DEUTSCHE BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 37/2017

A stress test framework for the German residential mortgage market – methodology and application

Thomas Siemsen
Johannes Vilsmeier

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978–3–95729–416–6 (Printversion)
ISBN 978–3–95729–417–3 (Internetversion)

Non-technical summary

Research Question

German house prices have been rising strongly since 2010 following decades of modest growth. These price dynamics are increasingly raising concerns about a real estate bubble in Germany which may pose a threat to financial stability. Recent estimates by the Bundesbank suggest overvaluations between 15% and 30% in German cities. We take this as a starting point for assessing the resilience of German less significant institutions (LSIs) to a severe decline in house prices.

Contribution

Our contributions are fourfold: first, using a unique and very recent survey data set, we assess the impact of a severe decline in house prices (up to 30%) on banks' solvency over a three-year stress horizon. The survey provides us with bank-specific information on current risk parameters and exposures in residential mortgage portfolios. Second, to map the macro-scenario to default probability (PD) dynamics, rather than applying one single model, we estimate a battery of different specifications to reduce model uncertainty. Third, we derive a relationship between loan to value (LTV) and loss given default (LGD) based on a review of the literature. And fourth, we study the effects of different models for risk-weighted assets (RWA) on stress test results by comparing stress impacts under the Standardized Approach for credit risk (SA), a potential revision to the SA and the internal rating based approach (IRB).

Results

We find German LSIs to be mostly well capitalized against a severe decline in house prices. However, stress test results are highly heterogeneous among banks. The median bank experiences - compared to its own planning - a reduction in CET1 ratio of 1.5pp over the stress horizon, induced by an isolated stress to banks' residential mortgage portfolios. About 12% of banks suffer a reduction of more than 3pp. Comparing stress effects under different RWA regimes, we find that under the SA the stress effect is up to 33% lower than under the IRB approach. Revisions to the SA which are currently und discussion would make the approach more risk-sensitive, implying an increase of the stress effect by 20% relative to the current SA.

Nichttechnische Zusammenfassung

Fragestellung

Nach langjährig flachen Preisdynamiken im deutschen Wohnimmobiliensektor, steigen die Preise seit 2010 stark an. Dies schürt zunehmend die Besorgnis über eine Überhitzung des Marktes. Schätzungen der Bundesbank halten Überbewertungen in Städten von 15% bis 30% für möglich. Vor diesem Hintergrund analysiert dieses Papier die Widerstandsfähigkeit kleiner und mittelgroßer deutscher Banken gegenüber einem starken Hauspreiseinbruch.

Beitrag

Das Papier quantifiziert, basierend auf aktuellen Umfrage-Daten, potentielle Verluste deutscher Banken infolge eines starken Einbruchs der Hauspreise (um bis zu 30%). Die Umfrage bietet Einblick in die aktuelle Risikoposition aus Wohnimmobilienkrediten von ca. 1500 Instituten, inkl. Ausfallwahrscheinlichkeiten (PD) und Beleihungsausläufe (LTV). Um den Effekt des Szenarios auf die PD zu schätzen, verwenden wir einen Ansatz, der die Modellunsicherheit reduziert, indem er eine große Menge an Regressionsmodellen geeignet kombiniert. Zur Berechnung gestresster Verlustquoten (LGD) leiten wir aus der Literatur eine empirische Beziehung zwischen LTV und der LGD her. Wir analysieren den Einfluss verschiedener Ansätze zur Modellierung von Risikogewichten auf Stresstestergebnisse indem wir die Stresswirkungen unter dem Kredit-Standard-Ansatz (KSA), einer möglichen Revision des KSA und unter dem auf bankinternen Ratings basierenden Ansatz (IRB) gegenüberstellen.

Ergebnisse

Die Ergebnisse zeigen, dass die meisten Institute ausreichend widerstandsfähig gegenüber einem starken Einbruch der Hauspreise sind. Jedoch herrscht eine starke Heterogenität zwischen den Instituten. Der isolierte Stress auf Wohnimmobilienkredite führt im Median zu einer Reduktion der harten Kernkapitalquote (CET1-Quote) um 1,5 Prozentpunkte (Pp) im Vergleich zu bankinternen Prognosen. Für 12% der Institute reduziert sich die CET1-Quote um mehr als 3 Pp. Der Vergleich der Stresstestergebnisse für verschiedene Risikogewicht-Modellierungen zeigt, dass im KSA der Stresseffekt um bis zu 33% unter dem IRB Ansatz liegt. Die aktuell diskutierte Revision des KSA würde diesen risikosensitiver machen, was zu einem Anstieg des Stresseffektes um 20% relativ zum aktuellen KSA führen würde.

Bundesbank Discussion Paper No 37/2017

A stress test framework for the German residential mortgage market - methodology and application*

Thomas Siemsen
Deutsche Bundesbank

Johannes Vilsmeier
Deutsche Bundesbank

Abstract

This paper exploits a recent and granular data set for 1,500 German LSIs to conduct a residential mortgage stress testing exercise. To account for model uncertainty when modeling PD dynamics we use a benchmark-constrained Bayesian model averaging approach that combines standard BMA with a benchmark derived from a quantile mapping between the historical PD distribution and the historical distribution of macro variables. To link LGD to current LTV we derive a reduced-form meta-dependency. In the baseline model, we quantify expected as well as unexpected losses. We show that German LSIs, though being mostly sufficiently capitalized, are susceptible to a corrective movement in house prices with a median CET1 ratio reduction of 1.5pp in the severely adverse scenario. We quantify the impact of RWA modeling on stress test results and show that the Standardized Approach leads to an up to 33% lower stress impact relative to the more risk-sensitive "pseudo-IRB" approach.

Keywords: stress test, Bayesian model averaging, quantile mapping, survey data, German residential mortgage market, model uncertainty

JEL classification: C11, C52, G21

*We would like to thank Thomas Kick (the editor) and seminar participants at Deutsche Bundesbank for their helpful comments. Contact address: Deutsche Bundesbank, Wilhelm-Epstein Strasse 14, 60431 Frankfurt am Main, Germany, e-mail: *thomas.siemsen@bundesbank.de*, *johannes.vilsmeier@bundesbank.de*.

1 Introduction

In Germany, house prices have been rising strongly since 2010 following decades of modest growth. This accelerating growth can partly be traced back to the low-interest-rate environment (LIRE), which leaves German banks with few profitable products, such as mortgage credit, while low interest payments keep housing affordable despite rising prices. The price dynamics are raising concerns about a real estate bubble in Germany which may pose a threat to financial stability. Recent estimates by the Bundesbank suggest that residential real estates may be overvalued between 15% and 30% in German cities. Real estate mortgage exposures make up a substantial percentage of total credit exposure in German less significant institutions' (LSIs) books with an aggregate share of 25% at the end of 2016. In such an environment, obtaining sound estimates of potential losses looming in banks' books due to a corrective movement in house prices is of major importance.

The proposed stress testing framework sheds light on the quantitative implications of a severe decrease in German house prices on banks CET1 ratios, transmitted through residential mortgage portfolios. In the adverse scenario, we assume a house price decrease of 20%, while in the severely adverse scenario we assume a decrease of 30%. Both scenarios are consistent with the recent estimates by the Bundesbank concerning overvaluation in German cities.

Our contributions are fourfold: *first*, there is little evidence on the resilience of German banks to residential mortgage stress, probably due to the lack of sufficiently granular data. We add to the stress testing literature by using a unique and very recent data set, collected through the Bundesbank's LIRE survey, to conduct a stress test for banks' residential mortgage portfolios. The data set provides timely (as of December 2016) and granular information on banks' internal planning data over the entire stress horizon, residential mortgage risk parameters and exposures at a rating class level, as well as (long-run) historical loss and default rates at the portfolio level.

Second, we add to the stress testing literature by employing a *benchmark constrained Bayesian model averaging* (BCBMA) framework to establish a sound mapping of the macro scenario to probability of default (PD) dynamics. Our BCBMA model combines the BMA approach of Henry and Kok (2013) with the quantile mapping idea of Bonti, Kalkbrener, Lotz, and Stahl (2006) in order to reduce model uncertainty and to provide sound estimates of the relationship between PDs and macro dynamics. Model uncertainty is an issue particularly in a fragile statistical environment with few observations and highly correlated macro covariates, as present in many top-down stress testing applications.

Third, we derive a continuous reduced-form dependency between current loan-to-value (CLTV) and loss given default (LGD) for German banks by combining and extending the studies of Qi and Yang (2009) and Palmroos (2016). Qi and Yang provide information on the dependency between CLTV and LGD for CLTV buckets that are too coarse for stress testing purposes. Palmroos derives a continuous relationship based on Finnish data. However, the level of this relationship seems not to be in line with German experience. We contribute to the stress testing literature by deriving a meta-dependency which equips the practitioner with a traceable tool for translating CLTV into LGD and vice versa without requiring granular data for calibration.

And *fourth*, we study the effect of different models for risk-weighted assets (RWA) on stress test results. To this end, we contrast the CET1 ratio impact under different RWA dynamics derived from our baseline "pseudo-IRB" approach, the current SA and from the SA revision as suggested in Basel Committee on Banking Supervision (2015).

The stress test is run on all banks which took part in the 2017 LIRE survey and which have exposure to residential mortgages. This leaves us with 1401 institutions. Owing to access to banks' internal planning data, we can compute stress effects as the additional loss stemming exclusively from the scenario-induced dynamics relative to banks' planning data.

The main findings of the stress test are threefold: *first*, we find German LSIs to be mostly well capitalized for weathering a severe decline in German house prices. However, the stress test results are highly heterogeneous across banks. While the median bank experiences a reduction in the CET1 ratio of 0.6pp (1.5pp) in the adverse (severely adverse) scenario, the 95%-quantile bank suffers from a reduction of 1.6pp (4.0pp). About 12% of banks suffer from a reduction of more than 3pp in the severely adverse scenario, the maximum reduction being 11.1pp.

Second, when comparing stress effects under different models for RWA, we find that RWA increase by 4.6% under the "pseudo-IRB" approach over the three-year stress horizon of the severely adverse scenario but by only 1.4% under the SA. The sluggish KSA RWA dynamics reduce the median stress effect from -1.3pp to -0.9pp. We model RWA dynamics to capture unexpected losses during the stress horizon. Since unexpected losses, similar to expected losses, increase continuously with deteriorating macro conditions, our results would tend to support the view that employing the risk-insensitive SA to approximate unexpected losses may substantially bias stress effects downwards, also for banks using the SA for regulatory purposes.

And *third*, when estimating the model space of the dependency of the residential mortgage PD time series on macro variables, we find a high heterogeneity in estimated coefficients, many showing a high R^2 but having economic "implausible" coefficient signs, low out-of-sample predictive power or implausible stress effects when compared to a benchmark. This result emphasizes the importance of looking at the entire model space and filtering for "plausible" model specifications, instead of employing ad hoc specifications.

While the residential mortgage market has been studied extensively (see Allen, 2004, for a survey of the literature), the literature on residential mortgage stress tests is sparse.[1] Rodriguez and Trucharte (2007) conduct a stress testing exercise for the aggregate Spanish mortgage market using a simulation approach to derive a credit loss distribution assuming constant LGD. Coleman, Esho, Sellathurai, and Thavabalan (2005) suggest a stress testing framework for the Australian mortgage market using a combination of reduced-form and ad-hoc dependencies to model credit risk parameters. For Ireland, Gaffney, Kelly, and McCann (2014) use a transition-based PD framework to model stressed PD dynamics which accounts for hysteresis effects in loan cures. Hott (2015) employs a calibrated structural model for the US and Switzerland to conduct a mortgage stress test on the macro level. While such an approach may serve as a top-down benchmark, it lacks the granularity necessary for microprudential stress testing.[2]

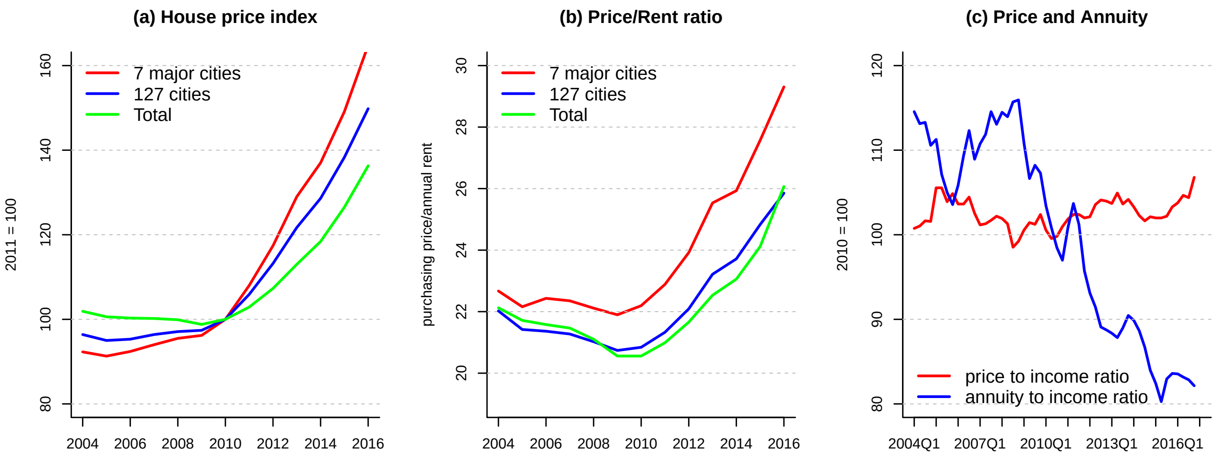
[1]For a survey of the credit risk stress testing literature in general, see Foglia (2009).

[2]See also Basel Committee on Banking Supervision (2012) for a discussion of the role of macro models in stress testing.

The remainder of the paper is structured as follows: Section 2 provides a brief overview of the current situation in the German housing market, Section 3 lays out our stress testing framework, Section 4 describes the data and the macro scenarios, and Section 5 discusses quantitative results including a comparison of different RWA models. Section 6 concludes.

2 The German housing market

This section provides a brief overview of recent developments in the German housing market. Germany has traditionally experienced only weak growth (if any) in house prices with an average annual growth rate of -0.7% between 1991 and 2009. However, as the left-hand panel in Figure 1 shows, since 2010 growth accelerated and the average annual growth rate between 2009 and 2016 increased to 2.3%. The increase in house prices sent price-rent-ratios soaring, especially in the seven major German cities[3] where it reached 30 in 2016 (see middle panel). One factor driving this development, besides fundamentals such as limited space in urban areas, is the low-interest-rate environment. With interest rates in most developed economies at all-time lows, investors are deprived of profitable outside options and are thus willing to accept low returns on housing investments. On the other hand, low interest rates mask increasing house prices through low annuities (see right-hand panel). Both effects are currently keeping demand up for housing and mortgage credit. Still, the situation can be fragile, as further increasing prices may at some point lead investors to start selling their properties or a tightening of monetary policy may put upward pressure on mortgage lending rates. This could trigger a downward movement of prices, raising losses (when selling a house bought at a high price) and default rates (when borrowers cannot entertain their interest payments) along the way. Current Bundesbank estimates indicate potential overvaluations in German cities between 15% and 30% (see Deutsche Bundesbank, 2017).



Source: Deutsche Bundesbank

Figure 1: Standard indicators for the German real estate market

[3]Munich, Stuttgart, Frankfurt am Main, Cologne, Hamburg, Berlin, Düsseldorf.

The LIRE survey shows that residential mortgages are material for German LSIs, which constitute the sample for our stress testing exercise. Banks participating in the LIRE survey had a combined residential mortgage exposure of about €640 billion, 25 % of their total credit exposure, at the end of 2016. Given this significant share, deteriorating real estate prices harbor the potential for high losses. The aim of this paper is to quantify these losses.

3 Methodology

This section sets out our stress testing framework. The framework consists of three model blocks: *first*, a model that maps the macro scenario to stress PD dynamics using a benchmark constrained Bayesian model averaging approach (see Section 3.1), *second*, a model that maps the scenario to stress LGD dynamics using a meta-dependency between current LTV (CLTV) and LGD (see Section 3.2) and *third*, a set of recursive equations that translates stressed risk parameters to P&L effects (expected losses) and RWA effects (unexpected losses) (see Section 3.3). The modeling of the stress effects follows the EBA methodology (see European Banking Authority, 2016) which considers impairments on new defaulted assets, old defaulted assets and risk-weighted assets. In addition, our granular data allows us to model foregone interest payments due to defaulted assets. Before going into greater detail, Figure 2 gives a high-level overview of our stress testing framework.

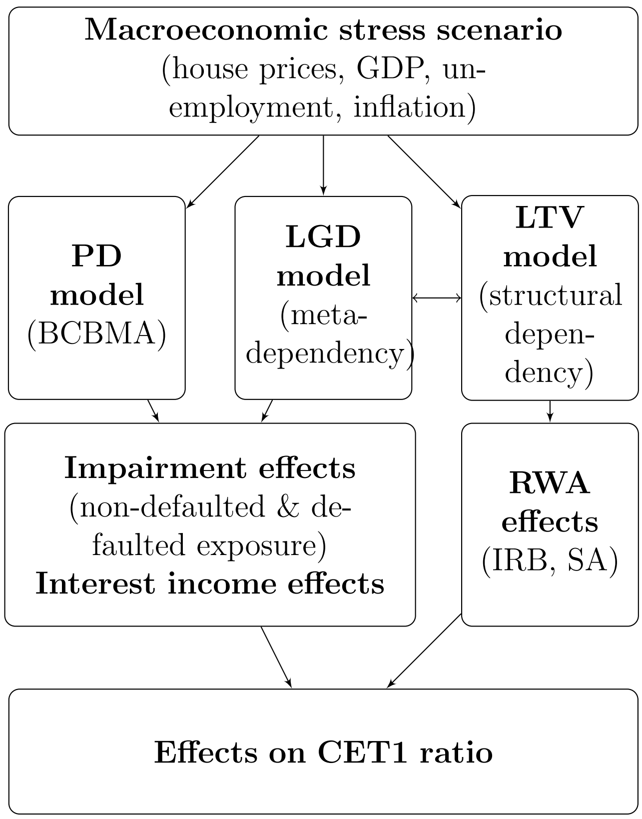


Figure 2: Overview stress testing framework

3.1 PD model

The PD model combines the standard BMA approach with a quantile mapping procedure suggested by Bonti et al. (2006), which maps quantiles of the macro variable distribution to quantiles of the PD distribution.

The rationale behind Bayesian model averaging (BMA) is the need to account for model uncertainty in statistical analyses. Model uncertainty is a concern especially when using short time series (as in many macro stress testing applications) and highly correlated covariates. In these cases, regressor coefficients may be non-unique, since variables catch up on effects of other variables. In such an empirical environment it is highly recommendable to analyze the entire model space rather than picking one model at random or by discretion and neglecting the information from all other models (for another application see Pelster and Vilsmeier, 2017). This is especially true in a stress testing context, where biased estimates can substantially affect stress test outcomes (see Gross and Poblacion, 2017).[4]

While model uncertainty is extensively discussed in the macroeconomic forecasting literature (see Wright, 2008; Hirano and Wright, 2017, among others) it is mostly neglected in the stress testing literature. Stress tests, as pointed out by Misina and Tessier (2008), face the additional problem that they forecast dynamics at the edge of or even beyond the "observation space" where regression-based models, which extrapolate average observations from the sample period, are likely to perform poorly (see also Corbae, D'Erasmo, Galaasen, Irarrazabal, and Siemsen, 2017). This induces an additional source of model uncertainty, since none of the (usually OLS-based) models can properly describe the true stressed relationship. Credit risk stress test models which deal explicitly with issues of model uncertainty are, to our knowledge, restricted to Misina and Tessier (2008), who emphasize theoretically the need to account for model risk and non-linearity in credit risk stress tests, and Henry and Kok (2013) and Gross and Poblacion (2017), who use a BMA approach to carry out their credit risk stress test exercise.

To derive time series for stressed PDs at the bank level we proceed in two steps: first, we derive a mapping between residential mortgage sector PD and the macro scenario, which we use to compute stressed time series for the sector PD. Second, we translate the stressed sector PD to bank-specific starting values in the distance-to-default space.

We consider a standard autoregressive distributed lag specification (ADL):

$$\Delta^4 \log\left(\frac{PD}{1-PD}\right)_t = \sum_{i=1}^{K} \alpha_i \Delta^4 \log\left(\frac{PD}{1-PD}\right)_{t-i} + \sum_{i=0}^{L} \beta_i' x_{t-i} + \varepsilon_t\,, \tag{1}$$

where PD denotes the quarterly residential-mortgage-sector probability of default derived from the German Credit Register and x_t denotes a vector of macro variables, which includes growth rates for GDP, inflation, unemployment rate, house prices and the 3M EURIBOR (the 3M EURIBOR is also included in levels). To derive robust estimates of this dependency despite the rather short time series for PDs (starting in 2008Q1) and the fact that the macro covariates are highly correlated (and estimates are thus very sensitive to the inclusion and exclusion of covariates), we estimate the entire model space for Equation (1). This implies

[4]Note moreover that using a panel data set is no remedy for the issue of model uncertainty, since the identification of variable dynamics requires sufficient observations in the time dimension.

estimating Equation (1) for all possible x_t, i.e. combinations of the macro covariates, for all possible lag lengths and definitions of growth rates (quarter-on-quarter, year-on-year and, for the 3M EURIBOR, also levels).[5] Then, we filter the model space for models which satisfy the following set of restrictions:

1. highly correlated covariates are not within the same model (exclusion restriction),
2. models are econometrically correctly specified (no autocorrelation in residuals),
3. estimated coefficients satisfy economic plausibility (sign restriction), and
4. the models have no significantly worse out-of-sample forecast performance than the best model in the model space as measured by the "leave-one-out" principle (Occam's window).[6]

We refer to models that survive filtering criteria 1 to 4 as "filtered model space". Under a standard BMA approach, one would now combine all surviving models, weighting each model with its posterior inclusion probability. In doing so, all statistically and economic relevant information would be used in the final model in order to derive a robust estimate for the dependency of the sector PDs on macroeconomic conditions. However, we pursue a different route. While restrictions 1 to 4 impose minimum requirements on statistical and economic soundness, they do not necessarily guarantee "plausible" stress test dynamics. For example, a model in the filtered model space may well imply that under the adverse macroeconomic conditions PDs increase by 1000% or decrease by 50%, which may be regarded as implausible judging from historical experience and the severity of the assumed scenario. To guarantee "stress testing plausibility" in addition to statistical and economic soundness, we benchmark constrain the filtered model space (benchmark constrained BMA, BCBMA). To this end, we impose an additional restriction following Bonti et al. (2006):

5. the model-implied stress forecast for PDs does not deviate "too strongly" from a benchmark forecast. The benchmark forecast is derived from a Merton-Vasicek one-factor model, in which the systematic factor is set to a stress-level consistent with the estimated probability of occurrence of the macro scenario.[7] We set the "plausibility" bounds to two standard deviations of the historical PD time series around the Merton-Vasicek-implied PD-increase. We set this region to be generous enough to guarantee some dispersion in the predicted PD increases of the surviving ADL models. This reflects our approach of combining the information from both frameworks, instead of using only one. Appendix B provides details.

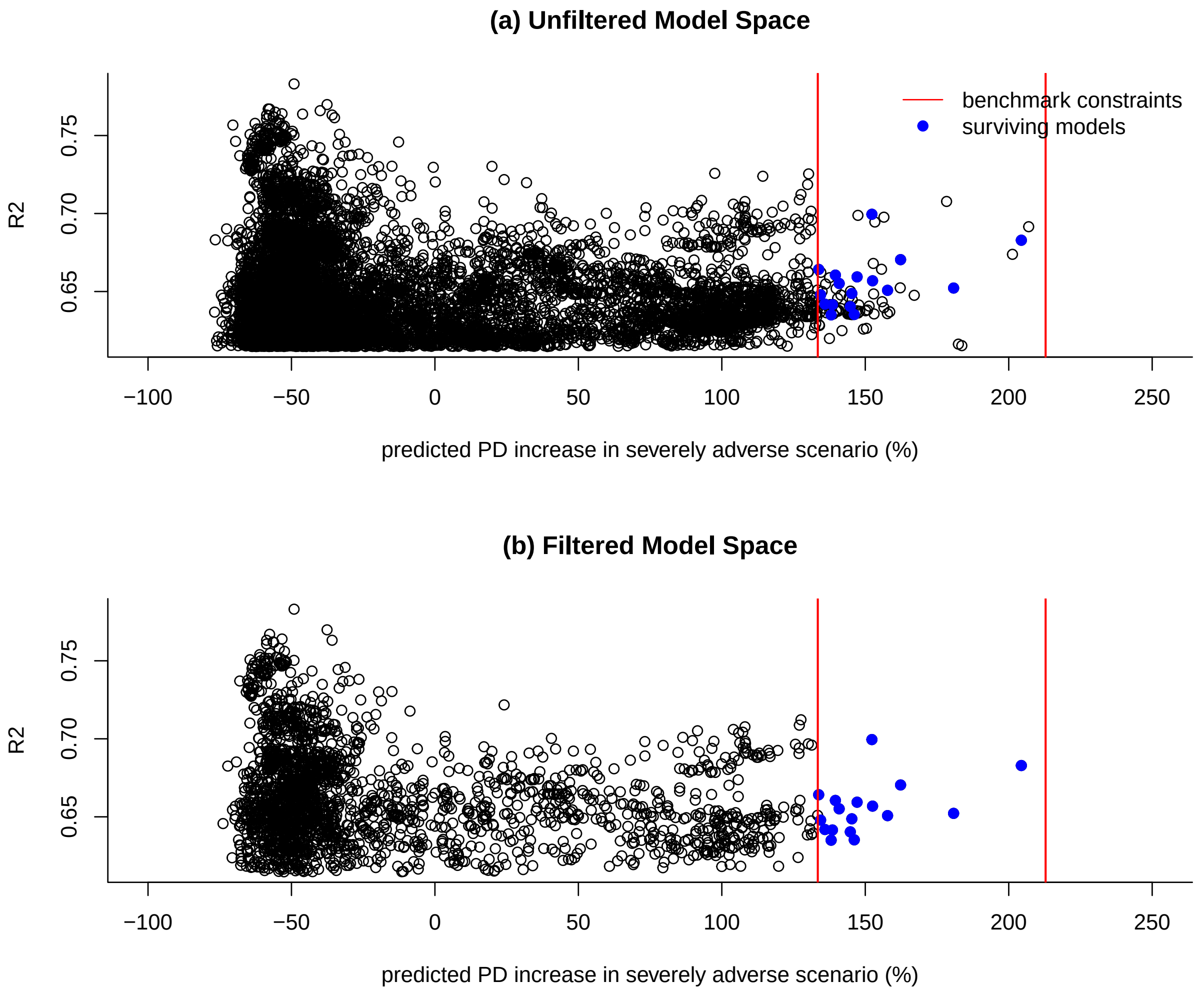
[5]When defining the model space we allow for a maximum of four covariates in each model ($nvmax = 4$), including lags and a maximum lag order of two ($L = K = 2$). This leaves us with $N = 26$ potential regressors. Thus, the entire model space features $\sum_{j=1}^{nvmax} \frac{N!}{j!(N-j)!} = 17,901$ different model setups. To economize on computation time we restrict our attention to a subset of the 10,000 best models according to the adjusted R^2 using the *leaps and bounds* algorithm (see Furnival and Wilson, 1974). Pelster and Vilsmeier (2017) provide additional details.

[6]See Appendix A for a detailed description of these filtering criteria. The "leave-one-out" principle implies that one repeatedly drops one observation from the sample, estimates the model using the remaining data and predicts the missing observation.

[7]The realization of the stressed macro variables is mapped to a realization of the systematic component in the one factor model that has the same probability of occurrence.

The advantage of using a benchmark constrained filtered model space lies in the fact that the benchmark is derived from a model that does not suffer from the same limitations as the estimated OLS models, such as scarce data or collinearity. Instead, the suggested quantile-mapping approach offers a simple alternative by just assuming a co-monotone relationship between the endogenous variable and at least one covariate (Bonti et al., 2006). By dropping all models from the model space, which predict stressed PD increases that lie outside a region deemed plausible by the non-OLS benchmark model, we attenuate the impact on the final BMA model of issues, which may not be filtered out by only considering OLS-type models. By assuming a co-monotone dependency between the PD dynamics and the macro scenario, zero stress-sensitivity always lies outside the benchmark constraints (for a more detailed discussion, see Siemsen and Vilsmeier, 2017).

Before any filtering criteria are applied, the unfiltered model space contains 10,000 different estimated models. Figure 3 shows the R^2 of all models in the unfiltered and filtered model space as a function of the scenario-implied average annual PD increase over the three-year horizon together with the quantile-mapping-implied benchmark constraints. We find a large dispersion of implied PD changes across the model space. Also, it shows that R^2 is not a good predictor of model plausibility, since a large percentage of models with high R^2 predicts a strong *decrease* in PDs over the stress scenario. Such non-intuitive model estimates are rather frequent in time series models based on historical PDs or default rates and are caused by correlated regressors (among the different macro-variables and also among the lags of the same variable) in combination with scarce data. In such cases, regressor coefficients may be non-unique, since variables catch up on effects of other variables. The benchmark constraints filter out such models. Of 10,000 estimated models in the unfiltered model space, 17 models survive the filtering and lie within the benchmark constraints.



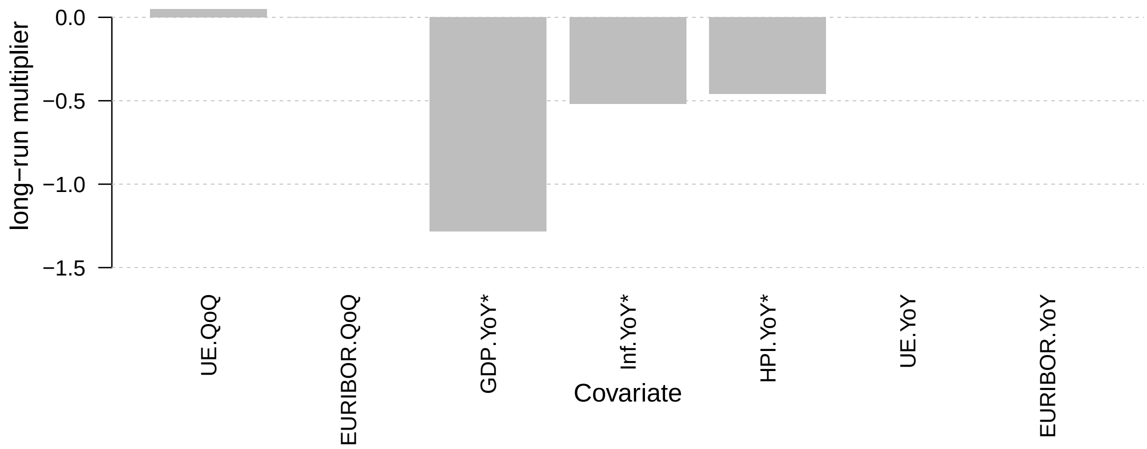
Notes: "Unfiltered Model Space" corresponds to the universe of all estimated ADL models. "Filtered Model" space corresponds to the model space post filtering criteria 1.-4. Colored dots indicate models that lie within the benchmark bounds.

Figure 3: Scenario Implied PD increases (average over 3-year stress-horizon) vs. R^2 of models in unfiltered and filtered model space

Figure 4 shows the normalized long-run multipliers of the BMA-combined final ADL specification. The long-run multiplier is defined as

$$\theta_i = \frac{\sigma_{X_i}}{\sigma_Y} \sum_{k=0}^{\infty} \frac{\partial EY_{t+k}}{\partial X_{i,t}} = \frac{\sigma_{X_i}}{\sigma_Y} \frac{\sum_{l=0}^{L} \beta_{i,l}}{1 - \sum_{m=1}^{M} \alpha_m},$$

with $Y = \Delta^4 \log(PD/(1 - PD))_t$ and can be interpreted as the increase in Y in terms of standard deviation if X_i is increased permanently by one standard deviation. All variables which are considered in our macro scenario (see Section 4.3) are selected by the BCBMA algorithm to be included in the final model. In the final specification, only contemporary dependencies and no lags are included. Given that macroeconomic variables tend to move sluggishly, this may be driven by the exclusion restriction (restriction 1), which prevents highly correlated covariates from being included in the same model. Except for the unemployment rate and 3M EURIBOR, all variables enter significantly according to the posterior inclusion probability (see Appendix A for a definition).



Notes: "QoQ"("YoY") denotes quarter-on-quarter (year-on-year) growth rate. All variables with an asterisk are significant according to the posterior inclusion probability. All variables which are featured in the macro scenario are included in the final model specification, although unemployment and EURIBOR enter insignificantly. The BCBMA algorithm includes only contemporary dependencies and no lags in the final model specification.

Figure 4: BCBMA-implied long-run multipliers

We combine the estimated coefficients of all models, M, remaining in the filtered and benchmark constrained model space to one model using each model's posterior inclusion probability

$$\beta_{BMA} = \sum_{i=1}^{M} \pi_i \beta_i \, ,$$

where β_i is the vector of coefficients of model i and π_i is the posterior inclusion probability, which can be interpreted as the probability of a model being the "best model" in the model space in repeated samples

$$Pr\left(M_i | Data\right) \equiv \pi_i = \frac{\exp(-0.5\Delta_i)}{\sum_{j=1}^{M} \exp(-0.5\Delta_j)} \, ,$$

with $\Delta_i = AIC_i - AIC_{min}$ and $AIC = n\log(\sigma_\varepsilon^2) + 2k$ (assuming $\varepsilon_t \sim N(\mu, \sigma_\varepsilon^2)$). We compute σ_ε^2 using the "leave-one-out" principle. The β_{BMA} vector is then used to derive the sector-aggregate PD dynamics over the stress horizon. Figure 8(a) shows the predicted scenario-implied evolution of the residential-mortgage aggregate sector PD.
To translate the stressed dynamics of sector s PDs to bank-specific PD dynamics for bank i, we apply a distance-to-default-transformation on starting values at $t = 0$:

$$PD_{i,t} = \Phi\left(\Phi^{-1}\left(PD_{i,0}\right) + \left[\Phi^{-1}\left(PD_{s,t}\right) - \Phi^{-1}\left(PD_{s,0}\right)\right]\right) ,$$

where Φ denotes the standard-normal cumulative distribution function and $PD_{i,t}$ is the bank-level PD computed as volume-weighted rating-class PDs.[8] Figure 6(a) shows the initial PD distribution at the bank level to which the BCBMA approach is applied. The median PD is 0.7% with a standard deviation of 0.4%.

3.2 LGD model

The LGD model maps the macro scenario to LGD dynamics. To this end, we first exploit the structural relationship between CLTV and house prices. Measured at market value, a decrease in house prices for a given loan volume translates directly to an increase in CLTV. In particular, for a bank i and rating class j the following recursion holds:

$$CLTV_{2016,i,j} = \frac{EaD^{pre}_{2016,i,j}}{C_{2016,i,j}}$$

$$CLTV_{t,i,j} = \frac{EaD^{pre}_{2016,i,j}}{C_{t,i,j}} \; , \forall t \geq 2017 \, ,$$

with

$$C_{2016,i,j} = EaD^{pre}_{2016,i,j} - EaD^{post}_{2016,i,j}$$

$$C_{t,i,j} = C_{t-1,i,j} \times \frac{HPI_t}{HPI_{t-1}} \; , \forall t \geq 2017 \, ,$$

with EaD^{pre} and EaD^{post} denoting exposure at default pre and post collateral, C denoting the collateral value and HPI denoting the house price index. Note that by keeping $EaD^{pre}_{2016,i,j}$ constant in the numerator of the CLTV, we impose a static balance sheet assumption that rules out amortizations during the stress horizon.
In the second step, we translate the CLTV to a corresponding LGD. We do so by using a reduced-form meta-dependency between CLTV and LGD based on international evidence.

[8] To ensure that banks' internal PD estimates for real estate exposures, which we use as starting values for the stress test, do not deviate "too strongly" from observed default rates, we benchmark reported one-year-ahead PDs with observed (long-run average) default rates from "exposures secured by real estate" (for fully secured residential mortgages) and "retail exposures" (for partially secured residential mortgages). If the reported, bank-level-aggregated PDs for fully and partially secured mortgages lie more than 20 % below historical default rates, we scale up reported PDs. For bank i and rating class j

$$PD_{i,j} = \Phi\left(\delta + \Phi^{-1}\left(\hat{PD}_{i,j}\right)\right)$$

with $\delta = \Phi^{-1}\left(PD^h_{k,i}\right) - \Phi^{-1}\left(\bar{PD}_i\right)$, with $\hat{PD}_j$ being the unadjusted PD, $PD^h_{k,i}$ being the historical (2014-2016) benchmark with $k \in \{$"secured by real estate", "retail" exposures$\}$ and $\bar{PD}$ being the bank-level aggregated real estate exposure PD. See Düllmann and Kick (2014) and Koziol, Schell, and Eckhardt (2015) for other stress tests, where initial PDs are based on bank-internal estimates instead of historical default rates.

Several studies argue that the CLTV is the single most important determinant of LGD (see, for example, Pennington-Cross, 2003; Calem and LaCour-Little, 2004; Qi and Yang, 2009, and the sources cited therein). While a more structural approach to LGD modeling offers the advantage of identifying the deep parameters that determinate LGD dynamics (see, for example, Gupton and Stein, 2002; Seidler and Jakubik, 2009), such determinants are often not necessary to compute expected and unexpected losses in a stress testing context. Moreover, calibrating a structural model often requires granular data such as debt type, seniority of the debt or debt market value after default, which may not be readily available (at the required level of granularity). Our reduced-form dependency remains silent about the particular nature of LGD determinants and simply exploits the correlation between CLTV and LGD found in different studies. Since it is not calibrated to any specific portfolio or jurisdiction, an application to different credit types or countries is feasible but requires additional adjustments and robustness checks.

For the meta-dependency, we employ two studies: Qi and Yang (2009) study the dependency between LGD and CLTV for the US using a granular loan-level data set from the *Mortgage Insurance Companies of America*, a trade association of six major US private mortgage companies. Their data covers a total of 241,293 mortgage insurance claims in the period 1990-2003. Qi and Yang bucket loans in seven CLTV classes: $CLTV \leq 0.80$, $0.80 < CLTV \leq 0.90$, $0.90 < CLTV \leq 0.95$, $0.95 < CLTV \leq 1.00$, $1.00 < CLTV \leq 1.10$, $1.10 < CLTV \leq 1.20$ and $CLTV > 1.20$. Thus, while they provide relatively dense information of the CLTV-LGD dependency for CLTVs between 80% and 120%, the information is more sparse in the higher and lower CLTV regions. In particular, they find that in the highest CLTV bucket ($> 120\%$), the mean LGD is 49.2%, far below the maximum possible value of 100%. Since, in a stress testing context, CLTV and thus LGD are likely to rise above historical heights, we are especially interested in the CLTV region that induces high LGDs. To this end, we draw on another study to find indications of the correlation between CLTV and LGD in high CLTV regions.

Palmroos (2016) uses Finnish mortgage data collected in a 2010 survey among all Finnish banks. The survey covers maturity structure at the loan level as well as LTV distributions. The dependency derived by Palmroos between CLTV and LGD covers the CLTV region between 0.60 and 20 and can therefore be used to shed light on CLTV and LGD correlation for high CLTV.[9] To derive our meta-dependency we split the CLTV space into four regions:

1. $0.85 \leq CLTV \leq 1.15$: the corresponding LGD values are directly taken from Qi and Yang (2009), Table 5.[10]

2. $1.15 < CLTV$: as discussed above in the high CLTV region, Qi and Yang (2009) provide only sparse information. We use the marginal rate of change between CLTV and LGD as shown in Palmroos (2016). The rate of change is anchored at CLTV=1.15

[9] We use the CLTV-LGD dependency derived if default can only occur for LGD>0. CLTV is derived via CLTV=1/collateral value.

[10] Assuming uniform distribution within each CLTV bucket, the CLTV is approximated by the mean bucket value. Values between buckets are linearly interpolated. To keep the dependency agnostic and due to a lack of corresponding information in our data set, we set all dummy variables in Qi and Yang (2009), except the relevant CLTV-bucket, to zero.

and iterated forward until CLTV=2.00, which corresponds to a LGD of about 1.0.[11]

3. $0 < CLTV < 0.85$: to derive the CLTV-LGD dependency in this region we draw on two pieces of information from Qi and Yang (2009): first, they show that overall the correlation between CLTV and LGD is 0.7 (see their Table 4) and, second, they show that in this CLTV region the mean LGD is 0.13 (see their Table 5).[12] We use these pieces of information to derive a linear dependency. We draw 10,000 synthetic normally distributed LGD series, which have a mean of 0.13 and a correlation with CLTV $\in [0, 0.85)$ of 0.7, and run a linear regression for each of these draws. We take the mean of the estimated coefficients.

We fit a polynomial of second order in regions 1 to 3 to filter any artifical jumps due to plugging regions 2 and 3 together.

4. $0 < CLTV < 0.10$: We floor the CLTV at 0.10, which corresponds to a LGD of about 0.07.[13]

Figure 5 shows the final meta-dependency between CLTV and LGD and Appendix F provides the corresponding data. The solid diamond shows average LTV and LGD of German banks in the EBA 2014 stress test for the "Retail – secured by real estate-of which: non-SME" portfolio. This data point for German significant institutions (SIs) seems to suggest that we can approximate the level of the CLTV-LGD dependency reasonably well.[14]
Given our granular LIRE data set, which features observed (long-run) LGD for retail and secured by real estate portfolios, we can anchor residential mortgage LGD starting values (pre-stress) at historical observations. Thus, we ensure that the level of the initial LGD in the residential mortgage portfolio, which we derive through the meta-dependency from the initial CLTV reported in the data set, does not deviate too much from observed LGDs in the "retail" portfolio (initial LTV > 0.8) or "secured by real estate" portfolio (initial LTV ≤ 0.8).[15] As a consequence, the level of the meta-dependency itself is not directly relevant for

[11] We use the rate of change instead of the level from Palmroos, since we find LGD to be relatively low in Finland compared to data for German banks that participated in the 2014 EU-wide stress test. This may be due to differences in bankruptcy law, culture or financial market regulation.

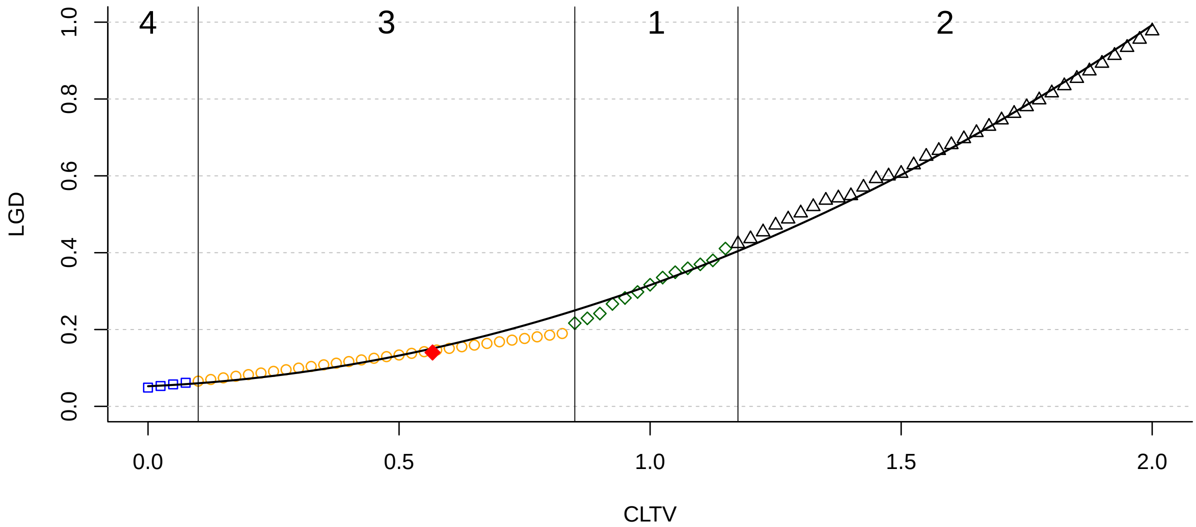
[12] Again, to keep the dependency agnostic and due to a lack of corresponding information in our data set we set all dummy variables except the relevant CLTV-bucket to zero.

[13] CRR Art. 164(4) prescribes a minimum LGD of 0.10 for retail exposures secured by real estate. Given our meta-dependency, this would correspond to a CLTV about about 0.30. Since such low CLTVs are not commonly observed, especially not under stressed conditions, the floor has no effect on stress results and is imposed only to account for potential data issues.

[14] The LGD of SIs participating in the EBA stress test does not necessarily have to be representative of the smaller LSIs participating in the LIRE survey. However, in the EBA 2014 stress test German SIs report an average LGD of 0.14, which lies very close to the average loss rate reported by LSIs of 0.12 (average 2014-2016).

[15] The LIRE survey data set features LTV only for partially secured exposures (LTV>1.0). We anchor the initial LGD for these positions by computing the meta-dependency-implied LGD and compare it to the historical retail LGD. If it deviates by more than 50% above or below, we replace the derived LGD with the historical counterpart. For fully secured positions no initial LTV is reported in the LIRE survey data. Therefore, we use the historical LGD in the secured by real estate portfolio directly as a starting point and map it to an initial LTV through the meta-dependency.

the computation of stress effects, since we always start off from a LGD that is anchored by the bank-specific loss history. We only exploit the estimated slope which maps percentage increases in CLTV during stress to percentage increases in LGD.[16]

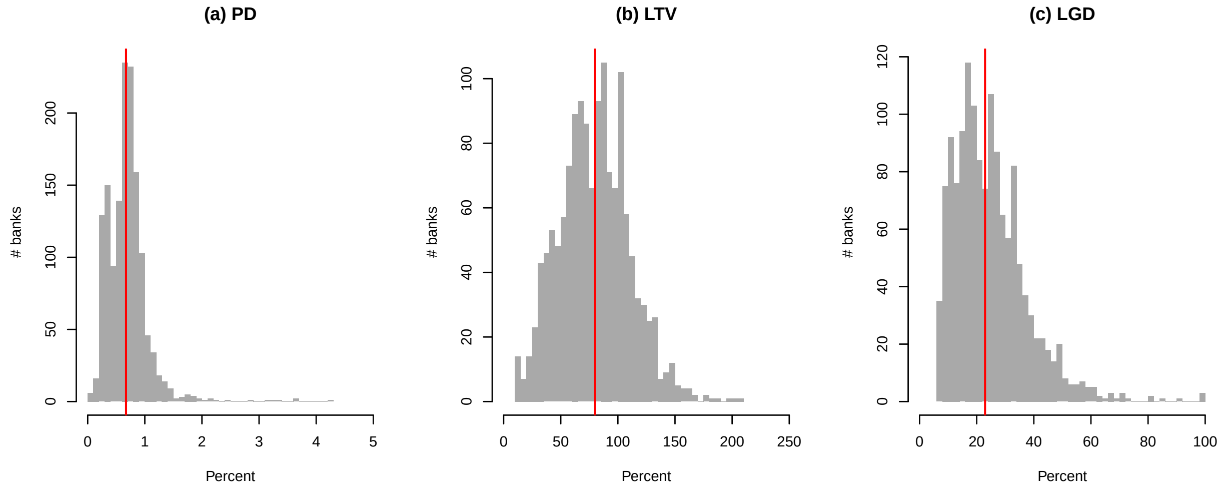


Notes: The solid diamond indicates average LTV and LGD of German banks in the EBA 2014 stress test for the "Retail – secured by real estate-of which: non-SME" portfolio. Numbers 1 to 4 indicate CLTV regions in which different methods of deriving the dependency are used. See main text for explanations.

Figure 5: Meta dependency between LGD and LTV

This completes the description of the two main model parts of our stress testing framework. Figure 6(b) shows the initial LTV distribution at the bank level and Figure 6(c) the corresponding initial LGD distribution derived from the meta-dependency. The median LGD is 22.9% with a standard deviation of 12.8pp.

[16]Note, moreover, that benchmarking non-market-value-based LTVs against 2014-2016 average realizations ensures that stress test starting values reflect recent market conditions.



Notes: Distributions show bank-level aggregated risk parameters. Vertical line indicates median.

Figure 6: Distribution of initial risk parameters

3.3 Stress impact calculation

The impact of the stressed risk parameters on banks' CET1 ratios follows the EBA credit risk methodology and is fourfold. First, elevated PDs and LGDs increase default flows and thus impairments on non-defaulted assets:

$$\begin{aligned}
defaultflow_{t+1} &= EaD_t \times PDpit_{t+1} \\
EaD_{t+1} &= EaD_t - defaultflow_{t+1} \\
impflow^{new}_{t+1} &= LGDpit_{t+1} \times defaultflow_{t+1} \,,
\end{aligned}$$

where *pit* indicates "point-in-time parameters" in contrast to "through-the-cycle" (*ttc*) parameters, EaD denotes exposure at default and $impflow$ denotes impairment flows. Second, a further deterioration in collateral quality during stress, reflected in increasing LGD, requires banks to also increase impairments on assets that already defaulted:

$$\begin{aligned}
defaultstock^{2016} &= \sum EaD^{2016}_{PD=1} \\
defaultstock_{t+1} &= defaultstock_t + defaultflow_{t+1} \\
impflow^{old}_{t+1} &= \max\{0, LGDpit_{t+1} \times defaultstock_t - provstock_t\} \\
provstock_{t+1} &= provstock_t + impflow^{old}_{t+1} + impflow^{new}_{t+1} \,.
\end{aligned}$$

where *provstock* denotes the provision stock. Note that we approximate the initial default stock by the exposure amount that was assigned by banks to the PD=100% rating class. Third, we assume that defaulted assets do not pay interest. Thus, elevated default flows induce foregone interest payments. From the LIRE survey, we know the interest rate on new business for real estate mortgage lending from 2014-2016. We assume that the interest rate remains constant during the stress horizon at the average 2014-2016 level $\bar{R}$. Thus, stressed interest income, *intinc*, is given by

$$intinc_{t+1} = \bar{R} \times [EaD_t - defaultflow_{t+1}] = \bar{R} \times EaD_{t+1}\,.$$

Fourth, we model RWA dynamics. We pursue a "pseudo-IRB" approach that applies the Basel IRB formula to calculate risk weights depending on PD^{ttc} and LGD^{ttc}, $\mathcal{F}$, to all banks.[17] This modeling choice implies that we consider expected as well as unexpected losses that accumulate during the stress horizon. The IRB formula is designed to calculate RWA as a buffer against unexpected losses as measured by the value at risk in excess of the expected loss (see Basel Committee on Banking Supervision, 2005). Under stressed conditions, both the mean of the loss distribution (expected losses), as well as the variance (unexpected losses) increase and have to be accounted for by thorough risk management. In that sense we pursue an economic in contrast to a purely regulatory view of stress losses. Below, we study how stress test results are affected by RWA modeling. To this end, in Section 5.2 we contrast the benchmark "pseudo IRB" approach with the regulatory SA as stated in CRR Art. 125 and a more risk-sensitive SA revision as proposed in Basel Committee on Banking Supervision (2015).

$$\begin{aligned}\Delta RWA_{t+1} &= \mathcal{F}^{Basel}\left[PD^{ttc}_{t+1}, LGD^{ttc}_{t+1}\right] - \mathcal{F}^{Basel}\left[PD^{ttc}_{t}, LGD^{ttc}_{t}\right]\\ RWA_{t+1} &= RWA_t + \Delta RWA_{t+1}\,.\end{aligned}$$

To compute the through-the-cycle adjustments we apply a pragmatic adjustment algorithm outlined in Appendix C. These four effects are then mapped to banks' profits, π^{Stress}_{t+1} through

$$\begin{aligned}\pi^{Stress}_{t+1} = \pi^{Plan}_{t+1} - intinc^{Plan}_{t+1} + impflow^{Plan}_{t+1} + intinc^{Stress}_{t+1}\\ - impflow^{new,Stress}_{t+1} - impflow^{old,Stress}_{t+1}\,,\end{aligned}$$

i.e. the counterfactual stress profit is computed as the profit as assumed in banks' planning data in the corresponding year of the stress horizon adjusted for stressed interest income and stressed impairments. Finally, we can compute the stressed CET1 ratio

[17] Most banks participating in the LIRE survey follow the credit standardized approach. We model RWA dynamics for those banks using the "pseudo-IRB" approach as well, since we are interested in continuous and economically relevant unexpected loss dynamics during the stress horizon. Under the SA, RWA dynamics occur when loan splitting is applied between the "secured by real estate" exposure class (with a risk weight of 35%) and the regulatory non-preferred exposure class (with a risk weight of 75%). The threshold LTV between these two classes is 80%. According to Art. 208(3)a CRR residential property collateral has to be revalued at least every three years, such that, even in the absence of loan splitting, LTVs are not constant during the lifetime of a mortgage, but likely to move only sluggishly. A more frequent revaluation is necessary if market conditions are subject to "significant changes". The Standardized Approach has been criticized as being too risk-insensitive and the Basel Committee on Banking Supervision (2015) proposed a revision to increase risk sensitivity. Section 5.2 shows stress test results for the SA.

$$CET1_{t+1} = CET1_t + \pi_{t+1}^{Stress}$$
$$CET1r_{t+1} = \frac{CET1_{t+1}}{RWA_{t+1}}.$$

4 Data

4.1 LIRE Survey

The stress testing framework draws upon a unique and granular data set collected through the LIRE survey among 1,500 German LSIs. These banks account for 88% of all German banks and 41% total assets. This survey has been conducted by the Bundesbank in cooperation with BaFin on a biannual basis since 2013. For the first time, the 2017 survey (with data as of 31/12/2016) includes a template on real estate mortgage lending. This template provides granular data on mortgage interest rates, as well as PDs, collateral and exposures on a rating class level.[18] These data are used to derive bank-specific starting values for credit risk parameters and exposures. Appendix D shows the design of the residential mortgage template.

Besides data on bank-individual starting values for PDs, collateral values and exposure, the LIRE survey provides further information which we use in our stress test exercise: *first*, it features bank-internal planning data spanning the entire stress horizon from 2016 to 2019. Therefrom, we can compute all CET1 changes anticipated by each bank's internal planning (interpreted as the baseline scenario). This allows us to determine the CET1 ratio level post residential mortgage stress as the difference between banks' anticipated 2019 CET1 ratios and the scenario-induced stress effect. With these values, we can assess banks' regulatory CET1 ratios under an isolated stress to their mortgage portfolios, keeping the remainder of banks' balance sheets in line with banks' expectations. *Second*, reported values on bank-specific long-run average historical default and loss rates in the mortgage portfolios allow us to check the conservatism of banks' reported starting values (and to adjust them if necessary).

4.2 German Credit Register

To derive a time series for residential mortgage sector PDs used in the BCBMA model, we draw on the German Credit Register (GCR). The GCR features quarterly observations since 2008Q1 on all loans from a bank to a given borrower of more than €1 million. For each IRB loan the regulatory PD is reported. We identify residential mortgage loans as loans with a residential mortgage loan share above 90%.[19] This gives us a total coverage for all

[18]Note that collateral values in the survey are not based on market values. While values based on market prices may be preferable, we are not aware of any data source that provides such data for Germany at the same level of granularity as the LIRE survey. Note, moreover, that according to Art. 208(3)a CRR residential property collateral has to be revalued at least every three years and more frequently if market conditions change significantly.

[19]This assumption is required as PDs are reported only at the loan level, but not at the sublevels *"of which: collateralized loans"* and *"of which: collateralized loans, of which: residential mortgage loans"*.

residential mortgage loans of about 10%.[20] The corresponding PDs are aggregated to a real-estate-sector PD by volume-weights, which is then linked to the macro scenario. Figure 8(a) shows the historical time series from 2008Q1 to 2016Q4 together with the BCBMA-implied stress dynamics.

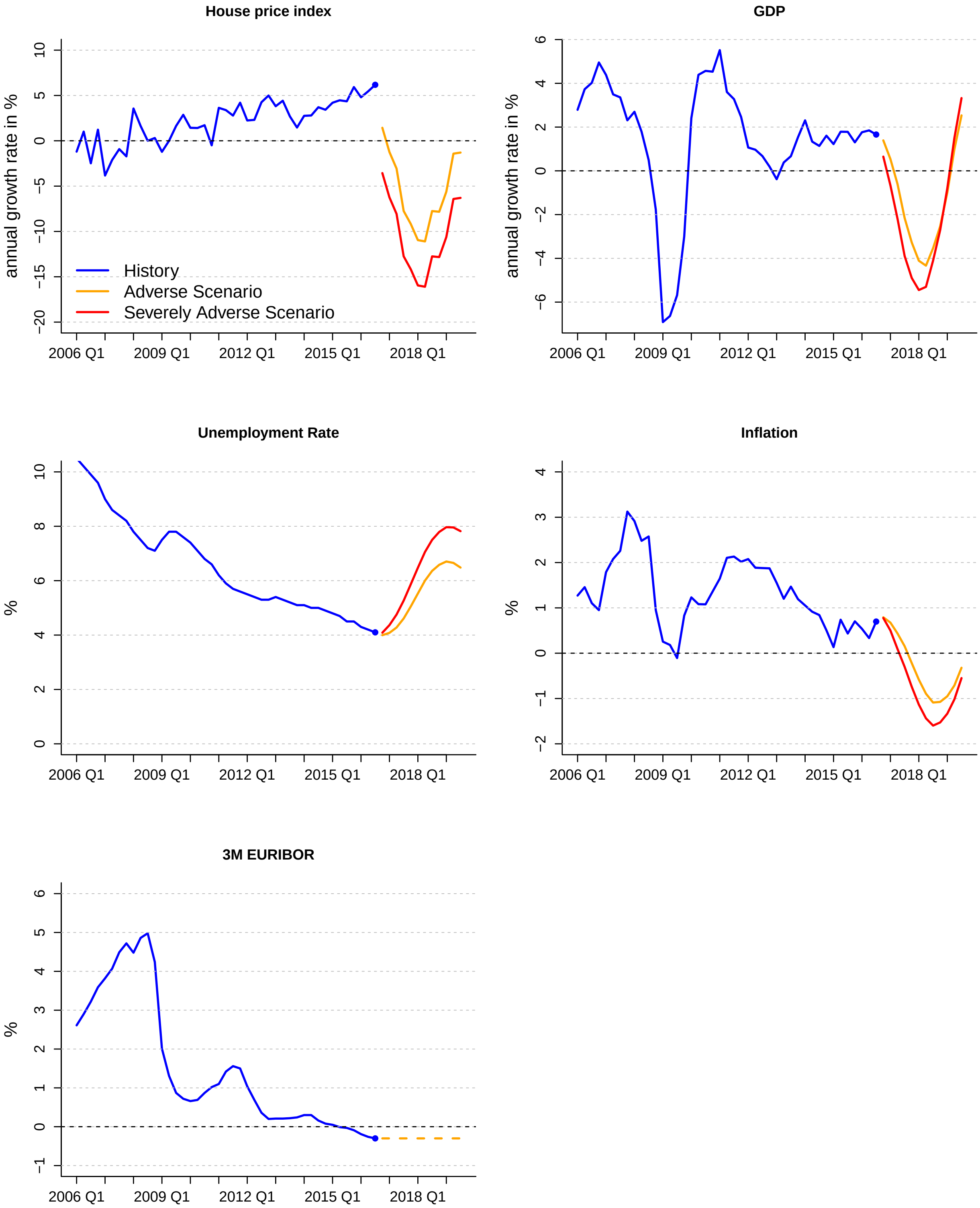
4.3 Macro Scenarios

As discussed in Section 2, current Bundesbank estimates indicate potential overvaluations in German cities of about 15% to 30%. We study the impact of a corrective movement in house prices of this magnitude on German LSIs' balance sheets. To derive adverse yet plausible scenarios for the drop in real estate prices, we use the dynamics observed in the Spanish housing market in the period 2011-2013. In this period, real estate prices dropped by about 30% peak to trough (*severely adverse scenario*). To study the sensitivity of German LSIs to changes in house prices we also consider a second less extreme scenario with a 20% house price drop peak-to-trough, which we derive through a parallel shift in annual growth rates of the house price index (*adverse scenario*).
For our stress test exercise, we consider a horizon of three years, ranging from 2017Q1 to 2019Q4. To derive macro dynamics consistent with the assumed drop in real estate prices, we use a standard VARX model with house prices and 3M EURIBOR as exogenous variables.[21] The endogenous variables include stationary time series for quarterly year-on-year GDP growth, quarterly year-on-year changes in the unemployment rate and the annual inflation rate. To control for the monetary policy stance, we include the level of the 3M EURIBOR interest rate, despite the fact that the EURIBOR level is not chosen by the BCBMA algorithm to be included in the final PD model (see Figure 4). The EURIBOR is kept constant during the stress horizon at its 2016Q4 level, assuming exogenous and constant monetary policy.[22] All data is taken from Eurostat, ranges from 2005Q1 to 2016Q4 and is seasonally adjusted, except for the house price index and the EURIBOR rate. Figure 7 shows the paths for the macro variables under the adverse and the severely adverse scenarios. Appendix E puts the assumed macro dynamics in historical perspective, showing how the assumed house price decline and corresponding GDP and unemployment dynamics compare to international experience.

[20] There is a possible concern that deriving the PD time series based on large loans from IRB banks may not be representative for the small to medium-sized banks participating in the LIRE survey. In this context it is worthwhile noting that the sector PD time series is used only to derive stress-implied relative PD increases which are applied to bank-specific initial PDs reported in the LIRE survey. That said, the sector PD is about 0.54% in 2016Q4, which lies close to the median residential mortgage PD of 0.67% reported by banks in the LIRE survey.

[21] Due to the short time-series we include two lags of endogenous and exogenous variables.

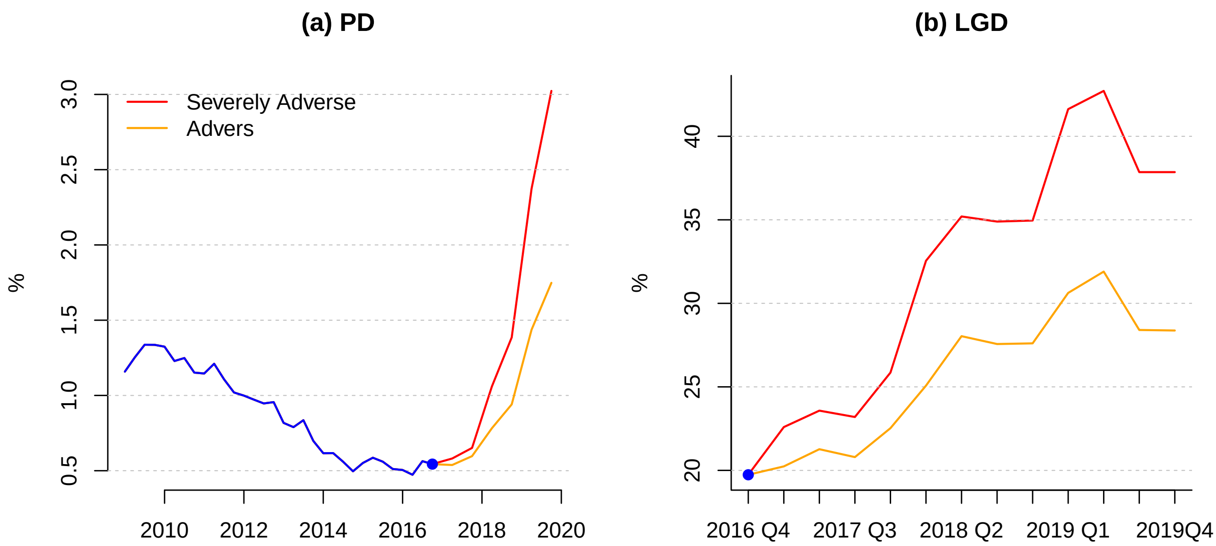
[22] Note that the EURIBOR interest rate is only kept constant during the stress horizon. Both the VARX and the BCBMA model use historical realizations for estimation. Detailed estimation results are available from the authors upon request.



Notes: All historical data from Eurostat. Stress dynamics are derived from VARX(2,2) using the house price index and EURIBOR as exogenous variables.

Figure 7: Macro Scenarios

Finally, Figure 8 shows the stressed credit risk parameter dynamics induced by the assumed macro scenario. We find that in the adverse (severely adverse) scenario the residential mortgage sector PD increases over the three-year stress horizon by about 206% (461%). Figure 8(b) shows the implied LGD dynamics for an initial LTV of 0.80. We find that, in the adverse (severely adverse) scenario, this initial LGD increases by 44% (92%).



Notes: Historical sector PD derived from GCR from 2008Q1 to 2016Q4. Panel (a): stressed PD dynamics as implied by BCBMA model. Panel (b): stressed LGD dynamics for initial LTV of 0.8 as implied by meta-dependency.

Figure 8: Scenario-implied residential mortgage sector risk parameter increases

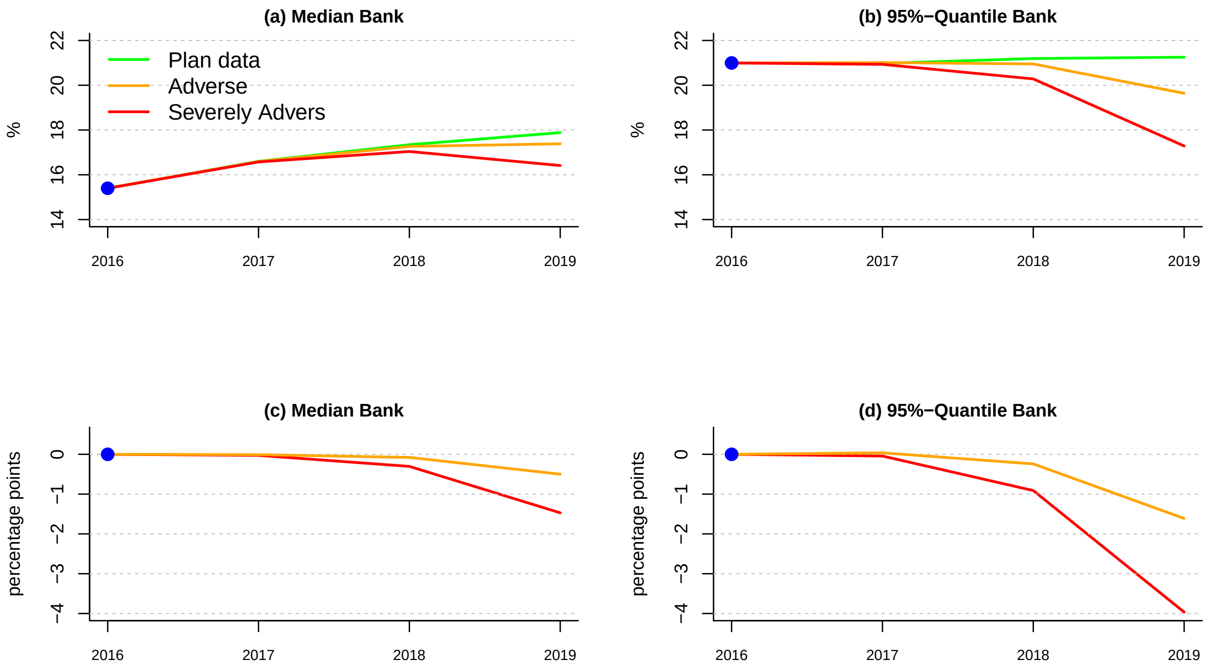
5 Quantitative Results

5.1 Stress test results

The stress test follows a top-down approach and has a purely microprudential focus.[23] Only the effects of the scenarios on residential mortgage portfolios of banks are considered, and only impairments, interest income and RWA effects within these portfolios are captured. PD and LGD dynamics in other portfolios and potential contagion effects between risk types, asset classes and banks are outside the scope of this stress test exercise. Changes in the impairments and RWA that are anticipated by the banks over the stress horizon (2017-2019) and which are independent of the scenarios are separately accounted for as "plan data effects". This means that all stress effects are to be interpreted as additional CET1 losses due to residential mortgage portfolios.

[23] In contrast to a "bottom-up" approach, where the stress calculations are performed by the banks themselves.

Figure 9 shows the stress dynamics in CET1 ratios for the bank with the median losses in the severely adverse scenario, as well as for the 95% quantile bank. Panels (a) and (b) show the CET1 ratio levels relative to banks' internal planning data. Panels (c)-(d) show the drop in the CET1 ratio relative to the 2016 initial level.



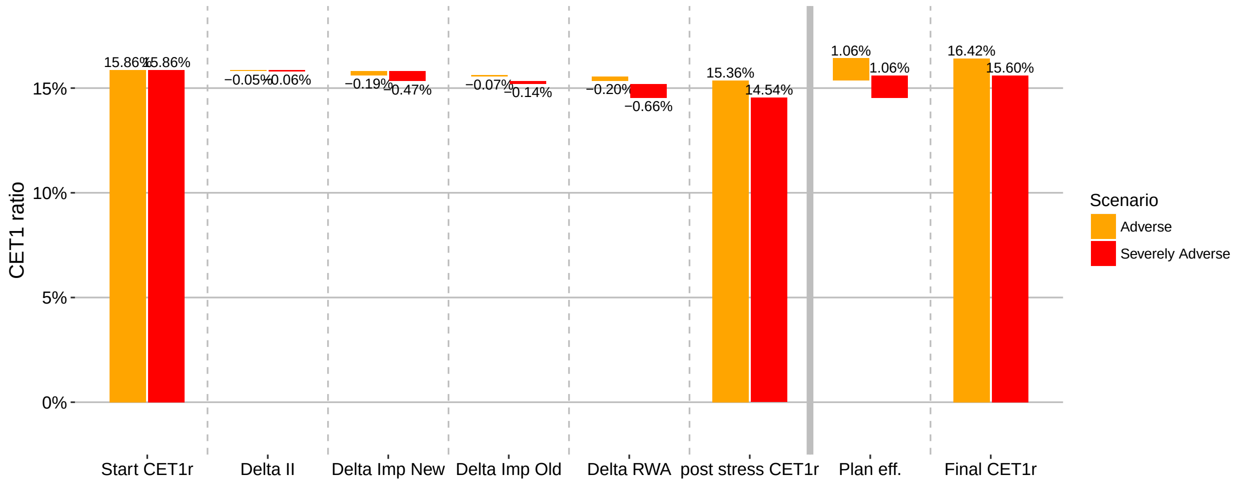
Notes: Top row: stress-implied CET1 ratio level relative to banks' internal planning. Bottom row: stress-implied CET1 ratio changes relative to initial 2016Q4 CET1 ratio.

Figure 9: Stressed CET1 ratio dynamics

In the adverse (severely adverse) scenario the CET1 ratio of the median bank decreases by 0.55pp (1.47pp) relative to 2016 (panel c). These additional CET1 losses reduce the CET1 ratio relative to the level anticipated in banks' plan data. In 2019, the framework predicts a CET1 ratio of 17.4% (16.4%) post stress for the median bank (panel a). The values include banks' internal predictions for CET1 growth over the three-year horizon from 2017-2019, which anticipate a substantial growth of more than 2pp.
The stress impact is highly heterogeneous across banks. About 12% of the banks experience a decrease by more than 3pp in the severely adverse scenario. The maximum decline is 11.1pp. The standard deviation of the CET1 ratio is 0.5pp (1.1pp) in the adverse (severely adverse) scenario. The heterogeneity of stress impacts becomes evident if one looks at the 95% quantile bank. For this bank the scenario-induced change in the CET1 ratio relative to 2016 is 1.6pp (4.0pp) in the adverse (severely adverse) scenario.
Comparing the two scenarios, we find a strong non-linear impact of real estate mortgage stress on banks' CET1 ratios. Relative to the adverse scenario, the drop in real estate prices in the severely adverse scenario is about 50% stronger. However, we find that the aggregate

(composite bank) CET1 ratio drops 1.65 times more than under the adverse scenario. This non-linearity is driven by the non-linear increase in PD and LGD (see Figure 8) and the non-linear IRB formula used for computing RWA. The non-linear increase in LGD is thus affected by the non-linear dependency between CLTV and LGD as suggested by the meta-dependency (see Figure 5).

Besides considering median results, it is also interesting to elaborate on aggregate stress effects. To this end, Figure 10 shows the composition of stress effects on the CET1 ratio of the composite bank, i.e. the sum of all participating banks. We find that the main stress drivers are RWA (40% in the adverse scenario, 50% in the severely adverse scenario) and impairments on newly defaulted loans (38% and 35%). Also, the effect of impairments on the default stock is, at 13% and 10%, considerable. The figure shows the plan data effect on banks' CET1 ratios separately. The composite bank predicts an increase in its CET1 ratio by 1.1pp over the three-year horizon.



Notes: "Delta II": effect due to interest income. "Delta Imp New": effect due to impairments on newly defaulted loans. "Delta Imp Old": effect due to impairments on default stock'. "Delta RWA": effect due to RWA. "Plan eff.": CET1 ratio growth over stress horizon according to banks' internal planning.

Figure 10: Stress effect composition

To summarize, we find that German LSIs, while being mostly sufficiently capitalized, are susceptible to a corrective movement in house prices. If banks were to raise their CET1 ratio post severely adverse stress back to the initial 2016 level, they would on aggregate require €19 billion, which corresponds to 7.3% of aggregate 2016 CET1 capital. It should be borne in mind that such expected and unexpected losses are due solely to exposures to residential mortgages. If contagion effects between portfolios, risk types and banks were additionally to be taken into consideration, losses would probably be much higher.

5.2 The impact of RWA modeling on stress test results

The stress test results show that RWA dynamics are a major determinent of stress outcomes. Under the benchmark “pseudo-IRB” approach we model RWA dynamics using the BCBS formula for IRB banks, with through-the-cycle-PDs and LGDs as main inputs. Despite the fact that the vast majority of banks participating in this stress test use the SA, we prefer this approach since it allows for economic relevant and continuous dynamics in unexpected losses. Under the SA, RWA dynamics are induced if (among others) the LTV passes an 80% threshold (CRR Art. 125(2)d) and/or loan splitting is applied.[24] For real estate-secured exposures with an LTV below 80% (among others) a preferred risk weight of 35% can be used; otherwise the risk weight is 75%. This approach is generally considered to feature only limited risk sensitivity and thus no economically relevant modeling of unexpected losses. As discussed in Basel Committee on Banking Supervision (2005), IRB risk weights are computed to create a buffer of capital in excess of *expected* losses in order to insure banks against *unexpected* losses. Unexpected losses can be measured by the value at risk (or expected shortfall), which provides an upper bound of losses that will occur in a given time frame with a certain probability. Since in a stressed condition both the mean and the variance of the loss distribution increase, quantifying total losses requires measuring both expected losses (impairments) and unexpected losses (RWA). From this point of view, the SA approach may be considered as a simplified IRB approach which only accounts for sluggish, discrete movements in unexpected losses, thus pursuing a regulatory rather than an economic approach to RWA modeling. In Basel Committee on Banking Supervision (2015), a revision to the SA is suggested (“SA revision”) which would increase risk sensitivity by using six instead of two RWA buckets for different LTVs. Table 1 shows risk weights for the SA and for the SA revision.

Table 1: SA risk weights

LTV (%)	SA risk weights (%)	SA revision risk weights (%)
[0,40)	35	25
[40,60)	35	30
[60,80)	35	35
[80,90)	75	45
[90,100)	75	55
[100,Inf)	75	75

Notes: SA risk weights according to Art. 125(2)d CRR. SA revision weights according to Basel Committee on Banking Supervision (2015).

[24]Loan splitting implies that a loan with LTV>0.8 is split up into two loans: one loan with LTV 0.8 and the remainder, thereby reducing RWA. For example, assume a loan with volume 1000 and LTV=1. This loan would induce RWA requirements according to the SA of $0.75 \times 1000 = 750$. Splitting this loan up would create a loan of volume 800 with LTV 0.8 and an unsecured loan of 200, implying RWA requirements of $0.35 \times 800 + 0.75 \times 200 = 430$. For the SA revision one can show that loan splitting is only optimal for loans with LTV >0.8. In this case the split is done between the LTV=0.8-bucket and the LTV≥1-bucket, as the LTV=0.8-bucket optimizes the trade-off between risk-weight and collateralizable volume. See also footnote 17.

We study how stress test results differ across different RWA models. Table 2 shows stress test moments for the benchmark "pseudo-IRB" approach as well as for the SA and the SA revision. Note that in all cases the initial RWA are the same 2016Q4 values as submitted in the LIRE survey templates; only the modeling of ΔRWA_{t+1} is affected (see Section 3.3). We find that, under the SA, stress effects are reduced by 18% (33%) in the adverse (severely adverse) scenario. This can be traced back to a substantially lower increase in RWA during the stress horizon. While, under the "pseudo-IRB" approach, RWA increase by 1.3% (4.6%) in the adverse (severely adverse) scenario relative to 2016, the increase in RWA under the SA is by 46% (70%) lower.

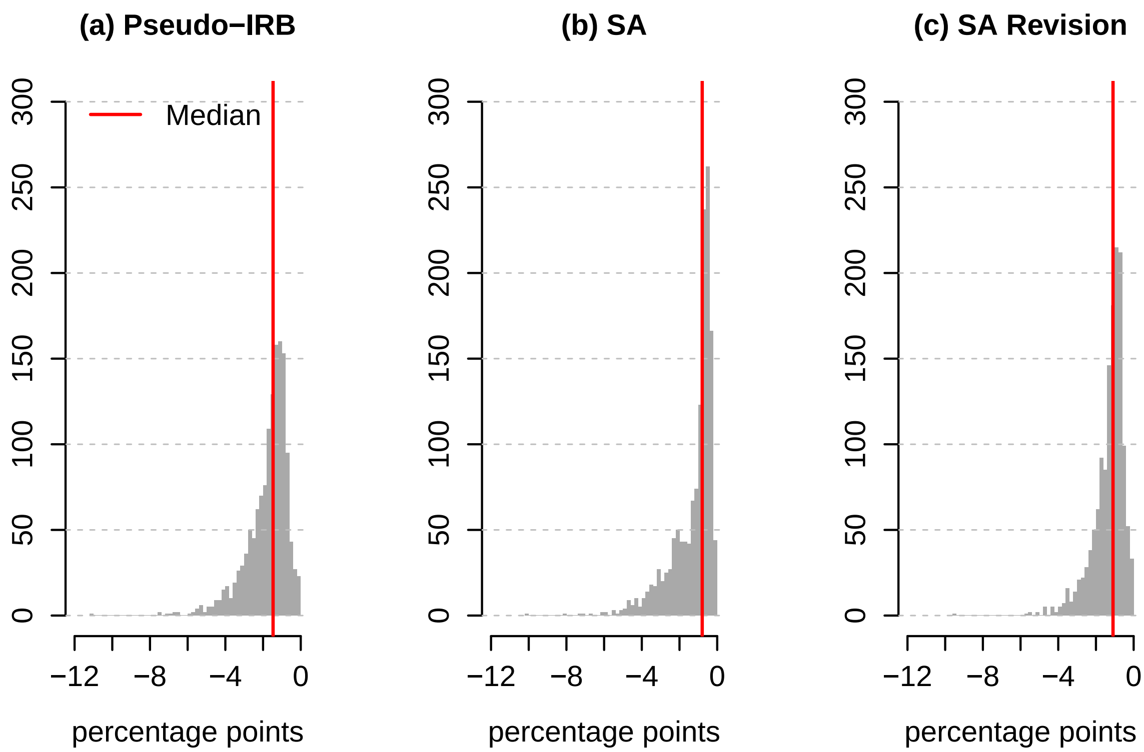
Looking at the stress test outcomes under the SA revision, we find that, under this approach, stress test outcomes move towards those of the "pseudo-IRB" approach. In the adverse (severely adverse) scenario, SA revision results lie 14% (20%) above those of the SA. The increase in RWA almost doubles across both scenarios. Under the adverse scenario, SA revision results lie very close to the "pseudo-IRB" results. This analysis highlights the important role of RWA modeling in stress test frameworks. Benchmarked against the "pseudo IRB" approach which is explicitly designed to approximate unexpected losses, employing a less risk-sensitive RWA model to approximate unexpected losses may induce substantially attenuated stress effects.

Figure 11 shows the CET1 ratio impact of RWA increases under different RWA regimes. Interestingly, we find that compared to the IRB approach (panel a) the SA (panel b) has a very similar stress distribution in the left-hand tail of the CET1 loss distribution (the 95% quantile of the loss distribution under the "pseudo-IRB" approach is 4.0pp, under the SA it is 3.7pp). However, the loss distribution features a higher mass of banks that experience low stress impact (the 5% quantile is 0.50pp under the "pseudo-IRB" approach, under the SA it is 0.25pp). This can be traced back to the sluggish RWA dynamics under the SA, which are only triggered once the CLTV rises above 80%. Thus, banks that start off with low CLTV and do not pass the 80% threshold (and thus do not apply loan splitting), within the three-year horizon, will not incur any RWA dynamics. Since these are the banks that will also experience low expected losses due to low CLTV, the SA induces a clustering in the right-hand tail of the distribution, which can induce an underestimate of aggregate economic losses. Under the SA revision (panel c), we find that the increased risk sensitivity moves the loss distribution closer to the IRB approach. This tends to support the view that a more risk-sensitive SA, benchmarked against the IRB approach, provides a better approximation of unexpected losses during the stress horizon and is less likely to underestimate economic losses.

Table 2: RWA effects

	IRB	SA	SA Revision
	Adverse scenario		
Aggregate stress effect (pp)	-0.50	-0.41	-0.49
Contribution ΔRWA (%)	39.8	26.3	39.3
RWA change (%)	1.3	0.7	1.2
	Severly adverse scenario		
Aggregate stress effect (pp)	-1.32	-0.88	-1.00
Contribution ΔRWA (%)	50.1	24.6	33.9
RWA change (%)	4.6	1.4	2.3

Notes: Stress effect measured in pp of CET1 ratio. "Contribution Δ RWA" corresponds to percentage share of total stress effect attributed to changes in RWA over the stress horizon. "RWA change" corresponds to percentage increase of RWA relative to 2016 within three-year stress horizon.



Notes: SA risk weights according to Art. 125(2)d CRR. SA revision weights according to Basel Committee on Banking Supervision (2015).

Figure 11: Distribution of stress effect depending on RWA model in the severely adverse scenario

6 Conclusion

The paper proposes a real estate mortgage stress testing framework for German LSIs. The framework combines a unique data set from the 2017 LIRE survey conducted among all 1500 German LSIs with a benchmark-constrained Bayesian model averaging approach to derive stressed risk parameter dynamics. Its main contributions are fourfold: *first*, the granular data set grants unique insights into German LSIs real estate mortgage exposures and credit risk parameters. *Second*, a benchmark-constrained BMA approach is used to map dynamics from a macro scenario to PDs, thereby reducing model uncertainty in a fragile statistical environment. *Third*, we derive a traceable reduced-form dependency between CLTV and LGD, which allows us to translate stressed CLTV dynamics to LGD dynamics and vice versa. And *fourth*, we elaborate on the effect of RWA modeling on stress test outcomes, contrasting a "pseudo-IRB" approach with the SA and the SA revision suggested in Basel Committee on Banking Supervision (2015).

We find that German LSIs are mostly well equipped to withstand a serve decline in house prices. However, stress test results are heterogeneous at the bank level. Following a house price drop by 30%, the median bank suffers a reduction in its CET1 ratio of 1.5pp. Raising the CET1 ratio back to a pre-stress level would on aggregate require about €19 billion, 7.3% of initial CET1 capital. We also show that using a regulatory RWA model that does not account for continuous dynamics in unexpected losses during the stress horizon can substantially attenuate stress effects. Comparing the "pseudo-IRB" approach with SA, we find that the latter induces stress effects which are up to 33% lower.

A Model space filtering criteria

When computing and filtering the model space as outlined in Section 3.1 we proceed as follows.

1. Due to computational constraints we have to define bounds of the model space. The bounds are

   - a maximum of 4 covariates in each model, including lags,
   - a maximum lag order of four, and
   - no lag for the endogenous variable[25]

2. We estimate 10,000 specifications for each set of models with the same number of covariates.

3. Of these models we consider only the 10,000 models with the highest adjusted R^2. We call these 10,000 models the "unfiltered model space".

4. We filter the unfiltered model space according to a set of statistical and economic criteria:

   - We allow only those covariates to be in the same model which have a pairwise correlation below 0.80 to limit imperfect collinearity, which may induce imprecise estimation.
   - We filter for models for which the Durbin-Watson test cannot be rejected at a 1% confidence level.
   - We filter for models with estimated coefficients that satisfy the sign restrictions in Table A.1.
   - We filter for models that do not have a significantly worse out-of-sample prediction performance than the best model in the model space according to the posterior model probability. Let $Pr(M_i|D)$ denote the probability that model M_i is the best model in repeated samples given data D. Then

     $$Pr(M_i|D) = \frac{exp(-0.5\Delta_i)}{\sum_{j=1}^{M} exp(-0.5\Delta_j)},$$

     where $\Delta_i = AIC_i - min(AIC_i)_{i=1}^{M}$ and, under the assumption that $\varepsilon_t \sim N(\mu, \sigma_\varepsilon^2)$, $AIC = n\log(\sigma_\varepsilon^2) + 2k$ with n denoting the number of observations and k the number of regressors. To compute σ_ε^2 we use the "leave-one-out" principle, such that $\hat{\sigma}_\varepsilon^2 = \frac{1}{n}\sum_{i=1}^{T}(\hat{y}_i - y_i)^2$, with $\hat{y}_i$ is computed by dropping y_i from the sample and predicting it, using the remaining observations to estimate the model. We drop all models with $Pr(M_I|D) < max[Pr(M_i|D)]/100$ from the model space.

[25] The quantitative results do not change when extending the model space to allow for more covariates or additional lags.

Table A.1: RWA effects in severely adverse scenario

Variable	sign restriction
HPI	-
GDP	-
Unemployment Rate	+
Inflation	0
3M EURIBOR	0

Notes: "+" denotes a positive coefficient sign restriction, "-" denotes a negative coefficient sign restriction and "0" denotes an unrestricted coefficient sign.

5. In the final model specification the significance of each included variable can be evaluated by comparing the prior with the posterior inclusion probability of each variable:

    - the prior inclusion probability (PrIP) is the probability that variable X_i is included in a model if each variable is picked randomly. Let n denote the maximum number of regressors allowed in each model specification and let K denote the number of potential covariates. Then

    $$PrIP_i = \frac{\frac{\sum_{j=1}^{n} j \frac{K!}{j!(K-j)!}}{\sum_{j=1}^{n} \frac{K!}{j!(K-j)!}}}{K}$$

    - the posterior inclusion probability (PoIP) is defined as

    $$PoIP_i = \sum_{j=1}^{N} I_{\beta_i \neq 0} \pi_i \ ,$$

      i.e. it corresponds to the sum over the posterior model probabilities of all N models in which the covariate i takes a non-zero value.

    - a covariate X_i is significant if $PoIP_i > PrIP_i$.

B Quantile-mapping benchmark constraints

For deriving the benchmark constraints for the filtered model space we pursue a quantile mapping approach as suggested by Bonti et al. (2006). Quantile mapping is a simple approach for deriving a non-linear dependency between variables, as the only required assumption is a co-monotone relationship. It does not require any OLS assumptions and would, for example, not suffer from collinear dependencies between covariates. Therefore, the advantage of using a quantile mapping approach to generate benchmarks for the BMA-derived model space is that the two model classes are independent of each other. The benchmarks are thus not affected by biased or imprecise estimation of the model space.

When computing the quantile mapping benchmarks we proceed as follows.

1. We start from a Vasicek (2002) one-factor model to derive a relationship between the macro scenario and the sector PD. The Merton-Vasicek model assumes that the value of borrower j in sector s, $A_{j,s,t}$ is driven by a systemic factor $Z_{s,t}$ and an idiosyncratic factor $U_{j,s,t}$, both of which are assumed to be standard-normally distributed:

$$A_{j,s,t} = \rho_s Z_{s,t} + \sqrt{1-\rho_s^2} U_{j,s,t} \,.$$

   We assume that a borrower defaults if her value drops below a constant threshold D_s. It is straightforward to show that under these assumption the sector PD can be derived as

$$PD_{s,t} = \Phi\left(\frac{D_s - \rho_s Z_{s,t}}{\sqrt{1-\rho_s^2}}\right) , \tag{2}$$

   where Φ is the standard normal cdf.

2. Solving Equation (2) for the systemic factor yields

$$z_{s,t} = \frac{D_s - \Phi^{-1}\left(PD_{s,t}\right)\sqrt{1-\rho_s^2}}{\rho_s} , \tag{3}$$

   which gives us the systemic factor $z_{s,t}$ for a given D_s, ρ_s and the time series of the observed sector PDs from the German credit register.

3. D_s is calibrated using the long-run average residential mortgage sector PD from 2008Q1 to 2016Q4:

$$D_s = \Phi^{-1}\left(\bar{PD}_s\right) \,. \tag{4}$$

   This calibration is based on the assumption that $A_{j,s,t}$ and thus $Z_{s,t}$ are standard-normally distributed. In particular, it requires that $Var(Z_{s,t}) = 1$. Thus for Equation (4) to be a valid calibration for D_s we have to ensure that this assumption is indeed satisfied. ρ_s is calibrated to satisfy this assumption. We use Equation (3) to compute $z_{s,t}$ for 100,000 $\rho_s \in [0,1]$. The optimal ρ_s implies $Var(z_{s,t}) = 1$.

4. Equipped with the calibrated parameters we link the macro scenario to the systemic factor $z_{s,t}$. We standardize all macro variables $x_{i,t} \in X_t$ through $\tilde{x}_{i,t} = I_i(x_{i,t} - \bar{x}_i)/\sigma_{x_i}$, where we impose a "sign restriction" I_t such that a negative co-monotonicity between $\tilde{X}_t$ and $PD_{s,t}$ can be expected.

5. Given the time series for $z_{s,t}$ implied by Equation (3) and the time series for the standardized macro variable $\tilde{X}_t$, we estimate the cdf $\hat{F}_z$ and $\hat{F}_{\tilde{X}_i}$ using kernel density estimation.

6. We can then compute the stressed systemic factor $\hat{z}_{s,t+k}$ for $k \in$ stress horizon, using the quantile mapping[26]:

$$\hat{z}_{s,t+k} = \hat{F}_z^{-1}\left(\hat{F}_{\tilde{X}_i}\left(x_{i,t+k}\right)\right), \forall i, k$$

and compute the implied stressed PD forecast

$$\hat{P}D_{s,t+k} = \Phi\left(\frac{D_s - \rho_s \hat{z}_{s,t+k}}{\sqrt{1-\rho_s^2}}\right)$$

7. We apply a univariate quantile mapping. To this end, we use the macro variable $x_{i,t} \in X_t$ which implies the lowest RMSE when predicting the sector PD in-sample[27]

8. We use the estimated stressed PD increases $\hat{P}D_{s,T}$ at the end of the stress horizon T as benchmark constraints for the BMA model space. In particular, we set the constraints to $\pm 2sd(\Delta PD)$ where ΔPD denotes the relative PD changes between 2008 and 2016.

C Through-the-cycle adjustment algorithm

To compute through-the-cycle (TTC)-adjusted sector PD time series we proceed as follows.

1. We apply a four-year rolling window on the stressed risk parameter time series to get a smoothed scenario forecast.

2. We calculate the implicit TTC-adjustment factor in the distance-to-default space:

$$\alpha_t^{TTC} = \Phi^{-1}\left(PD_t^{TTC}\right) - \Phi^{-1}\left(PD_0^{PIT}\right),$$

where $t \in$ stress horizon and PD_0^{PIT} is the initial (pre- stress) corresponding point-in-time value.

3. We apply the TTC-adjustment factor to the bank-specific initial PD. Thus, for bank i

$$PD_{i,t}^{TTC} = \Phi\left(\Phi^{-1}\left(PD_{i,0}^{PIT}\right) + \alpha_t^{TTC}\right)$$

For the TTC-adjusted LGDs, we stop at step 1, i.e. calculate bank-specific long-term averages via the moving window approach.

[26] Note that the quantile mapping simply maps the nth worst realization of the macro variable to the nth worst realization of the systemic factor.

[27] The results are robust to the application of multivariate quantile mapping, which weights each univariate PD forecast with the inverse of its RMSE.

D The residential mortgage template in the LIRE 2017 survey

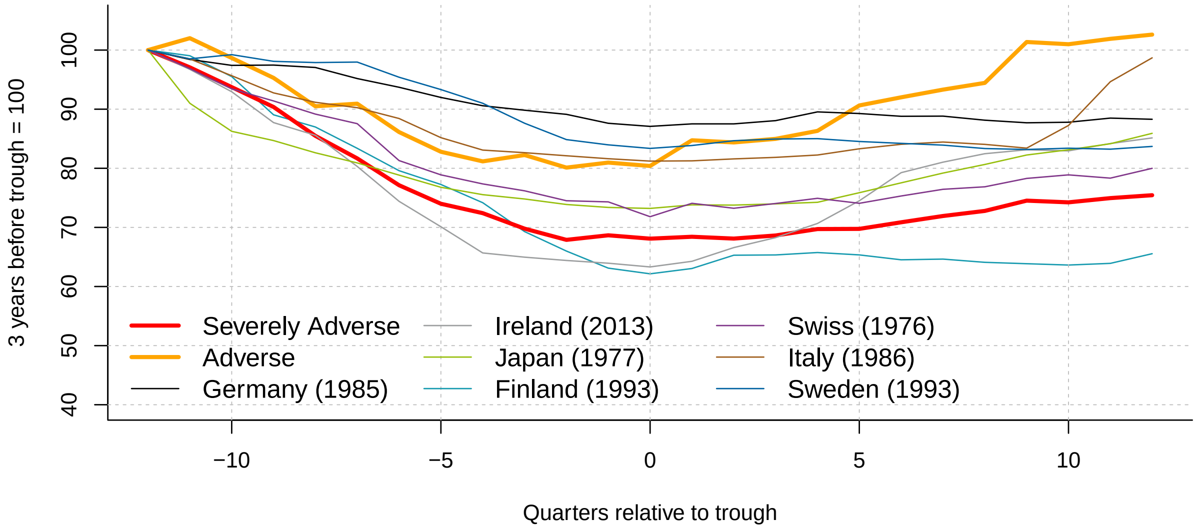
Figure D.1 shows the residential mortgage template of the LIRE 2017 survey. Banks provided data on up to 30 different rating classes (not shown in their entirety here) with corresponding volume-weighted PDs, pre- and post-collateral exposures and fully secured exposures. Exposures are provided for 2014, 2015 and 2016, of which we use only the 2016 values as initial exposures for the stress test.

		Volume-weighted PD	Exposure Pre-Collateral			Exosure Post-Collateral			Exposure Fully Secured		
ZNr	Residential mortgage loans by Rating-Classes	in %	in TEUR			in TEUR			in TEUR		
		31.12.2016	31.12.2014	31.12.2015	31.12.2016	31.12.2014	31.12.2015	31.12.2016	31.12.2014	31.12.2015	31.12.2016
13	Rating-Class 1										
14	Rating-Class 2										
15	Rating-Class 3										
16	Rating-Class 4										
17	Rating-Class 5										

Figure D.1: Example residential mortgage survey sheet

E The macro scenarios in a historical perspective

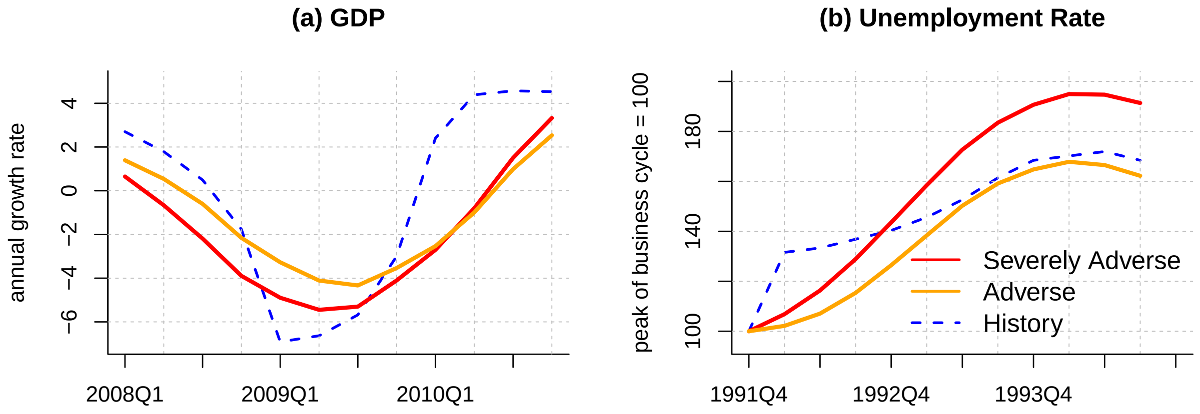
To provide some intuition about the severity of the assumed macro scenario this section compares the assumed macro dynamics with historical experience in Germany and in other countries. Figure E.2 shows the two assumed scenario paths for the house price index relative to a number of selected real estate crises. Given that the severely adverse scenario reproduces exactly the Spanish dynamics in 2011-2013 it is not surprising that the dynamics are in line with historical experience. Compared to the housing crisis in Germany in 1985 we find that both scenarios feature a stronger decline in house prices but are less persistent than the German experience. The house price scenario therefore seems to be adverse yet plausible for the German housing market.



Source: OECD and Eurostat.

Figure E.2: Stress scenarios compared to international experience

Figure E.3 compares the dynamics for GDP and unemployment derived from the VARX model to dynamics actually experienced in Germany. We find for GDP a decrease in growth that is less severe than during the Great Recession, although more persistent. This seems to be in line with the evidence in Claessens, Kose, and Terrones (2009) that recessions associated with house price busts appear to be deeper and more persistent than those without. We compare the unemployment rate dynamics to the German recession starting in 1992. We find that the initial scenario-implied increases are less strong. In both scenarios the unemployment rate peaks after about two years, which corresponds closely to the dynamics observed during the recession.



Source: OECD and Eurostat.

Figure E.3: Stress scenarios compared to German experience

F CLTV-LGD dependency

This section provides the data for the CLTV-LGD dependency based on Qi and Yang (2009) and Palmroos (2016) underlying Figure 5.

Table F.2: CLTV-LGD dependency

CLTV	LGD	CLTV	LGD	CLTV	LGD	CLTV	LGD
0.000	0.07	0.525	0.14	1.025	0.34	1.525	0.63
0.025	0.07	0.550	0.14	1.050	0.35	1.550	0.65
0.050	0.07	0.575	0.15	1.075	0.36	1.575	0.67
0.075	0.07	0.600	0.15	1.100	0.37	1.600	0.68
0.100	0.07	0.625	0.16	1.125	0.38	1.625	0.70
0.125	0.07	0.650	0.16	1.150	0.41	1.650	0.71
0.150	0.07	0.675	0.16	1.175	0.42	1.675	0.73
0.175	0.08	0.700	0.17	1.200	0.44	1.700	0.75
0.200	0.08	0.725	0.17	1.225	0.45	1.725	0.76
0.225	0.09	0.750	0.18	1.250	0.47	1.750	0.78
0.250	0.09	0.775	0.18	1.275	0.49	1.775	0.80
0.275	0.10	0.800	0.19	1.300	0.50	1.800	0.82
0.300	0.10	0.825	0.19	1.325	0.52	1.825	0.83
0.325	0.10	0.850	0.22	1.350	0.54	1.850	0.85
0.350	0.11	0.875	0.23	1.375	0.54	1.875	0.87
0.375	0.11	0.900	0.24	1.400	0.55	1.900	0.89
0.400	0.12	0.925	0.27	1.425	0.57	1.925	0.91
0.425	0.12	0.950	0.28	1.450	0.59	1.950	0.93
0.450	0.13	0.975	0.30	1.475	0.60	1.975	0.96
0.475	0.13	1.000	0.32	1.500	0.61	2.000	0.98
0.500	0.13						

Notes: Values correspond to data used for Figure 5. All LGD values rounded to two decimal places.

References

Allen, L. (2004). The Basel Capital Accords and International Mortgage Markets: A Survey of the Literature. *Financial Markets, Institutions & Instruments*.

Basel Committee on Banking Supervision (2005). An Explanatory Note on the Basel II IRB Risk Weight Functions. *Consultative Document*.

Basel Committee on Banking Supervision (2012). Models and tools for macroprudential analysis. *Bank of International Settlements Working Paper No. 21*.

Basel Committee on Banking Supervision (2015). Second consultative document - Revisions to the Standardised Approach for credit risk. *Standards*.

Bonti, G., M. Kalkbrener, C. Lotz, and G. Stahl (2006). Credit risk concentrations under stress. *Journal of Credit Risk 2*(3), 115–136.

Calem, P. S. and M. LaCour-Little (2004, March). Risk-based capital requirements for mortgage loans. *Journal of Banking & Finance 28*(3), 647–672.

Claessens, S., M. A. Kose, and M. E. Terrones (2009). What happens during recessions, crunches and busts? *Economic Policy 24*(60), 653–700.

Coleman, A., N. Esho, I. Sellathurai, and N. Thavabalan (2005). Stress testing housing loan portfolios: A regulatory case study. *Australian Prudential Regulation Authority, Working Paper 2005-01*.

Corbae, D., P. D'Erasmo, S. Galaasen, A. Irarrazabal, and T. Siemsen (2017). Structural stress tests. *University of Wisconsin Working Paper*.

Deutsche Bundesbank (2017, February). Monthly report.

Düllmann, K. and T. Kick (2014, November). Stress testing German banks against a global credit crunch. *Financial Markets and Portfolio Management 28*(4), 337–361.

European Banking Authority (2016). 2016 EU-wide Stress Test - Methodological note. Technical report, European Banking Authority.

Foglia, A. (2009, September). Stress Testing Credit Risk: A Survey of Authorities' Approaches. *International Journal of Central Banking 5*(3), 9–45.

Furnival, G. and R. Wilson (1974, December). Regressions by Leaps and Bounds. *Technometrics 16*(4), 499–511.

Gaffney, E., R. Kelly, and F. McCann (2014). A transitions-based framework for estimating expected credit losses. *Central Bank of Ireland Research Technical Paper*.

Gross, M. and J. Poblacion (2017). Implications of Model Uncertainty for Bank Stress Testing. *Journal of Financial Services Research*, 1–28.

Gupton, G. M. and R. M. Stein (2002). LOSSCALC: Model for predicting Loss Given Default. *Moody's KMV*.

Henry, J. and C. Kok (2013). A macro stress testing framework for assessing systemic risks in the banking sector. *Occasional Paper Series, No 152, ECB*.

Hirano, K. and J. H. Wright (2017, March). Forecasting with model uncertainty: Representations and risk reduction. *Econometrica 85*(2), 617–643.

Hott, C. (2015, February). A model of mortgage losses and its applications for macroprudential instruments. *Journal of Financial Stability 16*, 183–194.

Koziol, P., C. Schell, and M. Eckhardt (2015). Credit risk stress testing and copulas – is the Gaussian copula better than its reputation? *Deutsche Bundesbank Discussion Paper No 46/2015*.

Misina, M. and D. Tessier (2008). Non-Linearities, Model Uncertainty, and Macro Stress Testing. *Bank of Canada, Working Paper No. 2008-30*.

Palmroos, P. (2016). Correlation between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools. *Aalto University Working Paper*.

Pelster, M. and J. Vilsmeier (2017). The determinants of CDS spreads: evidence from the model space. *Review of Derivatives Research (forthcoming)*.

Pennington-Cross, A. (2003). Subprime and prime mortgages: Loss distributions. *Working paper, Office of Federal Housing Enterprise Oversight*.

Qi, M. and X. Yang (2009). Loss given default of high loan-to-value residential mortgages. *Journal of Banking and Finance 33*, 788–799.

Rodriguez, A. and C. Trucharte (2007, December). Loss coverage and stress testing mortgage portfolios: A non-parametric approach. *Journal of Financial Stability 3*(4), 342–367.

Seidler, J. and P. Jakubik (2009). The Merton Approach to Estimating Loss Given Default: Application to the Czech Republic. *CNB WP No. 13/2009*.

Siemsen, T. and J. Vilsmeier (2017). On a quest for robustness: About model risk, randomness and discretion in credit risk stress tests. *mimeo*.

Vasicek, O. (2002, December). Loan Portfolio Value. *Risk*, 160–162.

Wright, J. H. (2008, October). Bayesian Model Averaging and exchange rate forecasts. *Journal of Econometrics 146*(2), 329–341.