



Federal Reserve
Bank of Dallas

Understanding the Exposure at Default Risk of Commercial Real Estate Construction and Land Development Loans

Shan Luo and Anthony Murphy

Working Paper 2007

March 2020

Research Department

<https://doi.org/10.24149/wp2007>

Working papers from the Federal Reserve Bank of Dallas are preliminary drafts circulated for professional comment. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

Understanding the Exposure at Default Risk of Commercial Real Estate Construction and Land Development Loans*

Shan Luo[†] and Anthony Murphy[‡]

February 2020

Abstract

We study and model the determinants of exposure at default (EAD) for large U.S. construction and land development loans from 2010 to 2017. EAD is an important component of credit risk, and commercial real estate (CRE) construction loans are more risky than income producing loans. This is the first study modeling the EAD of construction loans. The underlying EAD data come from a large, confidential supervisory dataset used in the U.S. Federal Reserve's annual Comprehensive Capital Assessment Review (CCAR) stress tests. EAD reflects the relative bargaining ability and information sets of banks and obligors. We construct OLS and Tobit regression models, as well as several other machine-learning models, of EAD conversion measures, using a four-quarter horizon. The popular LEQ and CCF conversion measure is unstable, so we focus on EADF and AUF measures. Property type, the lagged utilization rate and loan size are important drivers of EAD. Changing local and national economic conditions also matter, so EAD is sensitive to macro-economic conditions. Even though default and EAD risk are negatively correlated, a conservative assumption is that all undrawn construction commitments will be fully drawn in default.

JEL Classification: G21, G28.

Keywords: Credit Risk, Commercial Real Estate (CRE), Construction, Exposure at Default, EAD Conversion Measures, Macro-sensitivity, Machine Learning.

* The views expressed are those of the authors, and not necessarily those of the Federal Reserve System, the Federal Reserve Bank of Chicago or the Federal Bank of Dallas.

[†]Shan Luo, Federal Reserve Bank of Chicago, 230 S. LaSalle St., Chicago, IL 60604, Shan.Luo@chi.frb.org.

[‡]Anthony Murphy, Federal Reserve Bank of Dallas, 2200 N. Pearl St., Dallas, TX 75201, USA. Anthony.murphy@dal.frb.org.

1. Introduction

The U.S. experienced one national house price boom and bust and two commercial real estate (CRE) booms and busts since the late 1980s. During the Great Recession and Financial Crisis, the scale of the CRE boom and bust was similar to that of residential real estate. Although, the peak in CRE delinquency rates was lower during the most recent recession, the rate of commercial bank charge-offs was similar, and the fall in CRE prices was greater as shown in Figures 1(a) to (c). Both residential and commercial real estate construction loans are highly pro-cyclical. The sub-prime boom and bust in residential house prices, sales, construction, defaults and foreclosures is well documented. However, the accompanying CRE boom and bust has received far less attention, in part because banks tended to extend their lending to distressed facilities due to complex loan structures and the financial regulators' 2009 policy statement on prudent CRE workouts (Federal Reserve et al., 2009).¹

-- Figures 1(a) to 1(c) --

Income producing CRE facilities are generally highly drawn - the average utilization ratio is above 97% and about 95% of defaulted facilities are fully utilized - so we study the variation in the draw rates and exposure at default (EAD) for construction and land development ("construction") facilities, which are riskier than income producing ones. Credit risk is commonly measured using an expected loss (EL) approach, the product of the probability of default (PD), loss given default (LGD), and exposure at default (EAD), i.e. $EL = PD \times LGD \times EAD$. Compared with PD and LGD modelling, EAD modelling of construction loans is very underdeveloped since the available data are generally proprietary and span relatively short, benign time periods. As a result, the data are not very informative about the determinants of stressed EAD, or cyclical variations in EAD. Under the advanced internal-ratings approach, the Basel Capital Accords allow banks to estimate their capital requirements using their own EAD risk parameters. However, it is difficult to judge the reliability and robustness of these estimates given the absence of a standard EAD modelling framework and the limitations of many datasets (Bank for International Settlements. 2016).

This paper focuses on the EAD risk of construction loans using a unique dataset for the U.S., and fills an important gap in the limited literature, which focuses almost exclusively on the EAD of retail and corporate facilities. We identify and model the determinant of the EAD of construction loans. We

¹ The 2009 policy statement provided new guidance for examiners and financial institutions working with CRE borrowers experiencing diminished operating cash flows, depreciated collateral values, or prolonged delays in selling or renting commercial properties. The financial regulators recognized that prudent loan workouts are often in the best interest of both financial institutions and borrowers, particularly during difficult economic conditions.

highlight the importance of lagged utilization rate, loan characteristics and macroeconomic conditions in explaining EAD.

Our research leverages confidential Federal Reserve supervisory data, as well as some earlier special collection, for the largest U.S. banks.² The data are used in the annual Dodd Frank Act stress test and Comprehensive Capital Assessment Review (CCAR) exercises conducted by the Federal Reserve. The supervisory data are at the facility level, and captures the risky characteristics of the CRE construction facilities above \$1 million held by the CCAR banks. We link the supervisory data with aggregate, macro-economic data (e.g. BBB corporate bond spreads), local economic variables (e.g. county level house prices and unemployment rates), and MSA and property type specific data (e.g. CRE property price indices and vacancy rates). Compared with studies of corporate or retail EAD, the number of construction defaults is relatively small. Even though our data do not span a full business cycle, our data for 2010 Q4 to 2017 Q4 includes quite a lot of stressed EAD observations from the Great Recession.

In contrast to the extensive literature on the usage of credit cards and credit lines, the EAD literature is relatively sparse – see the surveys in Jacobs and Bag (2013) and Gürtler et al. (2018) for example – and almost exclusively focuses on the EAD of corporate or retail facilities, e.g. home equity lines of credit, corporate credit lines, corporate and retail credit cards. Araten and Jacobs (2001), Asarnow and Marker (1995), Bag and Jacobs (2012), Jacobs (2010), Jiménez et al. (2008, 2009), Yang and Tkahenko (2010) and Zhou et al. (2014) study corporate loans or credit lines in U.S., Canada (Yang and Tkahenko) and Spain (Jiménez et al.). Kim (2008), Taplin et al. (20007) and Valvonis (2008) study corporate credit cards in Korea, Australia and the EU, respectively. Agarwal et al. (2006) study retail home equity lines of credit in the U.S. Banerjee and Canals-Cerdá (2012), Hon and Bellotti (2016), Leow and Crook (2016), Qi (2009), Tong et al. (2016) and Valvonis (2008) study retail credit cards in the EU, UK and U.S. Gürtler et al. (2018) study checking accounts at a private European bank. Bank for International Settlements (2016) studies corporate and retail facilities in 17 countries.

The literature suggests that previous utilization rates are by far the most important predictor of current EAD. Inter alios, Jacobs (2008) and Jiménez, Lopez and Saurina (2009) find high utilization rates several years prior to default. Hon and Bellotti (2016) find that lagged balances are the most important of future retail credit card balances. Barakova and Parthasarathy (2012) showed that banks seldom cut corporate limits or restrict draws until they rate the exposure as higher risk, or line use is very high. Barakova and Parthasarathy (2013) discussed the impact of risk rating, line utilization, exposure size and

² We refer to bank holding companies and international holding companies as “banks”.

economic conditions on the EAD risk of firms, and report that higher EAD is associated with default that is difficult to anticipate. A number of papers show that, in addition to high lagged utilization rates, low credit ratings or scores and greater time to default are significant drivers of EAD (e.g. Tong et al., 2016). We do not rely on heterogeneous bank internal ratings or a common external mapping of the internal ratings – we leave this as a topic for future research. Instead, we use a variety of loan characteristics, obligor and lender attributes as controls in our EAD models. A priori, the inclusion of time to default in an EAD model is rather problematic since it is unknown prior to default. Since, we focus on the determinants of EAD, and use a fixed four-quarter horizon when constructing our EAD conversion measures, we do not include time to default in our models.

In our study, we calculate and compare four commonly used, realized EAD risk measures: the loan equivalent (LEQ), credit conversion factor (CCF), exposure at default factor (EADF) and an additional utilization factor (AUF). The LEQ and CCF are both volatile EAD measures - when facilities are close to being fully utilized in the case of LEQ or unutilized in the case of CCF – so we focus on the closely related, more stable EADF and AUF measures. We model the EADF and AUF measures using a variety of parametric regression-based models, as well as various machine-learning-based approaches. We conducted extensive in-sample and out-of-sample performance tests to check the specification of our parametric models and compare the choice of different EAD estimators and modeling approaches. We find that the fit of parametric and machine-learning-based models of construction EAD is similar. The advantage of parametric models is that they are not a black box so the EAD risk drivers are explicitly identified and the plausibility of their EAD impact may be assessed. In line with the EAD literature for other types of lending, we find the current utilization rate is the single most important factor in driving future EAD risk. Loan characteristics, obligor and lender attributes, and national and local economic factors also help to explain EAD risk. A conservative assumption is that all undrawn construction commitments will be fully drawn.

The outline of this paper is as follow. The underlying FR Y-14 data are discussed in Section 2. The EAD conversion measures are described, tabulated and plotted in Section 3. The EAD conversion measures for construction and corporate loans are also compared in this section. EAD is often described as the outcome of the “race to default” between borrowers and lenders, reflecting the relative information of obligors and banks. Dynamic snapshots of this “race” are provided in Section 4. The estimation data, econometric models and results are set in Section 5, while some machine-learning robustness results are presented in Section 6. Finally, Section 7 summarizes our analysis and presents our conclusions.

2. Data

Data scarcity has been, and continues to be, the greatest challenge when modeling EAD. Our analysis uses detailed, confidential Federal Reserve supervisory data on CRE facilities (loans and credit lines) at the largest bank holding companies (BHCs) and international holding companies (IHCs) in the U.S. The data are used in the Comprehensive Capital Analysis and Review (CCAR) and Dodd-Frank Act stress test (DFAST) exercises conducted by the Federal Reserve (e.g. Board of Governors of the Federal Reserve, 2019). The CCAR and DFAST stress tests are the main macro-prudential tool of the Federal Reserve.

We use quarterly FR Y-14Q commercial real estate facility data starting in the December of 2010. The data are rich in breadth and depth, and provide a unique insight into the risk characteristics of individual CRE facilities held by large banks.³ Inter alia, detailed information on obligor characteristics, loan details (such as committed and utilized balances, maturity date, recourse), lender details including the line of business, underwriting information such as initial LTV and charge-offs, property type and location are collected. Please see the Appendix for further information about the data. In addition to the loan and obligor characteristics from FR Y-14Q data, our dataset includes national macroeconomic variables (such as the real GDP growth and the BBB corporate bond spread) and local economic variables (county-level house price indices and unemployment rates, MSA-level vacancy / occupancy rates and indices of CRE prices, net operating incomes, cap rates and rents by property type).⁴

-- Table 1 --

As of December 2017, the FR Y-14Q dataset had data on approximately 390,000 different CRE facilities at 39 large U.S. banks (BHCs and IHCs). We record a default event if any of the following conditions is met: a payment is 90 or more days past due, a positive net charge-off occurs, the facility is placed on non-accrual, an ASC 310-10 reserve is set aside or the facility is extended even though borrower is poorly rated. Since facilities may remain in default for more than one quarter, we restrict our analysis to the first default occurrence. We drop facilities that

³ Data are collected for all CRE facilities with original committed amounts greater than or equal to \$1 million.

⁴ The CRE data are supplied by CBRE.

enter the dataset in default, since we don't observe the utilized and committed amounts in earlier quarters. Our final EAD dataset has a total of 1,777 unique defaulted construction loans. The number of defaults by default event and year are tabulated in **Table 1**. Although our data does not span a complete business cycle, we believe the data are informative about stressed construction EAD since there are few defaults and EAD is low when the economy is expanding or booming.

3. EAD Conversion Measures

Exposure at default (EAD) captures the loss exposure at the time of default. It can be stated as a dollar amount, as well as a percentage of committed amount on a credit line. Directly modelling EAD dollar amounts is difficult since the EAD distribution ranges from zero or very low to extremely large amounts. Thus, the literature generally focuses on estimating and using more robust EAD conversion measure models (e.g. Yang and Tkachenko, 2012).⁵

A number of papers - including the recent Gürtler et al. (2018), Hahn and Reitz (2011) and Mural (2011) papers - discuss different approaches to measuring EAD. Four ex-post EAD conversion measures are widely used. We calculate these EAD measures using a four-quarter fixed time horizon approach since the fixed horizon approach (with a four-quarter horizon) is more widely used than cohort method or variable time horizon approaches. The most popular EAD conversion measure is the loan equivalent factor (LEQ), which is often called the credit conversion factor (CCR) in Europe. LEQ is defined as:

$$LEQ_t = \frac{\Delta_4 U_t}{C_{t-4} - U_{t-4}},$$

where t is the default date, U_{t-4} and C_{t-4} are the drawn or utilized amount and credit limit or committed amount four quarters prior to default, and $\Delta_4 U = U_t - U_{t-4}$ is the four-quarter change in the utilized amount prior to default. LEQ is undefined when the facility is fully drawn (i.e. the denominator of LEQ is zero), which is often the case for defaulted construction and land development loans. In this case, LEQ is often assigned a zero value.⁶ EAD is easily calculated using the LEQ measure:

$$EAD_t = U_{t-4} + LEQ_t \times (C_{t-4} - U_{t-4})$$

⁵ In our dataset, the median commitment is about \$3 million, with 10th and 90th percentiles of \$1.2m and \$18.6m respectively. The medians for fully and part drawn facilities are \$4.2m and \$2.2m, respectively.

⁶ Modeled LEQ is often restricted to the [0,1] range or, less frequently, to the [-1,1] range.

Note that LEQ can be unstable and volatile for high utilization loans, since the denominator of LEQ is small when the lagged utilization rate, U_{t-4}/C_{t-4} , is high. The other three common EAD conversion measures are the credit conversion factor, $CCF_t = U_t/U_{t-4}$, the exposure at default factor, $EADF_t = U_t/C_{t-4}$, and the additional utilization factor, $AUF_t = \Delta U_t/C_{t-4} = EADF_t - U_{t-4}/C_{t-4}$, i.e. AUF equals $EADF$ minus the lagged utilization rate. Note that the CCF conversion factor can be unstable when the utilization rate is low. The EAD calculations for the CCF , $EADF$ and AUF measures are:

$$EAD_t = CCF_t \times U_{t-4}, \quad EAD_t = EADF_t \times C_{t-4}, \quad EAD_t = U_{t-4} + AUF_t \times C_{t-4}$$

The medians of the four construction EAD conversion measures are set out in **Table 2**, along with their corporate equivalents. Even though defaulted constructions loans are less likely to be fully utilized than corporate loans one year prior to default, the 0.81 median construction $EADF_t = \frac{U_t}{C_{t-4}}$ is considerably higher than the corresponding 0.64 number for corporate facilities. Apart from this case, the median EAD conversion measures for construction and corporate loans in **Table 2** are very similar. A conservative EAD assumption is to assume that all construction facilities, including ones that are not fully drawn at quarter t- 4, are fully drawn when default occurs one year later at time t.

-- Table 2 --

-- Figures 2 and 3 --

The histograms of the four EAD conversion measures (**Figure 2**) clearly show the wide dispersion in the LEQ and CCF measures due to a small number of very high utilization, respectively low, rate facilities. This is the reason why we only model the other two, more stable $EADF$ and AUF measures in Section 5. Both the $EADF_t = U_t/C_{t-4}$ and $AUF_t = \Delta U_t/C_{t-4}$ measures have heavy, left hand tails with peaks at 1 and 0 respectively, since the utilized amount at default seldom exceeds the committed amount one year earlier and many defaulted loans are fully utilized. This pattern is even clearer in **Figure 3**, where separate $EADF$ and AUF histograms are plotted for fully drawn and partly drawn facilities at quarter t – 4, one year prior to default.

4. EAD and the “Race to Default”

The behavior of EAD approaching default is often characterized a race to default (**Figure 4**). As the obligor’s credit quality deteriorates, it has an incentive to increase the facility utilization rate by drawing on the available credit while it is still available. However, the lender’s incentive is to cut off or limit the obligor’s credit once it observes or infers a steady increase in credit risk. However, the effectiveness of

the clampdown on credit depends on the effectiveness and extent of monitoring and the strength of covenants. In practice, this is an empirical issue.

-- Figure 4 --

The change in construction and corporate utilization rates in the year prior to default, and on default, are set out in **Table 3**. The contrast in the behavior of construction and corporate utilization rates is interesting. As expected, utilization rates are much higher for construction facilities than for corporate facilities, and are more likely to be fully drawn one year prior to default (57% versus 15%). **Table 3** shows that utilization rates do indeed rise prior to default. The median utilization rate for part-utilized construction facilities rises from 0.86 one year prior to default to 0.95 on default. The corresponding increase in the median utilization rate for part-utilized corporate facilities is larger – from 0.54 one year before to 0.70 on default. The mean utilization rates are somewhat lower than the median rates, but the pattern is similar. Unsurprisingly, the median utilization rate for full-drawn corporate and construction facilities is 1.0.

-- Tables 3 and 4 --

Is there evidence that banks successfully cut off or reduce the credit lines of obligors who eventually default? The statistics in **Table 4** suggests that banks are only partly successful in doing so. The median level of construction and corporate commitments on default is 95% of the level of commitments one year before (C_{t-h}/C_{t-4} , $h = 0,1,2,3$), but almost all of the reduction occurs in the default quarter ($h = 0$). The time path of the mean of C_{t-h}/C_{t-4} is different from that of the median – it declines every quarter to 88% or 89% on default, with the largest decline in the default quarter. Overall, our results on the race to default for construction loans line are consistent with the evidence in Barakova and Parthasarathy (2013). They suggest that banks seldom cut limits or restrict corporate draws until they rate the exposure as higher risk, or line use is very high, and that firms that anticipate future deterioration are able to preempt banks by drawing more in advance of restrictions.

5. Data, Econometric Models and Estimation Results

Some summary statistics for the variables in our OLS and Tobit regression models are set out in **Table 5**. We model the four-quarter horizon EADF and AUF measures, since they are more regular, i.e. have fewer outliers than the other two EAD conversion measures. There are two types of explanatory variables in our models – loan specific variables such as the lagged facility size or an indicator for a

homebuilder or land developer facility, and macroeconomic / local economic variables. The latter consist of the year-on-year change in the BBB corporate bond yield spread over 10-year Treasuries and the changes in residential and commercial real estate prices - the county-level change in house prices and the MSA, property type specific change in CRE prices.

-- Table 5 --

Since we are interested in the determinants of construction EAD, and given the relatively small sample numbers compared to other studies of retail and corporate EAD, we focus on relatively simple OLS and Tobit model for $EADF_t = U_t/C_{t-4}$ and $AUF_t = \Delta U_t/C_{t-4}$, our more stable EAD measures. Some estimated models are presented in **Tables 6 and 7**. The explanatory variables were preselected from a large set of loan level characteristics, and macroeconomic / local economic variables, and the selection verified using automatic model selection and cross-validation procedures. The *EADF* and *AUF* results are consistent with each other, and plausible.

Consider the loan characteristic results in columns (1) and (2) of **Table 6**. *EADF* is negatively related to the size of the facility and size interacted with the bank's estimate of the probability of default (*PD*) of the facility, consistent with the notion that banks pay more attention to larger, riskier loans.⁷ *EADF* is also higher if the utilization rate one year ago is higher, but declines slightly as the ratio of interest reserves to commitments increases. It is also lower for homebuilders and land /lot developers. Those two property types are relatively high risk, so banks are likely to monitor their utilization rates closely.

-- Table 6 --

Local and national economic conditions also matter. *EADF* rises when local property prices fall, and when the BBB corporate bond spread widens. Wider corporate bond spreads are a proxy for wider economic and financial stress. The fact that changes in local residential property prices are significant while changes in local, property type specific CRE prices are not, may just reflect measurement problems.⁸ Column (2) shows that dropping $\Delta_4 \ln CRE Price_t$ leaves the other estimates unchanged.

⁷ The sample sizes in columns (1) and (3) are lower than in columns (2) and (4) since not all banks are required to report their internal *PD* estimates.

⁸ Both property prices – residential and commercial - tend to show up significant in CRE probability of default (*PD*) and exposure at default (*EAD*) models.

Other local CRE variables, such as the changes in CRE rents and vacancy rates, are correctly signed but insignificant, so they are not included in the regressions.

Banks differ in many dimensions, including specialization and risk appetite, so columns (3) and (4) of **Table 6** add bank fixed effects to the *EADF* specification. Although the fixed effects are jointly significant, the results are similar to those in columns (1) and (2), except that the lagged $PD \times \ln Facility Size$ interaction term is insignificant and the coefficient on the change in local house price halves in size. Overall, the effect of adding bank fixed effects on the fit of the model is rather modest. We include bank fixed effects in all subsequent models.

Very often, *EADF* numbers are restricted to the range [0,1] - negative *EADF* numbers are set to zero and *EADF* values greater than one are set to one. The two-limit Tobit model results in column (5) reflect this restriction and takes account of the different *EADF* distributions of fully utilized and other facilities. Specifically, the Tobit model includes an indicator for facilities that were not fully utilized one year earlier, as well as this indicator interacted with the lagged ratio of interest reserves to commitments. Both of these effects are significant. The remaining parameter estimates are similar to the ones in earlier columns, apart from the higher coefficient on the change in local house prices.

The OLS version of the Tobit specification is set out in column (6). The Tobit and OLS coefficients are very similar. After taking account of the different *EADF* distributions of fully utilized and other facilities, changes in local house prices are still a better proxy for changing local economic conditions than the change in MSA-level CRE prices by property type.

The adjusted R^2 s of the OLS models, which we have discussed so far, range from 34% to 41%. However, our suspicion is that the fit is mainly attributed to the fit to facilities that were not fully utilized one year earlier. The results in columns (7) and (8) confirm this suspicion – the adjusted R^2 of our *EADF* model for partly utilized facilities is 43% as opposed to 15% for fully utilized facilities (37% and 6%, respectively, without fixed effects). The parameter estimates for part utilized facilities in column (7), are also larger in absolute size and more significant than in column (8).

How important is the local house price effect in the *EADF* equation for part-utilized facilities in column (7)? Consider the Federal Reserve's CCAR severely adverse scenario for 2019. The maximum four-quarter fall in national house prices is 16 percent. If the 16 percent decline in national house prices applied everywhere, the OLS estimates in column (8) suggest that *EADF* for part-utilized facilities might fall by about 0.07 on average. Although the magnitude of the economic variables varies somewhat with

the model specification, the results suggest that EADF has a cyclical component – declining when property prices are falling and spreads widen and vice versa. A number of previous papers, including Jacobs (2008), also find a negative relationship between EAD risk and PD / default risk.

-- Table 7 –

Table 7 present a condensed set of OLS results for the *AUF* conversion measure. Since, apart from winsorizing, $AUF_t = EADF_t - U_{t-4}/C_{t-4}$, it should come as no surprise that the results in the table are qualitatively very similar to the *EADF* results in columns (6) to (8) of **Table 6**, the one exception being the negative coefficient on the lagged utilization rate. The same variables are generally significant (with similar t-stat rankings), and the coefficient on the lagged utilization rate is approximately equal to its **Table 6** coefficient minus one.

6. Machine Learning Based Robustness Checks

We built a number of machine learning based models of the EAD conversion measure models, and used the results as a check on the specification and performance of the models in the previous section (e.g. Breiman, 2001).⁹ The four parametric models were ordinary least square regressions, log-transformed linear regressions, gamma regressions and Tobit regressions. The non-parametric methods were regressions trees, bagging, random forests and boosting. We used stratified samples to randomly split our dataset 60:40 into an in-sample development or training dataset, and an out-of-sample test or validation dataset. We calculated R-squares, RMSEs and other unreported measures of fit for the datasets. Confidence intervals for the R-squares and RMSEs measures were generated using 100 resamples of our training and validation datasets. Machine learning approaches are generally employed using considerable larger datasets than ours, so the machine learning results are only indicative.

Regression trees were widely used in credit modeling, but have been superseded by better approaches including bagging, random forests and boosting (see James et al., 2013 or Kuhn and Johnson, 2013, for example). A decision tree uses features of the data to construct recursive binary splits of the data. The predicted EAD measure in a decision tree is just the average EAD measure in each terminal node of the tree. Bootstrap aggregation, or bagging, is used to reduce high variance and low prediction accuracy of decision trees. Bagging is an ensemble technique. Instead of building a single decision tree, multiple trees are built based on bootstrapped samples, and the final prediction is calculated by averaging the

⁹ The use of machine learning based EAD models is increasing. An early application is Yang and Tkachenko (2012).

individual tree predictions. A random forest further improves on bagging by constructing many weakly correlated trees, by randomly drawing a bootstrapping sample from the original data, and randomly drawing a subset of predictors for each tree. Boosting also builds many trees, but the trees are grown sequentially using information from the previously grown trees (by fitting their residuals). Boosting tends to learn slowly and employ small trees.

-- Table 8 --

Some indicative *EADF* and *AUF* machine learning results are presented in Table 8. The R^2 and RMSE results are not directly comparable with those in Tables 6 and 7 since, as a check on our specifications, we used a broader set of explanatory variables than the one in Tables 6 and 7. The training sample R^2 s are somewhat higher than the validation sample R^2 s, although the two sets of RMSEs are similar. The median validation sample R^2 s are 0.44 for *EADF*, and 0.28 for *AUF*, while the median RMSE's are 0.21 and 0.20. Note that the validation sample R^2 s and RMSEs are generally similar to those of the regression models in Tables 6 and 7, which suggest that we have not overlooked an important, systematic determinant of EAD.¹⁰

6. Summary and Conclusions

We examine the determinants of exposure at default (EAD) for large U.S. construction and land development loans from 2010 to 2017. EAD is an important component of credit risk, and construction loans are more risky than income producing CRE loans. In addition, most income producing loans in our dataset are fully drawn. The underlying EAD data come from a large, confidential supervisory dataset used in the annual Comprehensive Capital Assessment Review (CCAR) stress tests in the U.S.

This is the first study documenting and modeling the EAD of construction loans. There is a great deal of variation in EAD, even though the data do not cover a full business cycle. EAD reflects the relative bargaining ability and information sets of banks and obligors. We tabulate various EAD conversion measures, and show that the time paths of construction and corporate EAD differ as loans approach default. We present OLS and Tobit model results for EAD conversion measures, using a four-quarter fixed time horizon approach. The performance of these models is confirmed using a number of parametric and non-parametric machine learning based models.

¹⁰ Additional results for all four EAD conversion measures are available on request. No single measure dominates in the validation results, irrespective of whether we focus on the EAD conversion measures or EAD itself.

The popular LEQ conversion measure is unstable, so we focus on EADF and AUF measures. Property type, the lagged utilization rate and loan size are important drivers of EAD. Changing local and national economic conditions also matter, so EAD is sensitive to macro-economic conditions and negatively correlated with the probability of default. A conservative assumption is that all undrawn construction commitments will be fully drawn.

Appendix : FR Y-14Q Commercial Real Estate Data

The FR Y-14 collection began in September 2011 to support the Capital Assessment and Stress Testing regulatory requirements of the U.S. Federal Reserve. Inter alia, the data are used to assess the capital adequacy of large bank and intermediate holding companies (BHCs and IHCs) using forward-looking projections of revenue and losses, and to support supervisory stress test models. The respondent panel used in this paper consists of top-tier BHCs or IHCs that have \$50 billion or more in total consolidated assets on average over the most recent four quarters as reported on the FR Y-9C. The CRE panel expanded from 16 original respondents in 2010 Q4 to 33 as of 2017 Q4. The FR Y-14Q H.2 (CRE) schedule contains facility-level data on CRE loans and leases secured by real estate. The collection has grown significantly over time from approximately 50,000 CRE facilities in June 2012 to 80,000 in December 2017.

Total outstanding exposures reported in the FR Y-14Q CRE data collection were \$0.310 trillion as of 2012 Q2 and \$0.593 trillion as of 2017 Q4. The corresponding exposures for the same reporting entities in the FR Y-9C Consolidated Financial Statements for Holding Companies were \$0.454 trillion and \$0.612 trillion. Hence, aggregate FR Y-14Q CRE coverage ranged from 85% to 97% of outstanding FR Y-9C balances over this period. However, it is not possible to reconcile perfectly between these schedules due to differences in reporting requirements. For example, the FR Y-14Q schedule excludes facilities with less than \$1 million. Note that the FR Y-14-Q schedule and reporting instruction may be downloaded from the reporting forms page on the Board of Governors of the Federal Reserve website (<https://www.federalreserve.gov/apps/reportforms/default.aspx>).

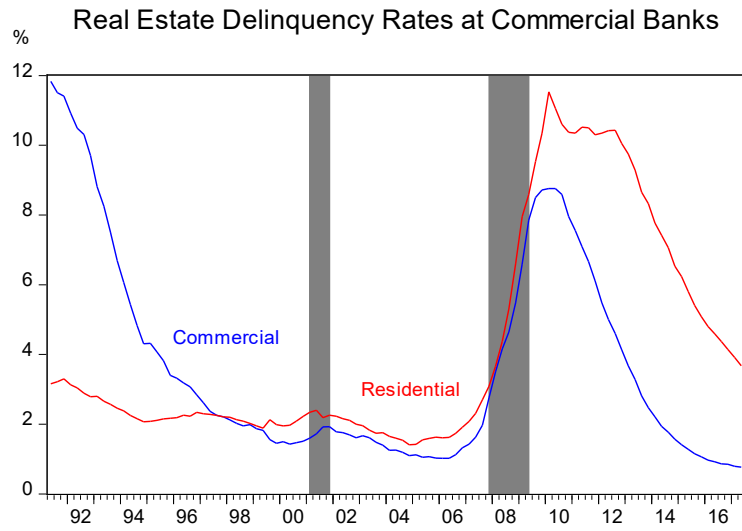
References

- Agarwal, S., Ambrose, B. W., Liu, C., 2006. Credit lines and credit utilization, *Journal of Money, Credit and Banking*, 38, 1-22.
- Araten, M., Jacobs, M. 2001. Loan equivalents for revolving credits and advised lines, *The Journal of the Risk Management Association*, May, 34-9.
- Asarnow, E., Marker J., 1995. Historical performance of the U.S. corporate loan market 1988-1993, *Commercial Lending Review*, 10, 13-32.
- Bag, P., and Jacobs, M., 2012, Parsimonious modeling of exposure at default for contingent credit lines, *The Journal of Risk Finance*, 13(1), January, 77-94. <https://doi.org/10.1108/15265941211191958>
- Bank for International Settlements, 2016. Regulatory consistency assessment programme (RCAP) – Analysis of risk-weighted assets for credit risk in the banking book. <https://www.bis.org/bcbs/publ/d363.pdf>
- Barakova, I., Parthasarthy, H., 2012. How committed are bank corporate line commitments, mimeo, Office of the Comptroller of the Currency. <https://doi.org/10.2139/ssrn.2056998>
- Barakova, I., Parthasarthy, H., 2013. Modelling corporate exposure at default, mimeo, Office of the Comptroller of the Currency. <https://doi.org/10.2139/ssrn.2235218>
- Banerjee, P., Canals-Cerdá, J. J., 2012. Credit risk analysis of credit card portfolios under economic stress conditions, working paper 12-18, Federal Reserve Bank of Philadelphia, Philadelphia. <https://doi.org/10.21799/frbp.wp.2012.18>
- Breiman, L., 2001, Statistical modeling: The two cultures, *Statistical Science*, 16(3), 199-231. <https://doi.org/10.1214/ss/1009213726>
- Board of Governors of the Federal Reserve, 2019. Comprehensive Capital Analysis and Review 2019: Assessment Framework and Results, June. <https://www.federalreserve.gov/publications/files/2019-ccar-assessment-framework-results-20190627.pdf>
- Gürtler, M., Hibbeln, M., Usselman, P., 2018. Exposure at default modeling –A theoretical and empirical assessment of estimation approaches and parameter choice, *Journal of Banking and Finance*, 91, 176-88. <https://doi.org/10.1016/j.jbankfin.2017.03.004>

- Federal Reserve System, Federal Deposit Insurance Company, National Credit Union Administration, Office of the Comptroller of the Currency, Office of Thrift Supervision, Federal Financial Institutions Examination Council State Liaison Committee, 2009. Policy statement on prudent commercial real estate workouts, October 30, 2009.
- Hahn, R., Reitz, S., 2011. Possibilities of estimating exposures, in B. Engelmann, R. Rauhmeier (eds.), *The Basel II Risk Parameters*, 2nd edition, Springer, Berlin. PDF.
https://doi.org/10.1007/978-3-642-16114-8_10
- Hon, P. S., Bellotti, T., 2016. Models and forecasts of credit card balance, *European Journal of Operational Research*, 249, 498-505. <https://doi.org/10.1016/j.ejor.2014.12.014>
- Jacobs, M., 2010. An empirical study of exposure at default, *Journal of Advanced Studies in Finance*, 1, 31-59.
- Jacobs, M., Bag, P., 2011. What do we know about exposure at default on contingent credit lines? A survey of the literature and empirical analysis, *The Journal of Advanced Studies in Finance*, 1:2(3), September, 26-46.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013, *An Introduction to Statistical Learning: with Applications in R*, Springer: New York. <https://doi.org/10.1007/978-1-4614-7138-7>
- Kuhn, M., Johnson, K., 2013, *Applied Predictive Modeling*, Springer: New York.
<https://doi.org/10.1007/978-1-4614-6849-3>
- Jiménez, G., Lopez, J. A., Saurina, J., 2008. Calibrating exposure at default for corporate credit lines, *Journal of Risk Management and Financial Institutions*, 2, 121-29.
- Jiménez, G, Lopez, J. A., Saurina, J., 2009. Empirical analysis of corporate credit lines, *Review of Financial Studies*, 22(12), 5069-98. <https://doi.org/10.1093/rfs/hhp061>
- Kim, M.-J., 2008. Stress EAD: experience of 2003 Korea credit card distress, *Journal of Economic Research*, 13, 73-102. <https://doi.org/10.17256/jer.2008.13.1.003>
- Leow, M., Crook, J., 2016. A new mixture model for the estimation of credit card exposure at default, *European Journal of Operational Research*, 249, 487-97.
<https://doi.org/10.1016/j.ejor.2015.10.001>

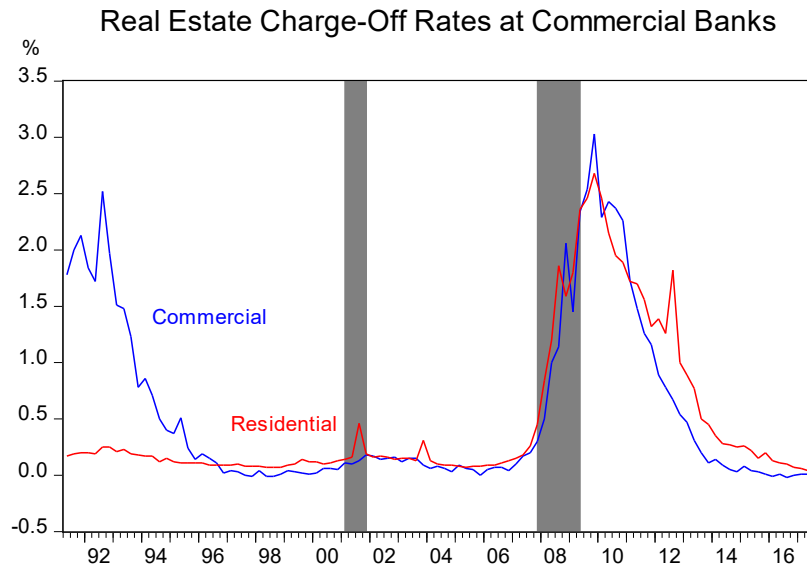
- Moral, G. (2011), EAD estimates for facilities with explicit limits, in Engelmann, B, Rauhmeier, R. (eds.), *The Basel II Risk Parameters*, 2nd edition, Springer, Berlin.
https://doi.org/10.1007/978-3-642-16114-8_11
- Qi, M., 2009. Exposure at default of unsecured credit cards, working paper 2009-2, Office of the Comptroller of the Currency, Washington, DC.
- Taplin, R., To, H., Hee, J., 2007. Modeling exposure at default, credit conversion factors and the Basel II Accord, *Journal of Credit Risk*, 3, 75-84. PDF. <https://doi.org/10.21314/jcr.2007.064>
- Tong, E., Mues, C., Brown, I., Thomas. L., 2016. Exposure at default models with and without the credit conversion factor, *European Journal of Operational Research*, 252(3), 910-20.
<https://doi.org/10.1016/j.ejor.2016.01.054>
- Valvonis, V., 2008. Estimating EAD for retail exposures for Basel II purposes, *Journal of Credit Risk*, 4(1), 79-109. <https://doi.org/10.21314/jcr.2008.069>
- Yang, B. H., Tkachenko, M., 2012. Modeling exposure at default and loss given default: empirical approaches and technical implementation, *Journal of Credit Risk*, 8(2), 81-102.
<https://doi.org/10.21314/jcr.2012.139>
- Zhao, J. Y., Dwyer, D.W., Zhang, J., 2014. Usage and exposures at default of corporate credit lines: an empirical study, *Journal of Credit Risk*, 10(1), 65-86. <https://doi.org/10.21314/jcr.2014.174>

Figure 1(a): Commercial and Residential Real Estate Delinquency Rates at Commercial Banks, 1990 Q1 to 2017 Q2



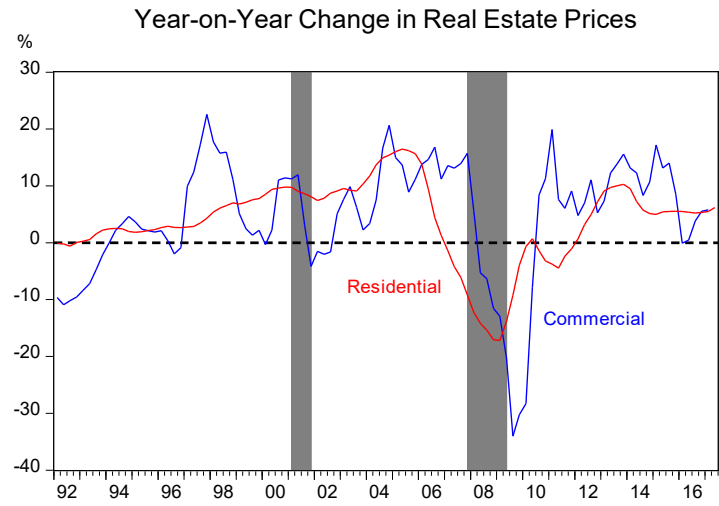
Source: FRB. Note: Data are seasonally adjusted, and NBER dated recessions are shaded.

Figure 1(b): Commercial and Residential Real Estate Charge-Off Rates at Commercial Banks, 1990 Q1 to 2017 Q2



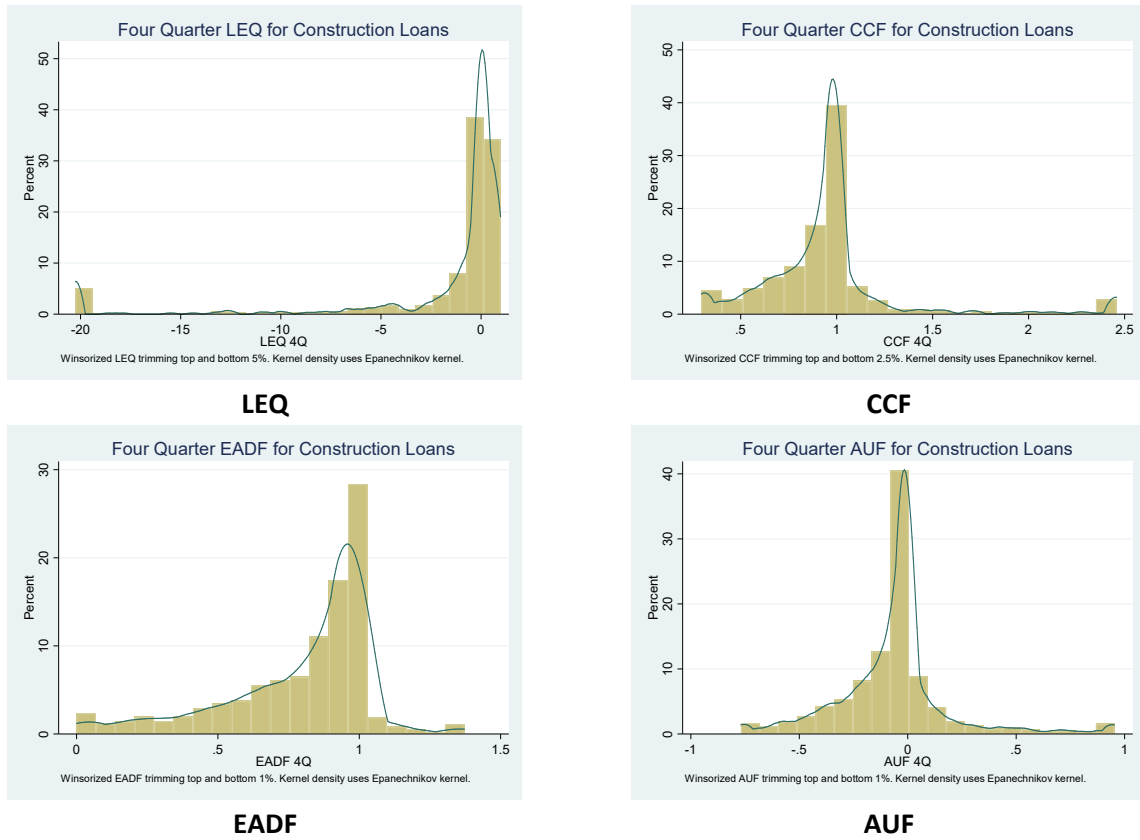
Source: FRB. Note: Data are seasonally adjusted, and NBER dated recessions are shaded.

Figure 1(c): Year-On-Year Changes in Commercial and Residential Real Estate Prices, 1990 Q1 to 2017 Q2



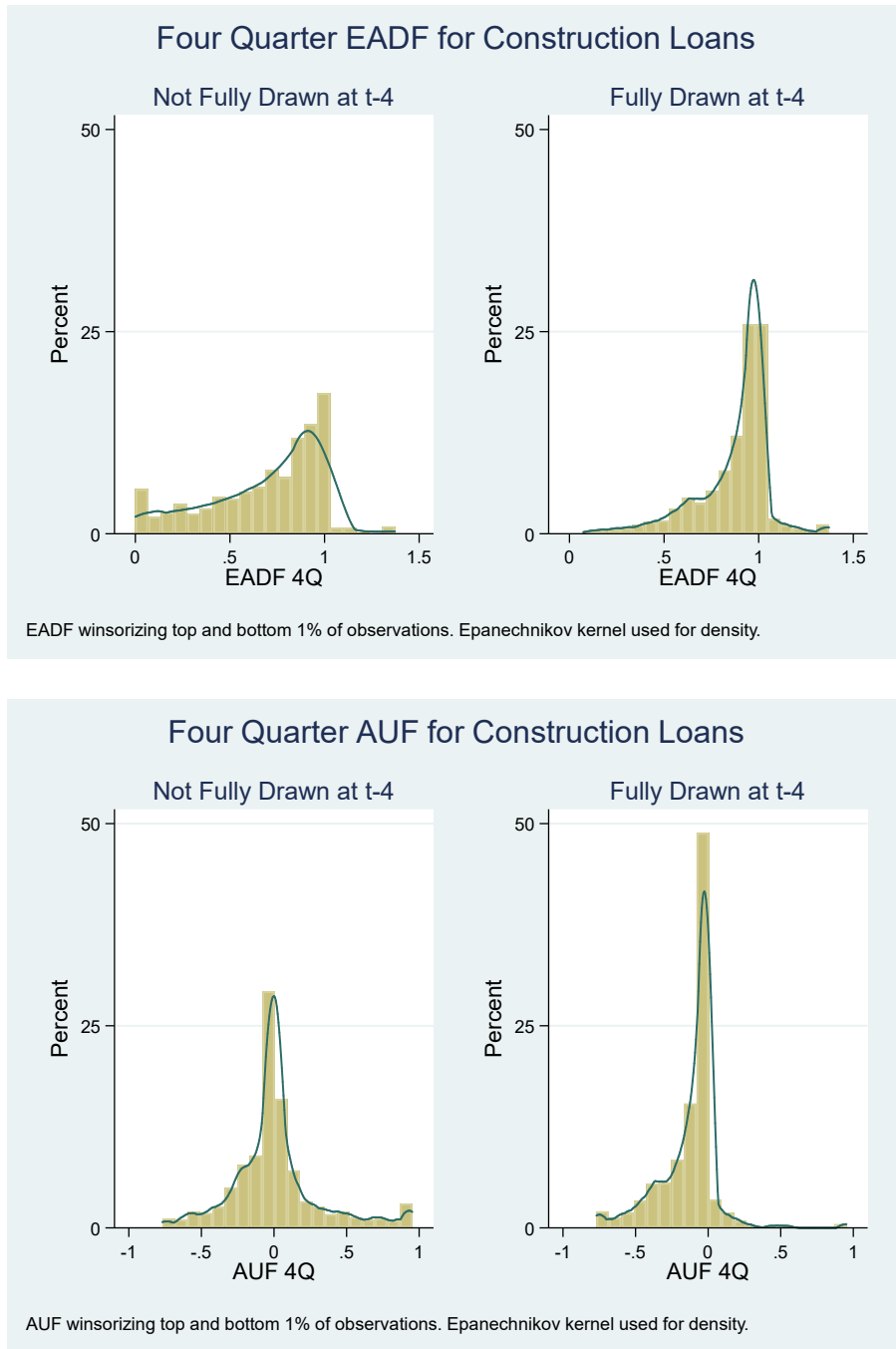
Sources: FRB and CoreLogic (single family, incl. distressed sales).
Notes: Data are seasonally adjusted and NBER dated Recessions are shaded.)

Figure 2: Histograms and Smoothed Densities of Four Construction EAD Conversion Measures
 Four-Quarter, Winsorized Measures



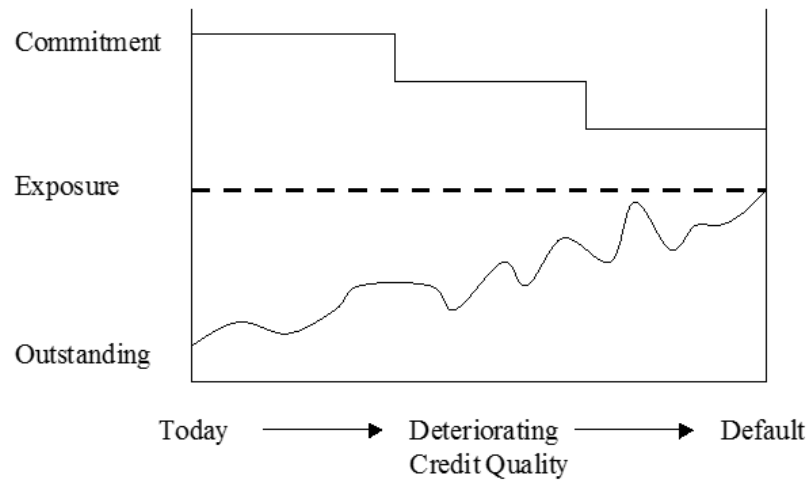
Notes: The EAD measures are winsorized, and the smoothed densities calculated using an Epanechnikov kernel (scaled to percentages). The top and bottom 5% of LEQ, 2.5% of LEQ, and 1% of the EADF and AUF observation were winsorized. Sources: FR Y-14Q H.2 and special collection data construction and land development facilities, and authors' calculations.

Figure 3: Four-Quarter Fully Drawn and Part Drawn EADF and AUF for Construction Facilities



Notes: The top and bottom 1% of four quarter EADF and AUF observations are winsorized. An Epanechnikov kernel is used for the smoothed densities (scaled to percentages). Source: FR Y-14Q H.2 and special collection data for construction and land development facilities, and authors' calculations.

Figure 4: EAD and the “Race To Default”



Source: Bank of America.

Table 1: Construction Facility Default Counts by Default Type and Year

Default Type	N	Default Year	N
90 Days Past Due	210	2010 Q4	216
Charge-Off	406	2011	564
ASC 310-10 Reserve	248	2012	515
Non-Accrual	297	2013	163
Extended and Rated "D" (Default)	922	2014	88
		2015	79
		2016	71
		2017	81
Total	1,777	Total	1,777

Sources: Federal Reserve FR Y-14 schedule H.2 and special collection data for construction and land development facilities, authors' calculations.

Table 2: Medians of EAD Conversion Measures for Defaulted CRE Construction Facilities
4 Quarter Horizon

Loan Type	Utilization	Medians			
		LEQ_t	CCF_t	$EADF_t$	AUF_t
CRE Construction	All Loan Facilities	-	0.97 (N = 1,723)	0.86 (N = 1,777)	-0.03 (N = 1,777)
	Fully Utilized at $t - 4$	-	0.94 (N = 1,019)	0.94 (N = 1,019)	-0.06 (N = 1,019)
	Part Utilized at $t - 4$	-0.04 (N = 758)	1.00 (N = 701)	0.81 (N = 758)	0.00 (N = 758)
Corporate	All Loan Facilities	-	0.99 (N = 2,972)	0.70 (N = 3,311)	0.00 (N = 3,310)
	Fully Utilized at $t - 4$	-	0.96 (N = 496)	0.96 (N = 496)	-0.04 (N = 496)
	Part Utilized at $t - 4$	0.00 (N = 2,814)	0.99 (N = 2,476)	0.64 (N = 2,815)	0.00 (N = 2,814)

Notes: $LEQ_t = \Delta U_t / (C_{t-4} - U_{t-4})$, $CCF_t = U_t / U_{t-4}$, $EADF_t = U_t / C_{t-4}$ and $AUF_t = \Delta U_t / U_{t-4}$, where t is the default quarter. The medians are weighted by commitments at quarter $t - 4$. Sources: Federal Reserve FR Y-14 schedule H.2 and special collection data for construction and land development facilities, authors' calculations.

Table 3: Do Utilization Rates Rise Prior to Default?

Loan Type	Utilization	Quarters to Default	Utilization Rates U_{t-h}/C_{t-h}			
			Mean	P10	Median	P90
Construction	Part Utilized At $t - 4$ (N = 758)	h = 0	0.84	0.52	0.95	1.00
		h = 1	0.83	0.50	0.93	1.00
		h = 2	0.80	0.44	0.90	1.00
		h = 3	0.78	0.41	0.88	1.00
		h = 4	0.75	0.30	0.86	0.99
	All Loans (N = 1,777)	h = 0	0.90	0.67	1.00	1.00
		h = 1	0.90	0.67	1.00	1.00
		h = 2	0.88	0.58	1.00	1.00
		h = 3	0.87	0.58	0.99	1.00
		h = 4	0.85	0.50	0.98	1.00
Corporate	Part Utilized At $t - 4$ (N = 2,814)	h = 0	0.63	0.08	0.70	1.00
		h = 1	0.57	0.00	0.63	1.00
		h = 2	0.56	0.00	0.62	0.98
		h = 3	0.53	0.00	0.58	0.95
		h = 4	0.50	0.00	0.54	0.92
	All Loans (N = 3,311)	h = 0	0.65	0.09	0.73	1.00
		h = 1	0.59	0.00	0.66	1.00
		h = 2	0.59	0.00	0.65	1.00
		h = 3	0.56	0.00	0.61	1.00
		h = 4	0.53	0.00	0.58	0.98

Notes: P10 and P90 are the 10th and 90th percentiles. The data are winsorized at the 1st and 99th percentiles, and weighted by the committed amounts at quarter $t - 4$. Source: Federal Reserve FR Y-14Q H.2 schedule and special collection data, and author's calculations.

Table 4: Are Credit Lines Cut Back Prior to Default?Ratio of Committed Amount at Quarter $t-h$ ($h = 0,1,2,3$) Relative to the Committed Amount at $t-4$

Loan Type	Quarters to Default	C_{t-h}/C_{t-4}			
		Mean	P10	Median	P90
CRE Construction (N = 1,777)	h = 0 (Default)	0.88	0.58	0.95	1.00
	h = 1	0.94	0.77	0.99	1.00
	h = 2	0.97	0.86	1.00	1.00
	h = 3	0.98	0.94	1.00	1.00
Corporate (N = 3,311)	h = 0 (Default)	0.89	0.50	0.95	1.14
	h = 1	0.94	0.64	1.00	1.17
	h = 2	0.97	0.75	1.00	1.11
	h = 3	0.99	0.85	1.00	1.06

Notes: P10 and P90 are the 10th and 90th percentiles. The data are winsorized at the 1st and 99th percentiles, and weighted by the committed amounts at time $t-4$. Source: Federal Reserve FR Y-14Q H.2 schedule and special collection data, and author's calculations.

Table 5: Summary Statistics for Regressions
Sample Size N = 1,777 or 1,581

Variable	Mean	Median	Std. Dev.	P5	P95
$EADF_t = U_t/C_{t-4}$	0.797	0.894	0.264	0.215	1.016
$AUF_t = \Delta_4 U_t/C_{t-4}$	-0.063	-0.036	0.262	-0.503	0.402
$\ln Facility Size_{t-4}$	15.276	15.111	1.101	13.935	17.235
$Utilization Rate_{t-4}$	0.862	1.000	0.264	0.171	1.000
$Interest Reserves_{t-4}/Committed_{t-4}$	0.203	0.000	7.364	0.000	0.054
<i>Homebuilder or Land Developer</i>	0.594	1.000	0.491	0.000	1.000
$\Delta_4 \ln House Price_t$ (County)	0.0063	0.0025	0.055	-0.074	0.098
$\Delta_4 \ln CRE Price_t$ (MSA x Property Type)	0.065	0.065	0.047	-0.011	0.145
$\Delta_4 BBB Spread_t$	0.196	0.153	0.640	-0.699	0.823

Notes: Two percent of the EADF and AUF observations have been winsorized. Facility size refers to the committed amount in millions. The utilization ratio is the ratio of the utilized balance to commitments, U_t/C_t . The homebuilder / land developer variable is a 0/1 dummy, consisting of categories 5 (homebuilders except condo), 6 (condo or coop) and 9 (land and lot development) of the property type field in the FR Y14-Q H.2 CRE schedule. The year-on-year change in the house price index is measured at the county level, while the year-on-year change in the CRE price index is measured at the MSA and property type (hotels, industrial, office, multi-family or retail) levels. The house price and CRE price data are from Corelogic and CBRE. The BBB corporate bond spread is measured relative to yield on 10 year Treasuries. The sample size is 1,777 except for the change in local house prices, where the sample size is 1,581. Sources: Federal Reserve FR Y-14Q H.2 schedule and special collection data, CoreLogic, CBRE, Federal Reserve Board and author's calculations.

Table 6: Four-Quarter Horizon EADF - OLS and Tobit ResultsDependent Variable: $EADF_t = U_t/C_{t-4}$

Explanatory Variables	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) Tobit	(6) OLS	(7) OLS	(8) OLS
<i>Ln Facility Size</i> _{t-4}	-0.043*** (7.17)	-0.040*** (7.28)	-0.042*** (6.87)	-0.040*** (6.89)	-0.043*** (6.56)	-0.041*** (7.02)	-0.053*** (5.24)	-0.034*** (4.79)
<i>PD</i> _{t-4} × <i>Ln Facility Size</i> _{t-4}	-0.003* (1.0)	-	-0.000 (0.02)	-	-	-	-	-
<i>Utilization Rate</i> _{t-4}	0.551*** (16.94)	0.540*** (17.57)	0.569*** (17.10)	0.570*** (18.30)	0.604*** (15.24)	0.558*** (15.25)	0.568*** (15.71)	-
<i>Not Fully Utilized</i> _{t-4}	-	-	-	-	-0.028* (1.81)	-0.009 (0.63)	-	-
<i>Interest Reserves</i> _{t-4} / <i>Committed</i> _{t-4}	-0.001*** (6.01)	-0.001*** (5.82)	-0.001*** (3.67)	-0.001*** (3.62)	0.005 (0.71)	-0.001*** (3.86)	0.490** (3.94)	-0.001** (2.10)
<i>Ditto</i> × <i>Not Fully Utilized</i> _{t-4}	-	-	-	-	0.457* (3.08)	0.512*** (4.71)	-	-
<i>Homebuilder or Land Developer</i>	-0.118*** (9.46)	-0.110*** (9.56)	-0.116*** (8.58)	-0.095*** (7.96)	-0.096*** (7.07)	-0.093*** (7.61)	-0.161*** (7.28)	-0.038*** (2.85)
Δ_4 <i>Ln House Price</i> _t (County)	0.434*** (4.47)	0.628*** (6.63)	0.311*** (2.94)	0.323*** (3.27)	0.355*** (3.21)	0.332*** (3.35)	0.431** (2.24)	0.266** (2.44)
Δ_4 <i>Ln CRE Price</i> _t (MSA x Property Type)	0.194 (1.47)	-	0.164 (1.23)	-	-	-	-	-
Δ_4 <i>BBB Spread</i> _t	-0.017* (1.74)	-0.015* (1.79)	-0.024** (2.37)	-0.024*** (2.59)	-0.023 (2.32)	-0.022 (2.39)	-0.026 (1.80)	-0.014 (1.61)
Bank Fixed Effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R ²	0.378	0.342	0.391	0.406	-	0.410	0.430	0.151
SE	0.207	0.213	0.205	0.203	0.055	0.202	0.232	0.151
N	1,300	1,582	1,300	1,582	1,582	1,582	676	906
Sample	All	All	All	All	All	All	Part Utilized	Fully Utilized

Notes: Two percent of the EADF observations have been winsorized. The PD variable is the bank's reported probability of default estimate, for banks using the advanced approach for regulatory capital. The intercept is not reported. Heteroscedastic robust t statistics are shown in parentheses. Upper and lower limits of 1 and 0 are used for the Tobit model, with respectively 29 and 250 censored observations. The superscripts ***, ** and * denote statistical significance at the 1, 5 and 10 percent levels respectively.

Table 7: Four-Quarter Horizon AUF - OLS ResultsDependent Variable: $AUF_t = \Delta_4 U_t / C_{t-4}$

Explanatory Variables	(1)	(2)	(3)
<i>Ln Facility Size</i> _{t-4}	-0.044 ^{***} (6.65)	-0.051 ^{***} (5.02)	-0.040 ^{***} (4.66)
<i>Utilization Rate</i> _{t-4}	-0.409 ^{***} (11.77)	-0.415 ^{***} (11.91)	-
<i>Not Fully Utilized</i> _{t-4}	-0.005 (0.33)	-	-
<i>Interest Reserves</i> _{t-4} / <i>Committed</i> _{t-4}	-0.001 ^{***} (3.29)	0.551 ^{***} (3.41)	-0.001 [*] (1.78)
<i>Ditto</i> × <i>Not Fully Utilized</i> _{t-4}	0.596 ^{***} (4.28)	-	-
<i>Homebuilder or Land Developer</i>	-0.090 ^{***} (7.46)	-0.154 ^{***} (7.29)	-0.038 ^{***} (2.76)
$\Delta_4 \ln$ <i>House Price</i> _t (County)	0.335 ^{***} (3.39)	0.430 ^{**} (2.31)	0.278 ^{***} (2.51)
Δ_4 <i>BBB Spread</i> _t	-0.021 ^{**} (2.23)	-0.026 [*] (1.84)	-0.012 (1.08)
Bank Fixed Effects	Yes	Yes	Yes
Adj. R ²	0.383	0.455	0.151
SE	0.204	0.228	0.181
N	1,582	676	906
Sample	All	Part Utilized	Fully Utilized

Notes: Two percent of the AUF observations have been winsorized. The PD variable is the bank's reported probability of default estimate, for banks using the advanced approach for regulatory capital. The intercept is not reported. Heteroscedastic robust t statistics are shown in parentheses. The superscripts ^{***}, ^{**} and ^{*} denote statistical significance at the 1, 5 and 10 percent levels respectively.

Table 8: Selected Machine Learning Based Goodness of Fit Results for EADF and AUF EAD Measures

Model / Algorithm		EADF Measure		AUF Measure	
		Training Sample	Validation Sample	Training Sample	Validation Sample
R Squared					
Parametric	Linear	0.50	0.43	0.28	0.18
	Log	0.53	0.53	0.17	0.07
	Gamma	0.44	0.40	0.28	0.19
	Tobit	0.31	0.36	0.27	0.13
Non-Parametric	Tree	0.45	0.38	0.44	0.37
	Bagging	0.89	0.45	0.88	0.43
	Random Forest	0.89	0.45	0.87	0.44
	Boosting	0.64	0.49	0.62	0.46
RMSE					
Parametric	Linear	0.21	0.21	0.20	0.20
	Log	0.57	0.52	0.54	0.56
	Gamma	0.22	0.21	0.20	0.20
	Tobit	0.24	0.21	0.22	0.21
Non-Parametric	Tree	0.19	0.21	0.18	0.19
	Bagging	0.10	0.19	0.09	0.18
	Random Forest	0.10	0.19	0.10	0.19
	Boosting	0.16	0.19	0.16	0.18

Source: Authors' calculations. Cross-validated 90% confidence intervals are available on request.