihe (in %)	0.100	**0.071**	0.042	0.054
	(0.012)	**(0.010)**	(0.012)	(0.012)
	***	***	***	***
10 – year Bund. 1. Lag	0.136	**0.151**	0.067	0.116
	(0.015)	**(0.012)**	(0.018)	(0.016)
	***	***	***	***
10 – year Bund. 2. lag	0.170	**0.112**	-0.016	0.053
	(0.020)	**(0.016)**	(0.020)	(0.024)
	***	***		**
Book Equity/total assets (in %)	0.042	**0.029**	0.006	0.009
	(0.009)	**(0.008)**	(0.011)	(0.009)
		***		
Total credit/total assets (in %)	0.017	**0.028**	0.018	0.020
	(0.003)	**(0.002)**	(0.003)	(0.002)
	***	***	***	***
Constant	0.302			
	(0.154)			
	*			
Number of observations	33,754	29,334	29,334	29,334
Number of banks	4,430	3,942	3,942	3,942
Number of instruments		158	22	23
Adj. R^2	0.55			

Table A3: estimation results for small banks: interest income (in % of total assets)

*Note: GMM1 - GMM3 are the GMM estimators as in Arellano und Bond (1991) using three sets of instruments for the model in first difference. GMM1 makes use of all available moment conditions. GMM2 uses only a subset of the moment conditions (first lag) , while GMM3 constructs a linear combination of the moment conditions (see Roodman (2009)). The preferred model is highlighted, which is the efficient GMM estimator as suggested by Arellano und Bond(1991). */**/*** indicates statistical significance of the estimated parameters at the 10%/5%/1%-level.*

	FE	**GMM1**	GMM2	GMM3
Lagged interest expense (in % of total assets)	0.438	**0.533**	0.568	0.631
	(0.038)	**(0.026)**	(0.029)	(0.027)
	***	***	***	***
3M EURIBOR (in %)	0.219	**0.223**	0.225	0.237
	(0.005)	**(0.005)**	(0.005)	(0.005)
	***	***	***	***
Lagged 3M EURIBOR	-0.036	**-0.027**	-0.028	-0.046
	(0.006)	**(0.005)**	(0.006)	(0.006)
	***	***	***	***
10 – year Bundesanleihe (in %)	-0.026	**-0.065**	0.066	-0.091
	(0.010)	**(0.006)**	(0.007)	(0.007)
	***	***	***	***
Lagged 10 – year Bundesanleihe	0.174	**0.130**	0.116	0.126
	(0.011)	**(0.007)**	(0.008)	(0.007)
	***	***	***	**
Book equity/ total assets (in %)	0.010	**0.018**	-0.001	0.006
	(0.004)	**(0.009)**	(0.012)	(0.009)
		*		
Total credit/total assets (in %)	0.003	**0.012**	0.004	0.007
	(0.002)	**(0.004)**	(0.005)	(0.004)
		***		*
Funding gap/ total assets (in %)	0.001	**-0.000**	0.005	0.005
	(0.002)	**(0.003)**	(0.003)	(0.003)
				*
Constant	0.010			
	(0.155)			
Number of observations	38,555	33,754	33,754	33,754
Number of banks	4,791	4,430	4,430	4,430
Number of instruments		160	24	24
Adj. R^2	0.67			

Table A4: estimation results small banks: interest expense (in % of total assets)

*Note: GMM1 - GMM3 are the GMM estimators as in Arellano und Bond (1991) using three sets of instruments for the model in first difference. GMM1 makes use of all available moment conditions. GMM2 uses only a subset of the moment conditions (first lag) , while GMM3 constructs a linear combination of the moment conditions (see Roodman (2009)). The preferred model is highlighted, which is the efficient GMM estimator as suggested by Arellano und Bond(1991). */**/*** indicates statistical significance of the estimated parameters at the 10%/5%/1%-level.*

	FE	**GMM1**	GMM2	GMM3
Lagged net fee income (in % of total assets)	0.394	**0.457**	0.639	0.535
	(0.120)	**(0.132)**	(0.147)	(0.143)
	***	***	***	***
GDP growth (in %)	0.006	**0.005**	0.007	0.006
	(0.003)	**(0.001)**	(0.001)	(0.002)
	**	***	***	***
lagged GDP growth	0.005	**0.003**	-0.002	-0.001
	(0.002)	**(0.001)**	(0.002)	(0.002)
	**	**		
Book equity/total assets (in %)	0.038	**0.036**	0,013	0.020
	(0.011)	**(0.010)**	(0,008)	(0.009)
	***	***		**
Log(Total assets)	-0.520	**-0.459**	-0.148	-0.246
	(0.211)	**(0.180)**	(0.079)	(0.109)
	**	***	*	**
Constant	10.388			
	(4.076)			
	**			
Number of observations	38,555	33,764	33,754	33,754
Number of banks	4,791	4,430	4,430	4,430
Number of instruments		157	21	21
Adj. R^2	0.24			

Table A5: estimation results small banks: net fee income (in % of total assets)

*Note: GMM1 - GMM3 are the GMM estimators as in Arellano und Bond (1991) using three sets of instruments for the model in first difference. GMM1 makes use of all available moment conditions. GMM2 uses only a subset of the moment conditions (first lag) , while GMM3 constructs a linear combination of the moment conditions (see Roodman (2009)). The preferred model is highlighted, which is the efficient GMM estimator as suggested by Arellano und Bond(1991). */**/*** indicates statistical significance of the estimated parameters at the 10%/5%/1%-level.*

	FE	**GMM**	GMM	GMM
Lagged provisions (in % of total loans)	0.203	**0.157**	0.209	0.145
	(0.079)	**(0.058)**	(0.105)	(0.082)
	**	***	**	*
Lagged GDP growth (in %)	-0.034	**-0.035**	-0.030	-0.026
	(0.006)	**(0.006)**	(0.009)	(0.008)
	***	***	***	***
Book equity / total assets (in %)	-0.018	**-0.018**	-0.012	-0.007
	(0.013)	**(0.012)**	(0.019)	(0.017)
RWA / total assets (in %)	0.009	**0.010**	0.009	0.010
	(0.001)	**(0.002)**	(0.003)	(0.003)
	**	***	***	***
Constant	0.080			
	(0.067)			
Number of observations	299	286	286	286
Number of banks	13	13	13	13
Number of instruments		7	9	11
Adj. R^2	0.35			

Table A6: estimation results large banks: net loan loss provisions (in % of total assets)

*Note: Season dummies are included, but are not shown. */**/*** indicates statistical significance of the estimated parameters at the 10%/5%/1%-level.*

	FE	GMM	GMM
Lagged net interest income		-0.173	-0.075
		(0.111)	(0.054)
3M EURIBOR (in %)	**-0.015**	-0.013	-0.016
	(0.006)	(0.004)	(0.005)
	**	***	***
3M EURIBOR 1. lag	**0.023**	0.026	0.025
	(0.005)	(0.004)	(0.004)
	***	***	***
3M EURIBOR 2. lag	**-0.012**	-0.014	-0.013
	(0.003)	(0.003)	(0.003)
	***		***
10J - Bundesanleihe (in %) 1. lag	**0.018**	0.013	0.017
	(0.007)	(0.007)	(0.007)
	**	*	**
10J - Bundesanleihe (in %) 2. lag	**-0.017**	-0.015	-0.017
	(0.006)	(0.006)	(0.006)
	**	**	***
Book equity / total assets (in %)	**0.011**	0.017	0.012
	(0.006)	(0.007)	(0.006)
	*	***	*
RWA / total assets (in %)	**0.002**	0.003	0.003
	(0.001)	(0.001)	(0.001)
	***	***	***
Constant	**0.073**		
	(0.030)		
Number of observations	348	286	286
Number of banks	16	16	16
Number of instruments		14	12
Adj. R^2	0.14		

Table A7: estimation results large banks: net interest income (in % of total assets)

*Note: Season dummies are included, but are not shown. The preferred model is highlighted, which is the static FE model, as the dynamic specifications (including dynamic FE, not shown) tend to be rejected. */**/*** indicates statistical significance of the estimated parameters at the 10%/5%/1%-level.*

	FE	**GMM**	GMM	GMM
Lagged net fee income (in % of total assets)	0.432	**0.467**	0.475	0.471
	(0.069)	**(0.059)**	(0.143)	(0.159)
	***	***	***	***
GDP growth (in %), 2. lag	0.001	**0.001**	0.001	0.001
	(0.000)	**(0.000)**	(0.001)	(0.001)
	*	**		
Book equity / total assets (in %)	0.008	**0.006**	0.006	0.007
	(0.003)	**(0.005)**	(0.005)	(0.004)
	**			
RWA / total assets (in %)	0.001	**0.000**	0.000	0.000
	(0.000)	**(0.000)**	(0.000)	(0.000)
	**			
Constant	-0.000			
	(0.015)			
Number of observations	348	332	332	332
Number of banks	16	16	16	16
Number of unstruments		12	11	10
Adj. R^2	0.37			

Table A8: estimation results large banks: net fee income (in % of total assets)

*Note: Season dummies are included, but are not shown. */**/*** indicates statistical significance of the estimated parameters at the 10%/5%/1%-level.*

Appendix 5: Regression results with additional macroeconomic factors

	FE	FE	GMM1	GMM1
Lagged interest income	0,261	0,250	0,338	0,300
	(0,043)	(0,043)	(0,048)	(0,046)
	***	***	***	***
3M EURIBOR (in %)	0,144	0,138	0,127	0,116
	(0,008)	(0,011)	(0,005)	(0,007)
	***	***	***	***
10J Bund. (in %)	0,100	0,037	0,071	0,043
	(0,012)	(0,004)	(0,010)	(0,008)
	***	***	***	***
10J Bund. 1. lag	0,136	0,172	0,151	0,175
	(0,015)	(0,016)	(0,012)	(0,012)
	***	***	***	***
10J Bund. 2. lag	0,170	0,045	0,112	0,028
	(0,020)	(0,017)	(0,016)	(0,014)
	***	***	***	**
1. Factor		-0,214		-0,172
		(0,020)		(0,018)
		***		***
2. Factor		-0,091		-0,077
		(0,014)		(0,008)
		***		***
3. Factor		-0,027		-0,023
		(0,005)		(0,003)
		***		***
Book equity/total assets (in %)	0,042	0,042	0,029	0,032
	(0,009)	(0,009)	(0,008)	(0,009)
			***	***
Loans/total assets(in %)	0,017	0,015	0,028	0,020
	(0,003)	(0,003)	(0,002)	(0,002)
	***	***	***	***
Constant	0,302	1,189		
	(0,154)	(0,207)		
	*	***		
Beobachtungen	33.754	33.754	29.334	29.334
Anzahl Banken	4.430	4.430	3.942	3.942
Anzahl Instrumente			158	161
Adj. R^2	0,550	0,554		

Table A9: estimation results for small banks: interest income (in % of total assets)

*Note: GMM1 - GMM3 are the GMM estimators as in Arellano und Bond (1991) using three sets of instruments for the model in first difference. GMM1 makes use of all available moment conditions. */**/*** indicates statistical significance of the estimated parameters at the 10%/5%/1%-level. The factors are estimated in a static factor model using 137 macroeconomic time series from Q1 1991 to Q4 2013. The quarterly estimates are averaged to obtain yearly estimates of the factors. The ICP2 criterion of Bai and Ng (2002) selects four factors, of which three turn out to be statistically relevant.*

	FE	**FE**	**GMM1**	**GMM1**
Lagged interest expense	0,438	0,414	0,533	0,446
	(0,038) ***	(0,041) ***	(0,026) ***	(0,034) ***
3M EURIBOR (in %)	0,219	0,216	0,223	0,205
	(0,005) ***	(0,006) ***	(0,005) ***	(0,005) ***
Lagged 3M EURIBOR	-0,036	0,022	-0,027	0,016
	(0,006) ***	(0,011) **	(0,005) ***	(0,008) **
10J Bundesanleihe (in %)	-0,026	-0,075	-0,065	-0,065
	(0,010) ***	(0,006) ***	(0,006) ***	(0,005) ***
Lagged 10J Bundesanleihe	0,174	0,107	0,130	0,111
	(0,011) ***	(0,007) ***	(0,007) ***	(0,006) ***
1. Factor		-0,130		-0,082
		(0,017) ***		(0,015) ***
2. Factor		-0,057		-0,050
		(0,009) ***		(0,007) ***
3. Factor		-0,057		-0,036
		(0,006) ***		(0,004) ***
book equity/ total assets (in %)	0,010	0,002	0,018	0,022
	(0,004)	(0,004)	(0,009) *	(0,009) **
total loans/ total assets (in %)	0,003	0,003	0,012	0,016
	(0,002)	(0,002)	(0,004) ***	(0,004) ***
funding gap/total assets (in %)	0,001	0,001	-0,000	-0,002
	(0,002)	(0,002)	(0,003)	(0,003)
Constant	0,010	0,423		
	(0,155)	(0,162) ***		
Number of observations	38.555		33.764	33.764
Number of banks	4.791		4.430	4.430
Number of instruments			160	163
Adj. R^2	0,671	0,675		

Table A10: estimation results for small banks: interest income (in % of total assets)

*Note: GMM1 - GMM3 are the GMM estimators as in Arellano und Bond (1991) using three sets of instruments for the model in first difference. GMM1 makes use of all available moment conditions. */**/*** indicates statistical significance of the estimated parameters at the 10%/5%/1%-level. The factors are estimated in a static factor model using 137 macroeconomic time series from Q1 1991 to Q4 2013. The quarterly estimates are averaged to obtain yearly estimates of the factors. The ICP2 criterion of Bai and Ng (2002) selects four factors, of which three turn out to be statistically relevant.*

	FE	FE	GMM	GMM
Lagged net interest income			-0,075	-0,087
			(0,054)	(0,046)
3M EURIBOR (in %)	-0,015	-0,025	-0,016	-0,026
	(0,006)	(0,007)	(0,005)	(0,007)
	**	***	***	***
3M EURIBOR 1. lag	0,023	0,018	0,025	0,019
	(0,005)	(0,004)	(0,004)	(0,004)
	***	***	***	***
3M EURIBOR 2. lag	-0,012	-0,004	-0,013	-0,005
	(0,003)	(0,003)	(0,003)	(0,003)
	***		***	
10J - Bundesanleihe (in %) 1. lag	0,018	0,031	0,017	0,031
	(0,007)	(0,009)	(0,007)	(0,009)
	**	***	**	***
10J - Bundesanleihe (in %) 2. lag	-0,017	-0,023	-0,017	-0,022
	(0,006)	(0,007)	(0,006)	(0,007)
	**	***	***	***
1. Factor		-0,014		-0,015
		(0,005)		(0,004)
		***		***
Book equity / total assets (in %)	0,011	0,010	0,012	0,011
	(0,006)	(0,007)	(0,006)	(0,007)
	*		*	
RWA / total assets (in %)	0,002	0,003	0,003	0,003
	(0,001)	(0,001)	(0,001)	(0,001)
	***	***	***	***
Constant	0,073	0,023		
	(0,030)	(0,029)		
Number of observations	299	299	286	286
Number of banks	16	16	16	16
Number of instruments			12	13
Adj. R^2	0,14	0,16		

Table A11: estimation results large banks: net loan loss provisions (in % of total assets)

*Note: Season dummies are included, but are not shown. */**/*** indicates statistical significance of the estimated parameters at the 10%/5%/1%-level. The factors are estimated in a static factor model using 137 macroeconomic time series from Q1 1991 to Q4 2013. The ICP2 criterion of Bai and Ng (2002) selects four factors, of which one turns out to be statistically relevant.*

Appendix 6: Forecasted bank income components for the small banks

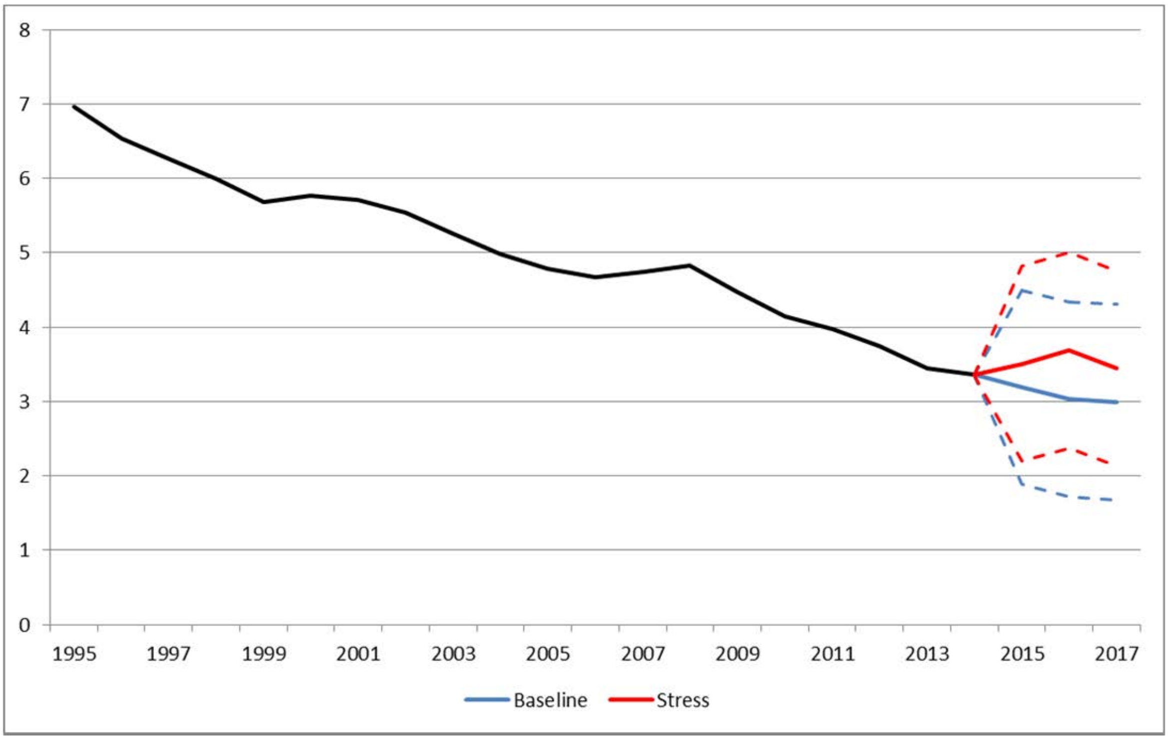

Figure A1: median interest income in baseline and stress scenario (in % of total assets) with 95% confidence bands (dashed lines)

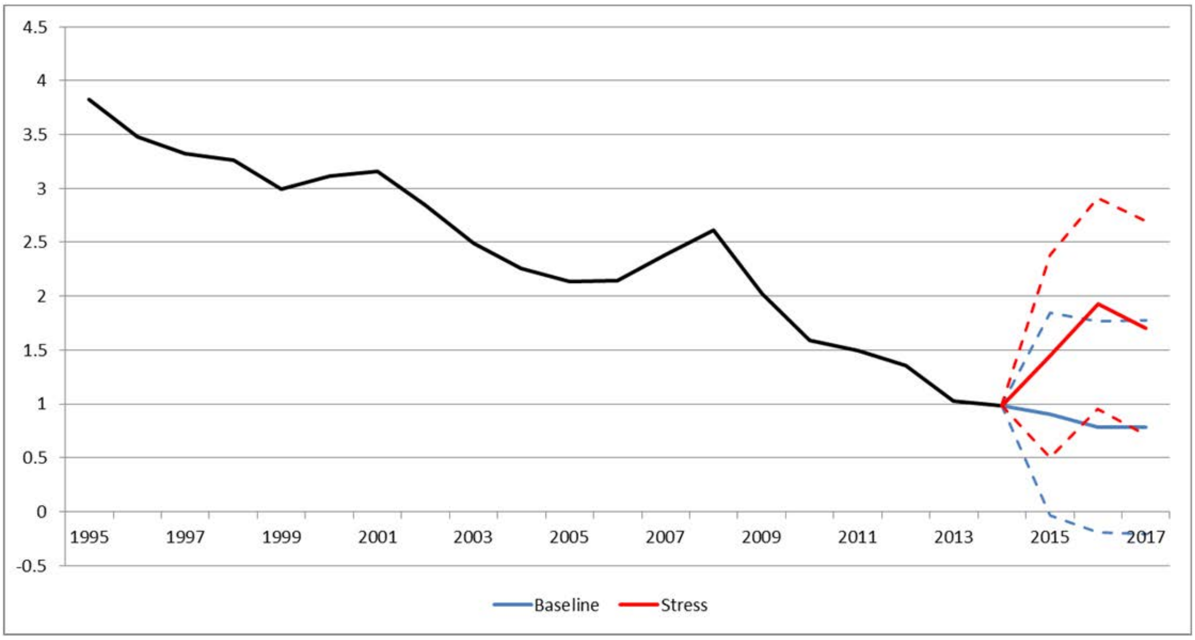

Figure A2: median interest expense in baseline and stress scenario (in % of total assets) with 95% confidence bands (dashed lines)

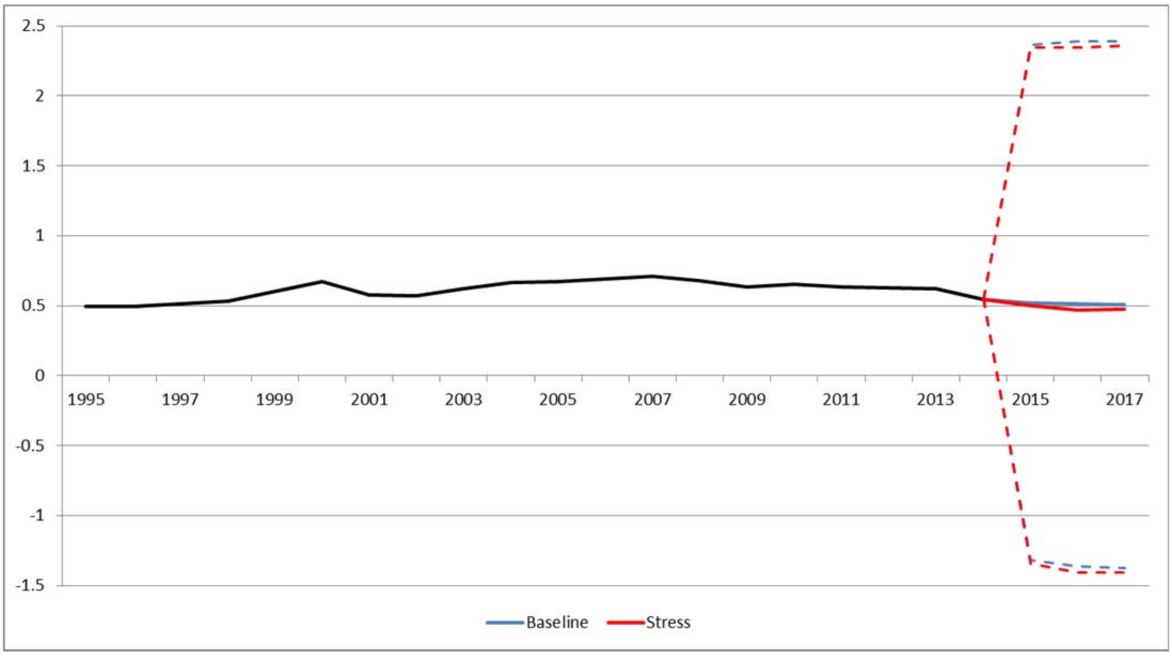

Figure A3: median net fee income in baseline and stress scenario (in % of total assets) with 95% confidence bands (dashed lines)

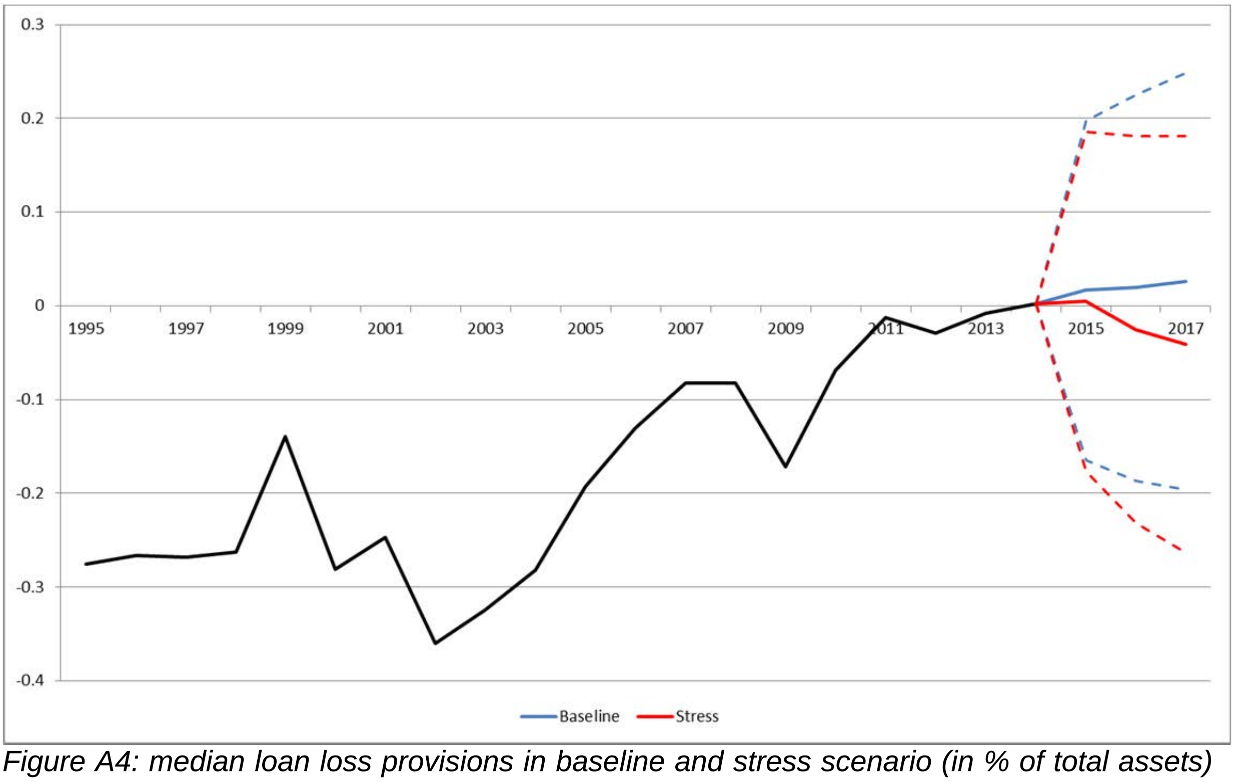

Figure A4: median loan loss provisions in baseline and stress scenario (in % of total assets) with 95% confidence bands (dashed lines)

Appendix 7: Forecasted operating result for small and large banks

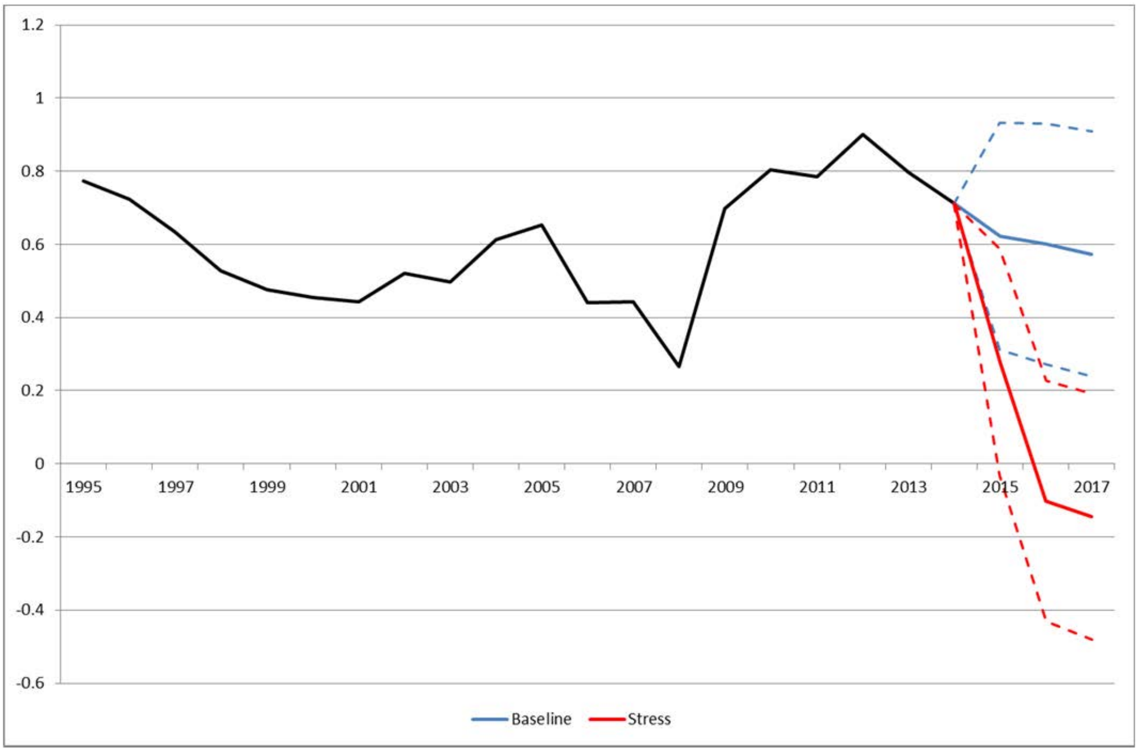

Figure A5: median operating result baseline and stress scenario (in % of total assets) with 95% confidence bands (dashed lines) for the small banks

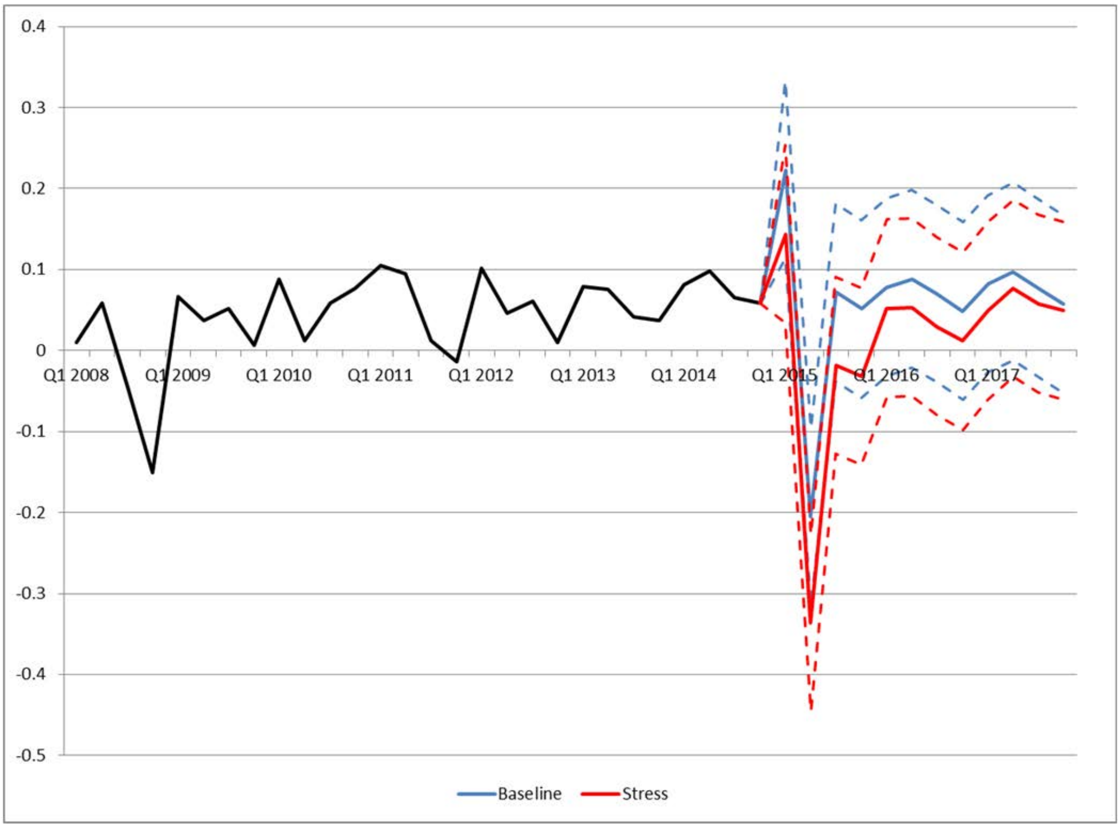

Figure A5: median operating result baseline and stress scenario (in % of total assets) with 95% confidence bands (dashed lines) for the large banks

Appendix 8 Forecast errors for small and large banks

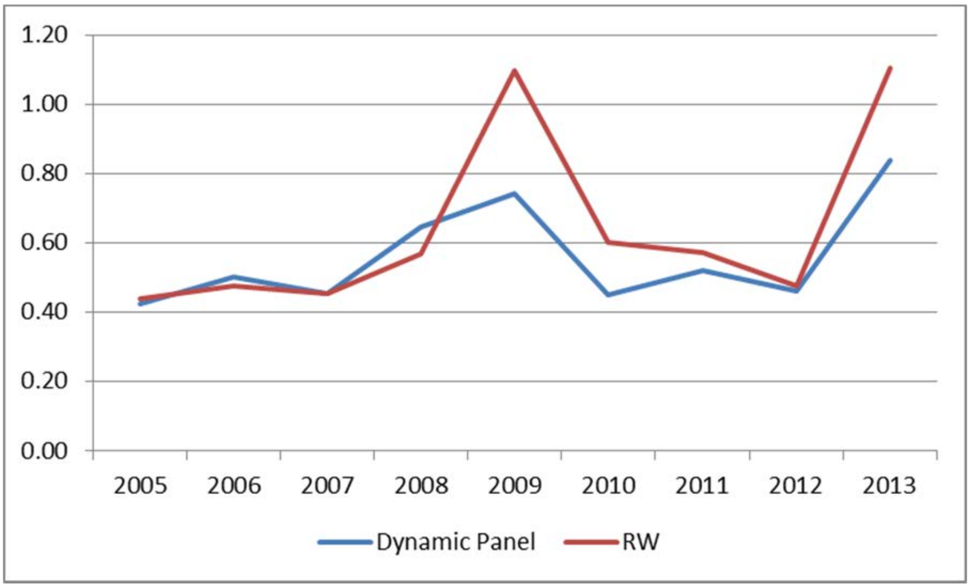

Figure A1:forecast errors for small banks: interest income

Note: The graph shows the evolution of the forecasting errors as given by (1), see Subsection 2.1.4. The forecast error resulting from the dynamic panel model is given by the blue line, while the forecast error from the approach which uses the most recent observation (RW) is given by the black line.

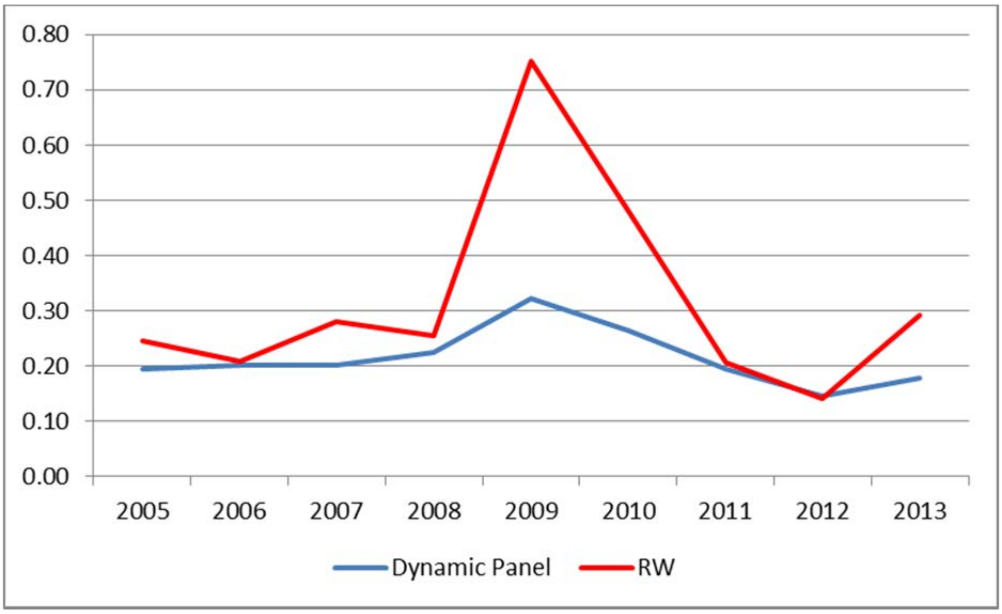

Figure A2:forecast errors for small banks: interest expense

Note: The graph shows the evolution of the forecasting errors as given by (1), see Subsection 2.1.4. The forecast error resulting from the dynamic panel model is given by the blue line, while the forecast error from the approach which uses the most recent observation (RW) is given by the black line.

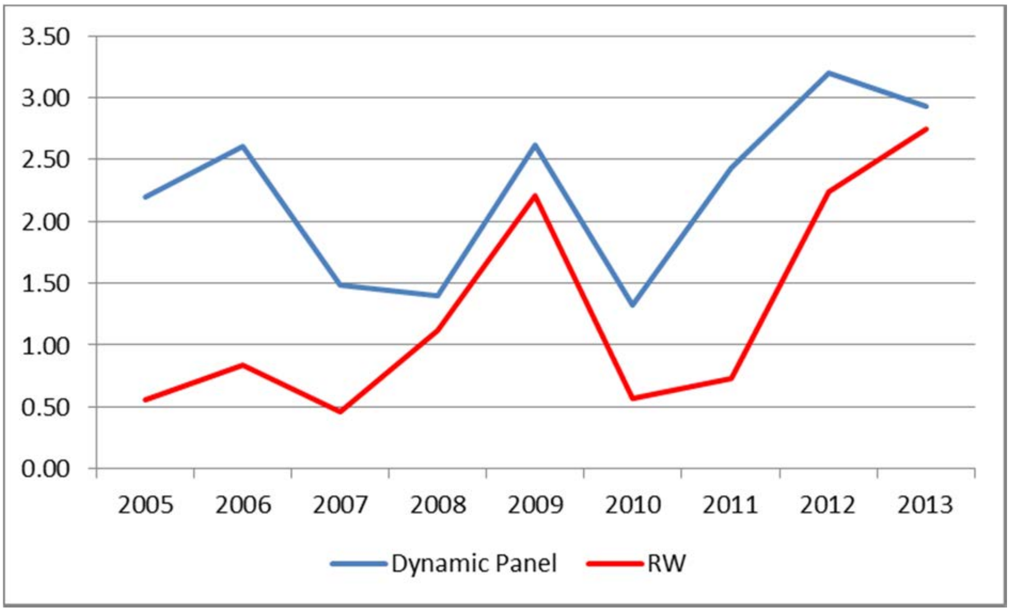

Figure A3:forecast errors for small banks: net fee income

Note: The graph shows the evolution of the forecasting errors as given by (1), see Subsection 2.1.4. The forecast error resulting from the dynamic panel model is given by the blue line, while the forecast error from the approach which uses the most recent observation (RW) is given by the black line.

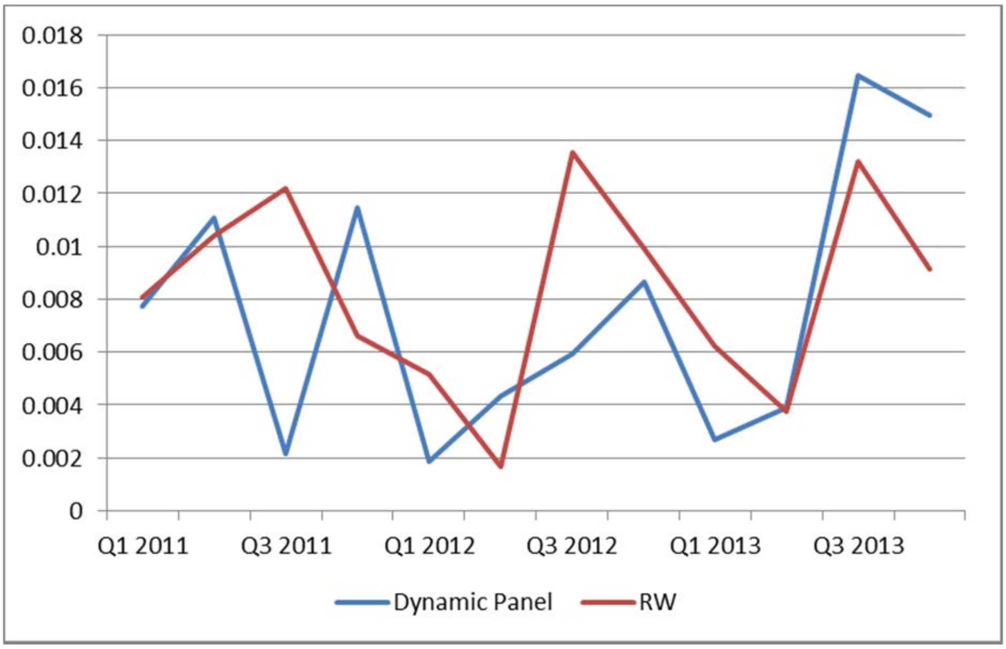

Figure A4:forecast errors for large banks:loan loss provisions

Note: The graph shows the evolution of the forecasting errors as given by (1), see Subsection 2.1.4. The forecast error resulting from the dynamic panel model is given by the blue line, while the forecast error from the approach which uses the most recent observation (RW) is given by the black line.

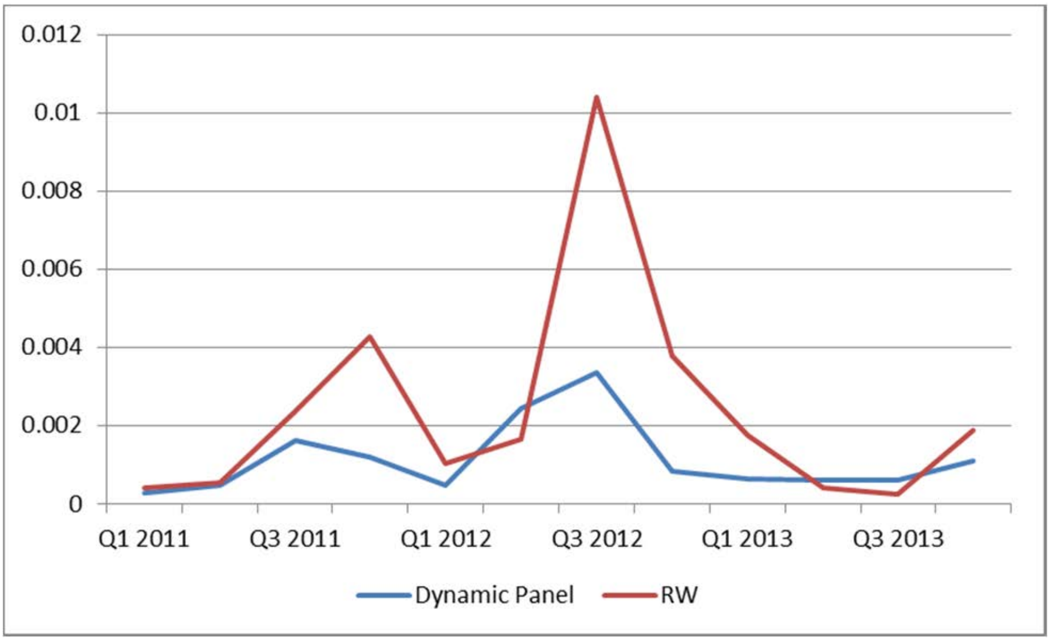

Figure A5:forecast errors for large banks:net interest income

Note: The graph shows the evolution of the forecasting errors as given by (1), see Subsection 2.1.4. The forecast error resulting from the dynamic panel model is given by the blue line, while the forecast error from the approach which uses the most recent observation (RW) is given by the black line.

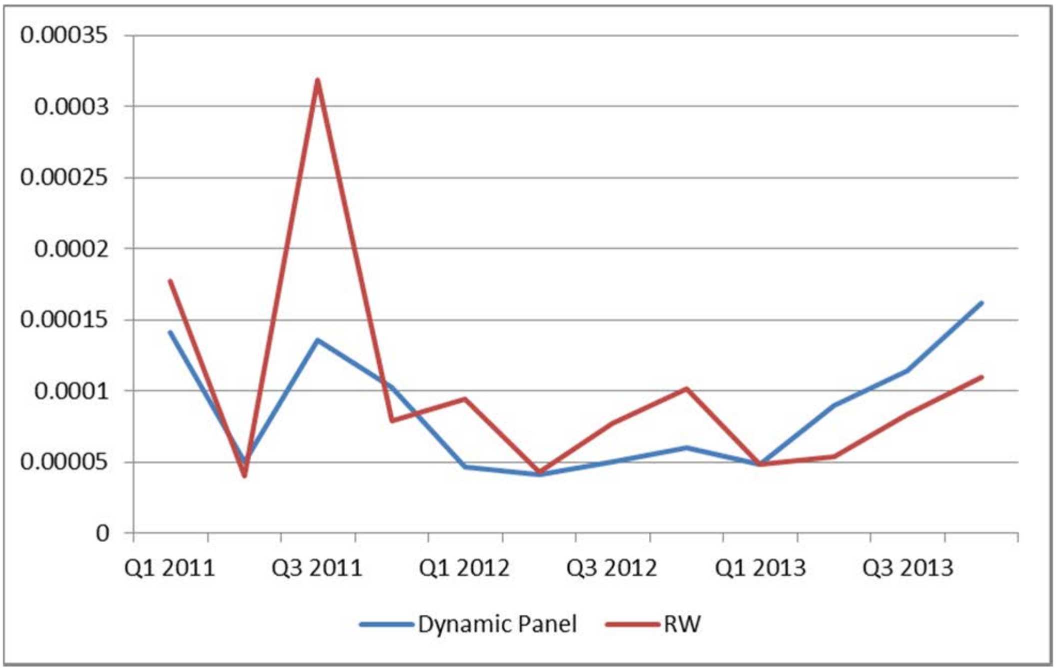

Figure A6:forecast errors for large banks:net fee income

Note: The graph shows the evolution of the forecasting errors as given by (1), see Subsection 2.1.4. The forecast error resulting from the dynamic panel model is given by the blue line, while the forecast error from the approach which uses the most recent observation (RW) is given by the black line.

Appendix 9: Sectoral regressions: Sensitivities to GPD growth and unemployment rate

Sector	GDP growth		Unemployment Rate		Const.	R^2
	Lag	Coeff.	Lag	Coeff.		
Agriculture and forestry	0	-0.00169	4	0.02960	0.00616	0.53
Energy and water supply	0	-0.00020	0	0.00211	-0.00218	0.18
Chemical industry	0	0.00000	0	0.12695	-0.01093	0.57
Plastic and rubber production	0	-0.00217	4	0.07339	-0.00075	0.77
Glass industry	3	-0.00236	0	0.00000	-0.00010	0.55
Metal production and processing	0	-0.00059	4	0.20843	-0.01514	0.75
Mechanical engineering	0	-0.00111	4	0.12033	-0.01101	0.65
Data processing	0	0.00000	3	0.18473	-0.02203	0.49
Wood products	0	0.00000	4	0.06299	-0.02259	0.10
Textiles and clothing	1	-0.00246	3	0.05828	-0.00911	0.48
Food industry	0	-0.00019	0	0.00193	-0.01161	0.03
Construction	0	-0.00515	0	0.00000	0.02049	0.91
Automobile industry	0	-0.00141	3	0.04058	-0.00262	0.80
Transportation and storage; Communications	0	0.00000	3	0.05936	-0.01206	0.05
Financial services and insurance	0	-0.00033	4	0.01350	0.00073	0.67
Housing services	0	-0.00266	4	0.00600	0.00843	0.87
Investments and associations	0	-0.00055	4	0.08869	-0.00726	0.47
Real estate services	0	-0.00267	4	0.00072	0.00756	0.80
Catering and hotel industry	1	-0.00291	4	0.01029	-0.00076	0.88
Information technology; Research and development	0	-0.00087	2	0.09843	-0.00879	0.71
Health care	4	-0.00099	4	0.00986	0.00078	0.76
Renting of moveable propert	0	0.00000	0	0.01821	-0.00873	0.00
Other services	0	-0.00124	4	0.14272	-0.00660	0.74
Private person (housing excluded)	4	-0.00113	2	0.02957	-0.00083	0.90
Installment loans (housing excluded)	3	-0.00232	4	0.05081	-0.00202	0.68
Housing	4	-0.00049	2	0.01617	-0.00206	0.67
Non - profit organisations	3	-0.00051	2	0.01165	0.00075	0.72

Table A12: Sensitivities to GPD growth and unemployment rate for the different industries and typs of household loans

Appendix 10: Risk weights according to Basle II

$$RW(pd) = c \cdot \left(\Phi\left(\frac{1}{\sqrt{1-R}} \cdot \Phi^{-1}(pd) + \sqrt{\frac{R}{1-R}} \cdot \Phi^{-1}(0.999) \right) - pd \right)$$

with

$$R = 0.12 + 0.12 \cdot e^{-50 \cdot pd}$$

where $\Phi(\cdot)$ is the cumulative density function of the standard normal distribution and *c* is a constant that depends on the LGD and several other fixed parameters which are not discussed in detail here. Please note that we implicitly assume that all credit exposures are corporate loans.

Appendix 11: Beta-distribution[17]

The density of the beta distribution is given by

$$f(x) = \frac{1}{B(\alpha,\beta)} \cdot x^{\alpha-1} \cdot (1-x)^{\beta-1} \quad x \in (0,1)$$

with

$$B(\alpha,\beta) = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

where Γ(·) is the gamma function. The parameters $\alpha > 0$ and $\beta > 0$ determine the shape of this distribution. The beta distribution is especially suited to model LGDs because (i) the domain is confined to the economic sensible interval from 0 to 1, (ii) it is highly flexible, and (iii) it nests other distributions. For instance, when both parameters equal 1, then the beta distribution becomes a uniform distribution. When both of the parameters are smaller than 1, the probability density function is u-shaped, with a large portion of the probability mass close to 0 and 1. For parameter values close to 0, this distribution converges to the binomial distribution. By contrast, the density is unimodal in the case of both parameters α and β being greater than 1. For very large parameter values, it converges to the degenerate distribution, where the entire probability mass is concentrated on one point.

We use the parameter estimates derived in Memmel et al. (2012) and set α to 0.28 and β to 0.35, which, as said above, yields a u-shaped density function with an expectation at 0.45.

[17] These explanations follow Memmel et al. (2012).

Appendix 12: Tables concerning the real estate risk

BelA-categories in AFS survey	< 50%	50-<60%	60-<70%	70-<80%	80-<90%	90-<100%	100-<110%	110-<120%	≥ 120%
Assumed mean BelA in respective class	40%	55%	65%	75%	85%	95%	105%	115%	130%

Table A13: Mapping of BelA-categories in AFS survey into mean BelA values for loss estimation

	ΔB	Δf
Baseline scenario	15%	10%
Adverse scenario	15%	40%

Table A14: Additional parameter inputs

References

Aikman, D., B. Eklund, P. Gai, S. Kapadia, E. Martin, N. Mora, G. Sterne and M. Willison (2009). Funding Liquidity Risk in a Quantitative Model of Systemic Stability, in: Financial Stability, Monetary Policy, and Central Banking, edited by R. A. Alfaro, Central Bank of Chile.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies 58, pp. 277 – 297.*

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70, pp. 191 – 221.*

Blundell, R. and S. Bond (1998). Initial conditions and moment restricitons in dynamic panel data models. *Journal of Econometrics 87, pp.115 – 143.*

Bun, M. and V. Sarafidis (2013). Dynamic panel data models. *University of Amsterdam Discussion paper 2013/01.*

Bundesbank (2013), Financial Stability Report, November 2013.

Covas F., B. Rump and E. Zakrajsek (2014). Stress – testing US bank holding companies: a dynamic panel quantile regression approach. *International Journal of Forecasting 30, pp. 691 – 713.*

De Jonghe, O. and Ö. Ötzekin (2010). On Regulation, Supervision, (De)leveraging and Banking Competition: Harder, Better, Faster, Stronger?, *Working Paper Bocconi 13/2010.*

EBA (2012). Update on the implementation of Capital Plans following the EBA's 2011 Recommendation on the creation of temporary capital buffers to restore market confidence, *Presse release 11 July 2012.*

Eisenberg, L. and T. H. Noe (2001). Systemic Risk in Financial Systems. *Management Science 47 (2), pp. 236–49.*

ESRB (2015). Second-round effects of the 2014 EBA stress test on banks, Document for the meeting on 19 March 2015 (confidential).

Fink, K., U. Krüger, B. Meller, L. Wong (2014). BSLoss - a comprehensive measure for interconnectedness, Mimeo, (available on the EBA-webpage: https://www.eba.europa.eu/news-press/calendar?p_p_id=8&_8_struts_action=%2Fcalendar%2Fview_event&_8_eventId=846258).

Francis, W. and M. Osborne (2009). Bank regulation, capital and credit supply: Measuring the im-pact of Prudential Standards, *FSA Occasional Papers in Financial Regulation 36.*

Gropp, R. and F. Heider (2010). The Determinants of Bank Capital Structure, *Review of Finance, Vol. 14, pp. 587-622.*

Jokipii, T. and A. Milne (2008): The cyclical behaviour of European bank capital buffers, *Journal of Banking and Finance, Vol. 32, pp. 1440-1451.*

Jokipii, T. and A. Milne (2011): Bank capital buffer and risk adjustment decisions, *Journal of Financial Stability, Vol. 7, pp. 165-178.*

Kripfganz, S and C. Schwarz (2013). Estimation of linear dynamic panel data models with time – invariant regressors. *Deutsche Bundesbank Discussion Paper, No 25/2013.*

Maddala, G., R. Trost, H. Li, and F. Joutz (1997): Estimation of short – run and long – run elasticities of energy demand from panel data using shrinkage estimation, *Journal of Business & Economic Statistics, Vol. 15, pp. 90-100.*

Memmel, C., Y. Gündüz and R. Raupach (2015). Common Drivers of Default, *Journal of Financial Stability, Vol. 16, pp. 232-247.*

Memmel, C. and P. Raupach (2010). How do banks adjust their capital ratios?, *Journal of Financial Intermediation, Vol. 19, pp. 509-528.*

Memmel, C., A. Sachs and I. Stein (2012). Contagion at the Interbank Market with stochastic Loss Given Default, *International Journal of Central Banking, Vol. 8 (3), pp. 177-206.*

Memmel, C. and A. Schertler (2013). Banks' management of the net interest margin: New measures, *Financial Markets and Portfolio Management, Vol. 27(3), pp. 275-297.*

Nickell S. (1981). Biases in dynamic models with fixed effects. *Econometrica 49, pp.1417 – 1426.*

Roodman, D. (2009). How to do xtabond2: an introduction to difference and system GMM in Stata. *The Stata Journal 9, pp. 86 – 136.*

Wooldridge, J. (2012). Introductory econometrics: a modern approach. 5th edition, South Western.