MODELING METHODOLOGY

Authors
May Jeng
Pierre Xu
Tomer Yahalom

Contact Us

Americas
+1.212.553.1658
clientservices@moodys.com

Europe
+44.20.7772.5454
clientservices.emea@moodys.com

Asia (Excluding Japan)
+85 2 2916 1121
clientservices.asia@moodys.co

Japan
+81 3 5408 4100
clientservices.japan@moodys.com

What Do 20 Million C&I Loan Observations Say about New Origination Dynamics? — Insights from Moody's Analytics CRD Data

Abstract

We construct and examine new origination of C&I loans to middle-market borrowers using the Loan Accounting System data extracted from Moody's Analytics Credit Research Database (CRD/LAS).

We find that C&I loan origination declines during the Great Recession and recovers soon after. The magnitude of the decline and the speed of the recovery varies across segments. For example, new lending to the financial industry decreases more than to the non-financial industry during the recession and recovers faster afterwards. Another example, new originations during the recession consists predominantly of short-term loans, while long-term lending becomes more dominant post crisis. This finding suggests that banks are using loan tenor as a means to mitigate risk during crises, at times even more so than credit quality.

Our analysis indicates that loan origination dynamics differ by loan type, loan tenor, borrower industry, and credit rating. The heterogeneous response of segmented C&I loan origination to macroeconomic shocks suggests the need to model loan origination separately for these segments.

We also conduct a validation of the CRD/LAS data by comparing its aggregated C&I portfolio against the aggregated FR Y-9C reports of the contributing banks and against aggregated FR Y-9C reports of all CCAR banks. Results show that the CRD/LAS dataset is representative of both the contributing banks' and CCAR banks' aggregated C&I portfolios.

Table of Contents

1. Introduction

Modeling new origination is an important part of forecasting a portfolio's future dynamics. These modeling processes are becoming more and more prevalent for capital and risk management, stress testing, and strategic planning. As the originated amount and other loan characteristics depend on the macro environment, banks have been focusing more in recent years on modeling the relationship between new origination and the macro environment. These types of origination models are important, particularly in light of recent regulatory stress testing frameworks

For example, in the United States, the Federal Reserve implemented the Comprehensive Capital Analysis and Review ("CCAR") and Dodd-Frank Act Stress Test ("DFAST") to assess the capital adequacy of large U.S. bank holding companies (BHCs). As part of the assessment, BHCs must develop models that estimate how Pre-Provision Net Revenue ("PPNR") and portfolio losses respond to various macroeconomic shocks. Portfolio loss and interest income, an important component of PPNR, strongly depend upon loan origination forecasts under different macroeconomic scenarios. Building models for new origination is often challenging due to the lack of sufficient internal and external historical data on loan origination with appropriate levels of granularity.

This paper studies C&I loan origination dynamics for different segments. We create the loan origination historical time series based on the Loan Accounting System (CRD/LAS) data extracted from Moody's Analytics Credit Research Database (CRD™). The CRD/LAS data includes quarterly, loan-level variables such as loan type, loan origination and maturity dates, interest charged and outstanding balance, as well as borrower-level information such as industry, credit rating, and size. The data covers multiple U.S. banks from 2000–2015 and allows for customized segmentation. An important question for new origination models is what type of data segmentation is important. To answer this question, we segment our data into different subsets according to borrower- and loan-level characteristics, such as credit rating and tenor. We then compare the loan origination dynamics across segments and examine whether their cyclical patterns, if any, differ sufficiently enough to warrant modeling them separately.

Note, the time span of data coverage for different banks in the CRD/LAS dataset depends on when each bank begins contributing data and when it stops. Without adjustment, when the data period for a particular bank begins or ends, the aggregated loan origination amount typically jumps or drops sharply. Similarly, when an existing bank in the dataset experiences an M&A event, the loan origination amount can also jump. In our analysis, we adjust the aggregate loan origination growth rates to remove the impact of these events — we only account for the "organic" growth of loan origination, typically what stress testing models attempt to capture. We then adjust the growth rates to create loan origination indices, the focus of our study.

While there is very limited public data to compare, we use Call Report and FR Y-9C data to show that the CRD/LAS data represents the C&I portfolio of the contributing banks well. The data also matches the aggregated portfolio dynamics for all CCAR banks. Given the dataset's size and the similarity in new origination dynamics across banks, we find that the CRD/LAS data also reasonably represents the aggregate portfolio dynamics for all banks (i.e., all depository institutions).

Results show that some segmentation is important when modeling C&I loan origination. A few examples:

1. Lending to short-term loans grows primarily during the recession and remains stagnant after the crisis. Post-crisis lending to investment-grade borrowers recovers more strongly than to non-investment-grade borrowers. When segmenting the data by both credit quality and tenor, we further find that tenor seems to explain the origination dynamics better than credit quality.

2. The financial sector experiences a stronger decline in loan origination during the crisis and stronger growth during recent years compared to the non-financial sectors. For the Mining, Oil & Gas sector, loan origination is correlates highly with the price of crude oil. This finding explains 2014's sharp drop in loan origination, unique to this sector.

3. Lending to credit lines increases significantly at the beginning of the recession. Separating term loans from credit lines allows accounting for the income from the undrawn amount.

We organize the remainder of the paper as follows:

» Section 2 describes and validates the CRD/LAS data.

» Section 3 discusses the methodology for constructing indices of loan origination amount.

» Section 4 shows the loan origination indices we construct for different segments and examines their cyclical patterns.

» Section 5 concludes.

2. Data Description

2.1 Data Source – Moody's Analytics CRD/LAS Data

We extract the data used for this study from Moody's proprietary Credit Research Database (CRD), the world's largest private firm database of middle market borrowers. It consists of data contributed by more than 75 financial institutions (including 19 U.S. Banks) and various vended data sources. Total assets of these institutions range from $20 billion to $2 trillion USD. CRD participants provide quarterly portfolio snapshots with obligor- and loan-level data.

The CRD/LAS dataset used in this study is one component of the CRD. It contains quarterly loan-level data with observations beginning in the second quarter of 2000 and ending in the fourth quarter of 2015. There are more than 50 fields for the CRD/LAS dataset. At the borrower level, the data includes size, region, rating, and industry. At the loan level, fields such as balance, commitment, tenor, and spread are available.

The CRD/LAS dataset is uniquely suitable for modeling new origination for a few reasons. First, it is large and representative. As of the end of 2015, the balance of the loans in the final data set (after data cleaning and filtering) is around 350 Billion USD, which represents about 20% of the entire US banking industry as illustrated in Figure 1. Second, the dataset comes from multiple U.S. banks, and the idiosyncratic data noise of individual banks is largely washed out when the data is aggregated. Third, it is the only dataset available in the market that includes detailed origination, as well as interest rate information, on individual C&I loans. More importantly, the dataset includes both borrower- and loan-level information, which provides flexibility in custom segmentation, based on borrower or loan characteristics. For example, the dataset contains loan type information to separate term loans from credit line, important for PPNR modeling, as capturing the dynamics of both drawn and undrawn amounts of credit lines allows forecasting the income generated from the undrawn amount. Additionally, the data covers a full economic cycle, imperative for accurately estimating the historical relationship between loan origination and the macroeconomic environment.

Figure 1   The CRD/LAS C&I Loan Balance represents a significant percentage of the National C&I Loan Balance.
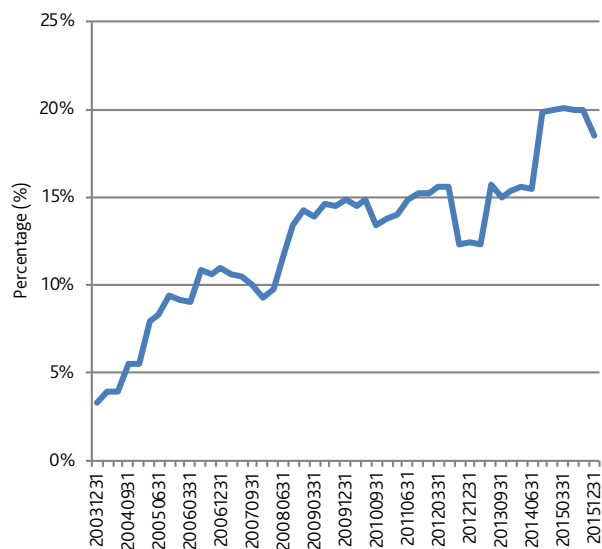
2.2 Data Cleaning and Validation

The contributing banks self-report the CRD/LAS data, and it may contain errors. To ensure the quality of the data used to construct loan origination time series, we compare individual bank's C&I loan balance seen in the dataset to the bank's C&I loan balance, reported in the FR Y-9C report.[1]

We conduct the data validation at the balance level, because there is no public data source for loan origination at the bank level. Since there may be differences between the C&I portfolio a bank contributes to the CRD/LAS dataset and the C&I portfolio the

---

[1] Since we are only interested in the C&I loans in this study, we remove from the CRD/LAS dataset non-C&I loans, such as CRE loans and loans lent to non-profit organizations.

bank reports in its FR Y-9C, we do not expect the two balances to match precisely. However, we validate that the CRD/LAS portfolio represents the bank's C&I portfolio well by making sure its balance does not differ too much from that reported in the FR Y-9C report, and that it moves together (i.e., in parallel) with the FR Y-9C balance over time, exhibiting a similar relationship with the macro environment.

We remove periods where the loan balance seen in the CRD/LAS dataset does not move in parallel with that reported in the bank's FR Y-9C report from the dataset. The resulting CRD/LAS dataset contains data from 18 contributing U.S. banks, with 670,293 borrowers and 1,398,955 facilities [2] and a total of 20,494,726 loan-level observations. Figure 2 shows the sample coverage over time, including the number of contributing banks, unique borrowers, and unique facilities.

Figure 2   Sample Coverage of the CRD/LAS Dataset.
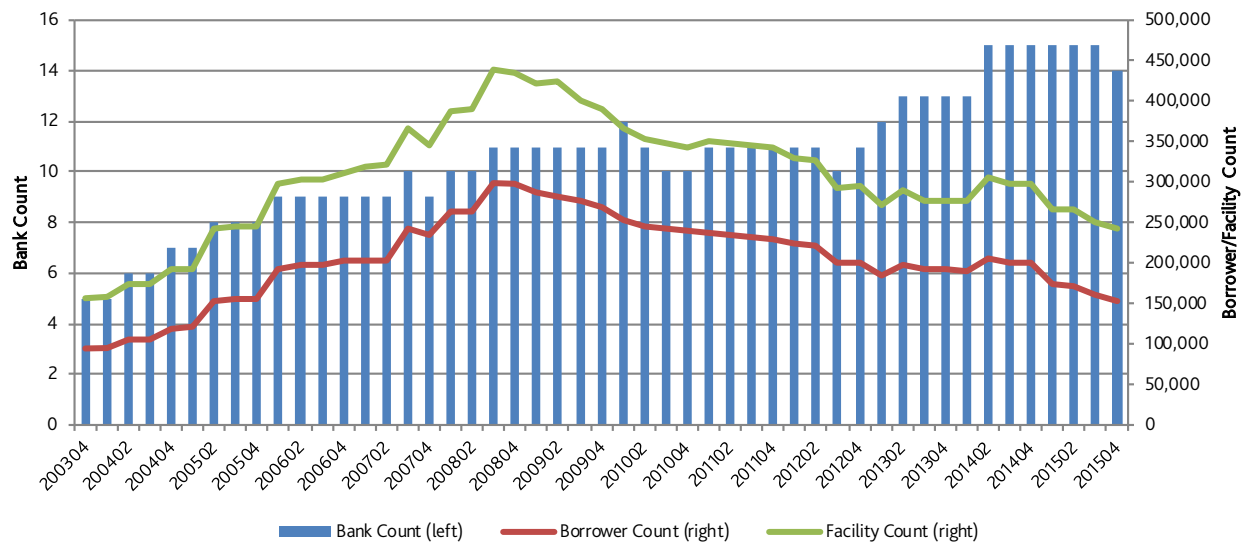
For confidentiality reasons, we cannot show individual bank's data. Instead, we construct an aggregated C&I loan balance index to illustrate how representative the CRD/LAS data is of the contributing bank's C&I portfolios. Figure 3 compares the C&I loan balance index constructed from CRD/LAS dataset with that reported in corresponding bank holding companies' FR Y-9C. In both cases, we adjust the aggregated growth rates for banks' entry, exit, and mergers and acquisitions. We can see that the indices from the two sources largely resemble each other.

---

[2] In the CRD/LAS dataset, a facility contains a set of loans, possibly with different characteristics such as loan type and maturity, under a formal agreement between the banks and the borrowers to help companies obtain operating capital. The agreement details the lending amounts, interest rate, payment terms, and conditions. A facility may apply to different projects or departments in a business.

Figure 3   CRD/LAS C&I Loan Balance is Representative of the Contributing Banks' Aggregate C&I Balance.
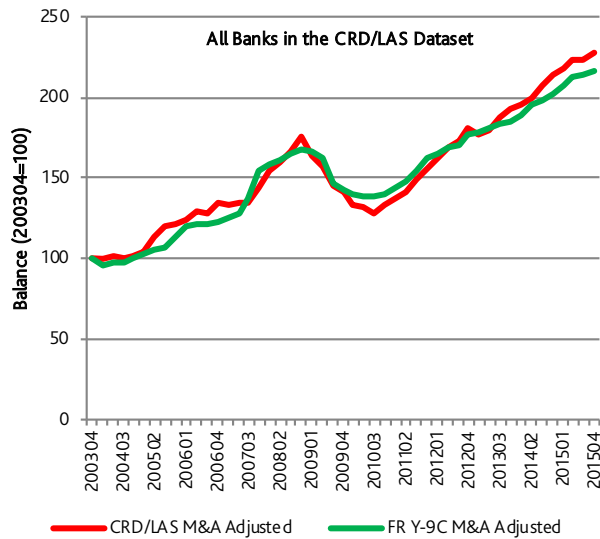
Figure 4 compares the C&I loan balance index constructed from CRD/LAS dataset with that reported in the CCAR banks' FR Y-9C. The orange line represents the index of aggregated outstanding C&I loan balance of the 33 CCAR banks. Similar to the previous comparison, we adjust the aggregated growth rates for banks' entry, exit, and mergers and acquisitions. The indices of the two sources co-move together throughout most periods, suggesting that CRD/LAS dataset represents CCAR banks' C&I loan portfolios as well. The correlation of the balance growth rate between the CRD/LAS data and the CCAR banks is 66%.

Figure 4   CRD/LAS C&I Loan Balance is Representative of the CCAR Banks' Aggregate C&I Balance.
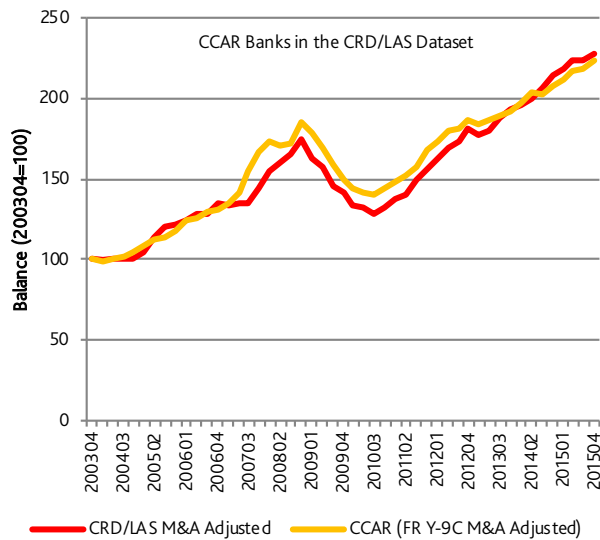
Figure 5 shows a comparison of the indices of aggregated loan balances between the CRD/LAS data and the entire banking industry. The blue line represents the index of aggregated outstanding C&I loan balances of all depository institutions (6,000+) that file Call Reports. While the CRD/LAS dataset contains only 18 banks, they appear reasonably representative of the overall industry. Note, the gap between the two indices covers 12+ years (49 quarters), so the difference over nine quarters, the CCAR projection period, is smaller than indicated in the figure. The correlation of balance growth rates between the CRD/LAS data and the entire industry is 63%.

Figure 5 CRD/LAS C&I Loan Balance is Representative of the National C&I Loan Balance.
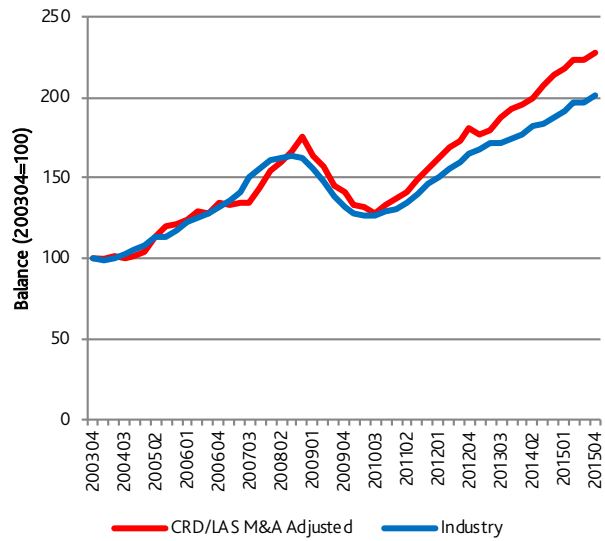
3. Methodology for Constructing Loan Origination Indices

3.1 Loan Origination Index

Banks' entries and exits can result in sudden jumps in the aggregated new origination amount. Similarly, banks' mergers and acquisitions may result in similar jumps and require adjustments. Using the unadjusted data for modeling would wrongly attribute these jumps to the macro environment and result in a misleading model. Figure 6 illustrates accounting for mergers and acquisitions. The green line represents the total C&I loan balance of one participating bank. The blue line represents the C&I loan balance excluding the loans from mergers and acquisitions. Unlike the green line, the blue line does not show a sudden increase at the time of mergers and acquisitions, especially during 2007, and better represents the intrinsic growth of the loan balance. [3]

New origination models aim to capture organic loan origination growth. We begin by estimating the growth rate in new origination during each quarter, clean of the effects of banks' entries/exits and M&A. To better illustrate the dynamics of new origination over time, we then translate these growth rates to level indices that represent the new origination dynamics for the aggregate set of banks. The index level at a given quarter t is computed as

$$Index_t = Index_{t-1} * (1 + New\ Orig\ QoQ\ Growth_t)$$

where New Orig QoQ Growth$_t$ is the quarter-over-quarter growth in the adjusted new origination amount, discussed in detail in Section 3.3. All the indices begin from 100 at Q4 2004 due to the scarce data during the early part of the sample.

Figure 6   Effect of Accounting for Mergers and Acquisitions.
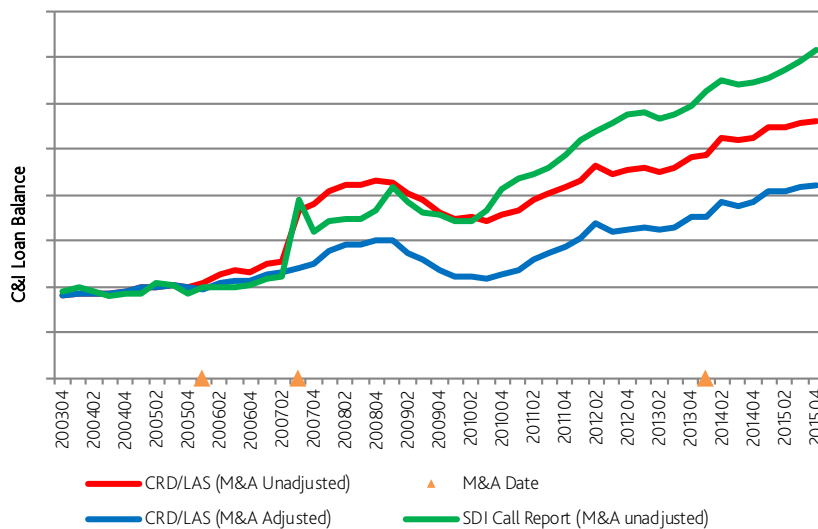
3.2 Identifying Loan Origination

In this study, a loan-level observation is identified as new origination when it represents the first appearance of the loan in the dataset and the date of this appearance is in the same or the following quarter as its recorded origination date. The qualification criteria on the consistency between first observation date and recorded origination date is placed to avoid mistaking existing loans with missing data history for newly originated loans.

We define the origination amount as the outstanding balance amount, since the origination amount is typically equal to the balance amount at origination. We aggregate the origination amount based on the time the loans originate. For example, for a loan whose origination date is in the quarter prior to the quarter of its first appearance, the origination amount is counted in the quarter of its origination date.

---

[3] We use the balance to illustrate the effect of mergers and acquisitions but the impact to the loan origination is similar.

For the purposes of studying loan origination dynamics, we filter out all snapshots of each loan, except for its original snapshot at origination. The resulting dataset contains origination data from 18 banks. The dataset includes 697,685 observations — each representing a loan snapshot.

Figure 7 shows the distribution of the size of a single loan origination. The median size of a new origination is around S100,000 USD.

Figure 7   Distribution of the Size of Individual Loan Origination.
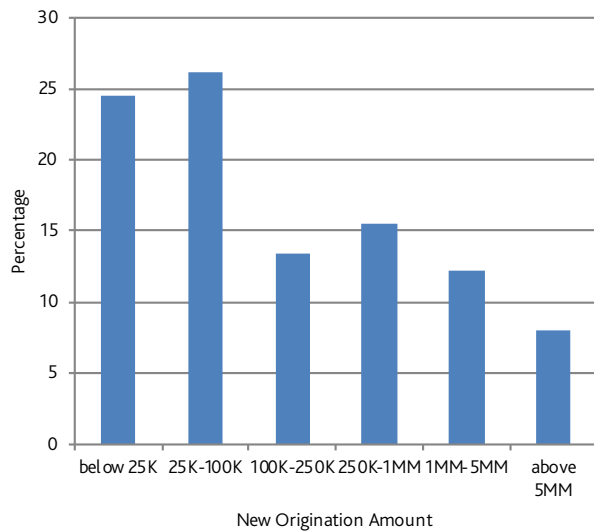
3.3 Method for Computing New Origination Growth Rates

The timespan of data coverage for different banks in the CRD/LAS dataset depends on when each bank begins submitting data and when it stops. When the data period for a particular bank begins (referred to as "bank entry") or ends (referred to as "bank exit"), the aggregated loan origination amount can exhibit a sharp jump. Similarly, when an existing bank in the dataset experiences an M&A event, the loan origination amount can also exhibit a jump. In our analysis, we adjust the quarterly growth rates of aggregate loan origination to remove the impact of these events. In other words, we only account for the "organic" loan origination growth, typically what stress testing models aim to predict. We then use the adjusted growth rates to create loan origination indices to demonstrate the origination dynamics over time.

To account for bank entry and exit, we make sure that we use the same set of banks for the beginning and end of each quarter when calculating the growth rate. Specifically, for a quarter where a bank entry occurs, we only have information for a particular bank at the end of the quarter, but not at the beginning. For consistency, we calculate the total new origination amount at both the beginning of the quarter and the end of the quarter, excluding data from the newly entered bank. We use these two values of new origination amount to compute the new origination growth rate for this quarter. For the following quarter, we already have the bank's data for both the beginning and end of the quarter, so we include the newly added bank's data in the growth rate calculation.

For a quarter where a bank exit occurs, we only have information for this particular bank at the beginning of the quarter, but not at the end. For consistency, we calculate the total new origination amount at both the beginning and end of the quarter, excluding data from the exiting bank. We use these two values of new origination amount to compute the new origination growth rate for this quarter.

Similarly, when a contributing bank has an M&A event[4] in a quarter, we exclude the bank's data when calculating the growth rate for the quarter. In addition, since we allow the first observation date to be one quarter later than the recorded new origination date when identifying new origination, we also exclude the bank's data when calculating the aggregated growth rate in the

---

[4] We obtain M&A event dates from the list of mergers and acquisitions that have occurred since 1976 provided by Federal Reserve. See "BHC Merger Data," Chicago Federal Reserve.

previous quarter. When we calculate the growth rate for the next quarter, we then include the bank's data. The step-by-step procedure for calculating the growth rate of new origination during any quarter follows:

a. Compute the total new origination amount at time t ($NewOrig_t$) and t-1 ($NewOrig_{t-1}$) using all data in the dataset during these two periods.

b. Compute the new origination amount attributed to banks whose data period starts at time t ($Entry_t$).

c. Compute the new origination amount attributed to banks whose data period ends at time t-1 ($Exit_{t-1}$).

d. Compute the new origination amount at time t ($M_{t,0}$) and t-1 ($M_{t,-1}$) attributed to banks with M&A events during time t.

e. Compute the new origination amount quarter over quarter growth (New Orig QoQ Growth) from t-1 to t as

$$\text{New Orig QoQ Growth}_t = \frac{\text{Adjusted NewOrig}_{t,0}}{\text{Adjusted New Orig}_{t,-1}} - 1 = \frac{\text{NewOrig}_t - \text{Entry}_t - \text{M}_{t,0}}{\text{NewOrig}_{t-1} - \text{Exit}_{t-1} - \text{M}_{t,-1}} - 1$$

4. Dynamics of New Origination

4.1 Aggregated New Origination

Based on the methodology described in the previous section, we construct new origination indices. Figure 8 shows the aggregate new origination index. In general, loan origination shows a significant decline during the Great Recession followed by a strong recovery. Since 2012, growth has been more sluggish.

Note, a strong seasonality effect is present in the origination growth rate. The growth rate tends to be much higher in the second and fourth quarters of each year than in the first and third quarters. To remove the seasonal variation, we smooth the new origination growth rate by using a backward-looking, four-quarter moving average value. Figure 8 also plots the resulting smoothed new origination index. We can see that both the smoothed and unsmoothed indices exhibit the same general pattern over time. Because the seasonality effect typically has little to do with the macroeconomic environment, it is usually not the focus of stress testing models. Therefore, to allow the data patterns to show clearly, the following sections only provide smoothed new origination indices, with seasonality effects removed.

Figure 8   New Origination Indices for All Loans.
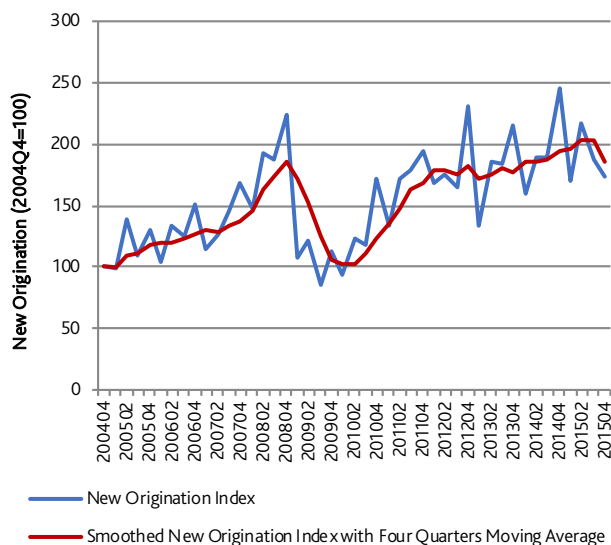
Banks use various ways to mitigate their lending risk in response to macroeconomic shocks. For example, risk can be limited by shortening the contractual maturity of the loans or by lending to safer borrowers. Demand for loans may also vary. For these reasons, we segment loans according to loan and borrower characteristics such as loan tenor, borrower's industry, and credit quality to study the dynamics of loan origination to different segments.

4.2 Segmentation by Loan Type

Term loan and credit lines are two different financing tools. With a term loan, the borrowers must repay the full amount at maturity. Lines of credit allow borrowers to draw down, repay, and redraw as needed within an agreed upon period. The repaying and re-borrowing features provide businesses with the flexibility to fund capital needs. This type of financing can be especially helpful during times of income/expense uncertainty. This motivates us to segment our sample based on loan type in order to explore whether the new origination of different loan types have different dynamics over time. Figure 9 plots the new origination indices for term loans and lines of credit.

In general, loan origination to both term loans and l credit lines grows gradually before the financial crisis, with lending to credit lines increasing more significantly at the beginning of the recession. Lending to both loan types decreases sharply during the recession and recovers after the crisis. One explanation for these dynamics may be that the flexibility of credit lines allows borrowers to obtain the funds quickly, in order to meet their short-term capital needs. A few quarters into the recession, lending to both term loan and lines of credit decreases when banks begin tightening credit.

Figure 9   New Origination Indices for Term Loans and Lines of Credit.
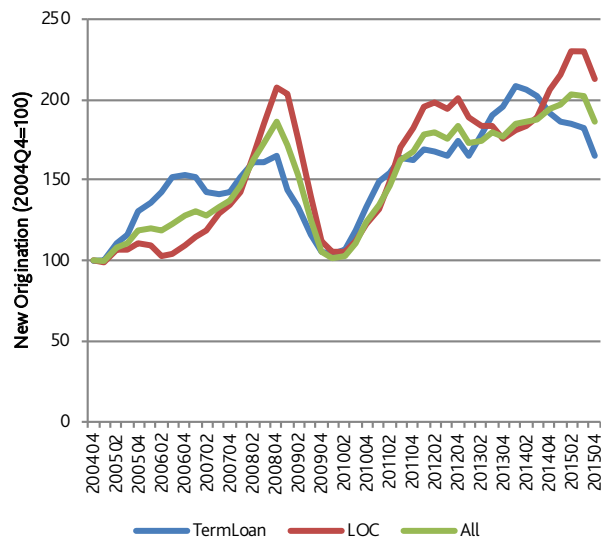
4.3 Segmentation by Borrower's Industry

The CRD/LAS dataset includes NAICS codes for borrowers, which can be used for a detailed segmentation by industry. To illustrate the importance of segmentation by industry, we examine the origination dynamics for two sectors as examples: the financial sector and the Mining, Oil & Gas sector. We compare separately the loan origination of the two sectors with the rest of the sectors.

We begin by segmenting the data into financial and non-financial sectors. This step is mainly due to the observation that some macro-environments can have a different impact on financial and non-financial firms. For example, the most recent recession was, at its core, a financial sector crisis, which propagated to other sectors. Figure 10 plots the new origination indices for loans given to borrowers in the financial sector and the non-financial sector.

In general, loan origination to both the financial and non-financial sectors begins to decline around the same time and decreases sharply at the beginning of the Great Recession. During the financial crisis, lending to both industry segments decreases, with the financial segment experiencing significantly more decline. Lending to financial firms continues to decline until Q2 2010, while lending to non-financial firms begins recovering one quarter earlier. In recent years, loan origination to financial firms has seen stronger growth.

Figure 10   New Origination Indices for Loans Lent to Financial and Non-Financial Industries.
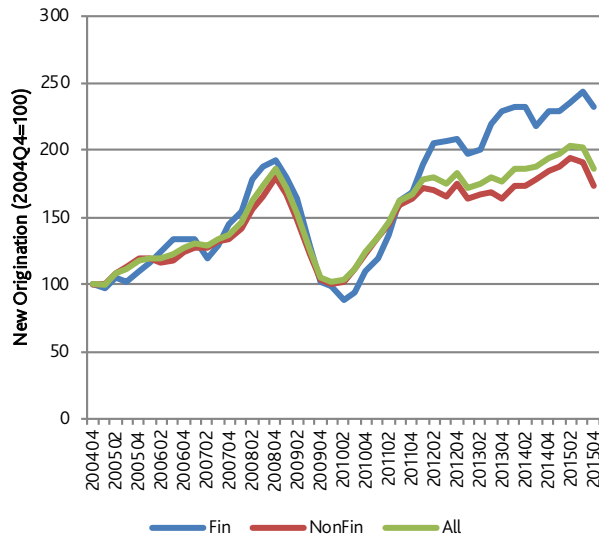
The sharp drop in oil prices around 2014 had a significant impact on several industries. While some industries benefit from lower prices, others are severely affected by it. This industry-specific reaction to oil price change is a good example for the importance of segmentation by industry when modeling new origination. We illustrate this point by comparing new origination to Mining, Oil & Gas versus all other industries.

Figure 11 plots the new origination indices for loans made to borrowers in the Mining, Oil & Gas sector and the rest of the sectors. Loan origination to Mining, Oil & Gas sector (plotted in red) is highly correlated with the crude oil price (plotted in green). The significant decline of lending in 2015 corresponds to the oil price drop. For other industries, origination remains stable after recovering from the financial crisis. Results show the unique dynamics of the Mining, Oil & Gas sector and suggest the need for industry segmentation.

Figure 11   New Origination Indices for Loans Made to Mining, Oil & Gas vs. Other Industries.
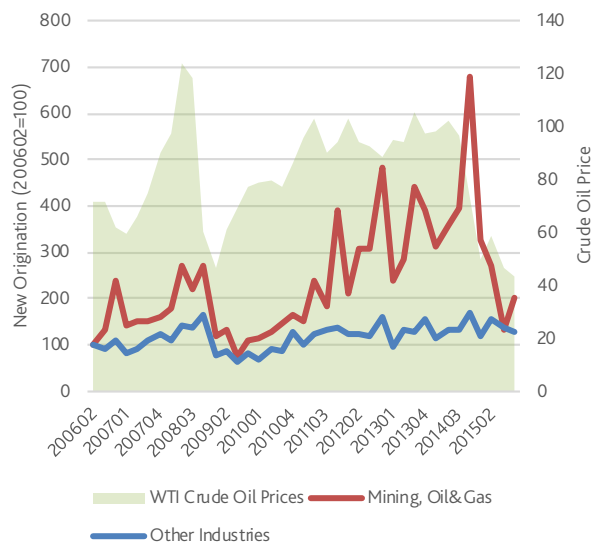
4.4 Segmentation by Tenor

Two of the main uses of C&I loans are to finance borrowers' short-term liquidity needs and to fund long-term investments. During different macroeconomic environments, the demand for short-term and long-term funding is likely different. For example, during an economic crisis, borrowers' needs for short-term liquidity may spike, which can increase origination of short-term loans. Similarly, during an economic boom, the funding needed for long-term investments may become relatively more important. In addition, banks may use loan tenor as a way to control for risk, and they are more willing to extend shorter-term loans during economic downturns. As a result, we expect origination of short-term and long-term loans to exhibit different cyclical patterns. We segment our origination data by loan tenor to explore the new origination dynamics of short-term and long-term loans over time. Specifically, we classify all loans into two segments: those with less than two-year tenor at origination (referred to as "Short-Tenor" loans) and those with more than two-year tenor at origination (referred to as "Long-Tenor" loans). Figure 12 plots new origination indices for these tenor segments.

Prior to the crisis, loan origination of the Long-Tenor segment experience steady growth while lending to Short-Tenor segment only starts to increase at the beginning of the crisis. A few quarters into the recession, both segments experience significant decline in origination with Short-Tenor segment declining slower initially. The drop is more severe for Long-Term lending. After the recession, the origination of Long-Tenor loans bounces back quickly exceeding the pre-crisis level, and then stays at a high level. In contrast, new origination of Short-Tenor loans remain stagnant post-crisis.

The different dynamics observed for short- and long-tenor lending remain consistent with the idea that in times of crisis, banks use short-term lending as a way to mitigate credit risk, as well as to meet the increase in demand for short-term borrowing from firms trying to stay afloat. In good economic times, firms look for longer-term loans to fund investments, and banks are more open to lend for longer term when they are less worried about credit risk, resulting in a sharp increase in long-term loan originations. The stark differences in origination dynamics across tenor segments suggests that this type of segmentation is extremely important to capture when modeling new origination.

Figure 12   New Origination Indices Segmented by Loan Tenor.
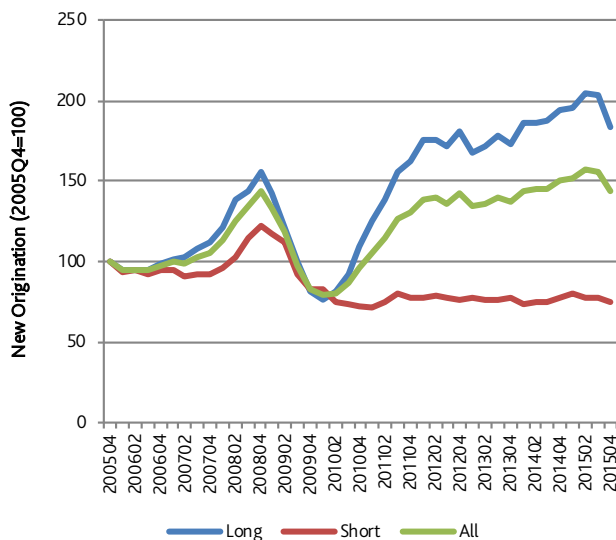
4.5 Segmentation by Credit Quality

A bank's internal rating captures its assessment of borrower credit quality. Banks use internal ratings for credit approval and portfolio monitoring. When the economy is expanding, banks tend to lend more to riskier firms to increase the revenue. During an economic downturn, banks are more concerned with credit risk, and tend to lend more to safer firms. As a result, we expect originations of different credit quality segments to exhibit different cyclical patterns. In this section, we separate our sample based on borrower's rating to study how credit quality affects new originations. Figure 13 plots the new origination indices for loans with investment-grade ratings versus those with non-investment grade ratings.

Loan origination to non-investment-grade borrowers grows faster in the years leading to the Great Recession. During the crisis, loan origination to both rating segments exhibits a sharp decline. Percentage-wise, it appears that loan origination to investment-grade borrowers declines relatively more during the crisis. We believe this sharper decrease is partly due to the effect of credit migration. During the crisis, the credit quality of many firms deteriorates, resulting in many rating downgrades from investment-

grade to non-investment grade. These downgrades make the universe of investment grade borrowers smaller, causing a further decrease in new origination to investment-grade borrowers during downturns. Last, post-crisis lending to investment-grade borrowers recovers more strongly than to non-investment-grade borrowers and passes its pre-crisis level.

Figure 13   New Origination Indices for Loans Segmented by Credit Quality.
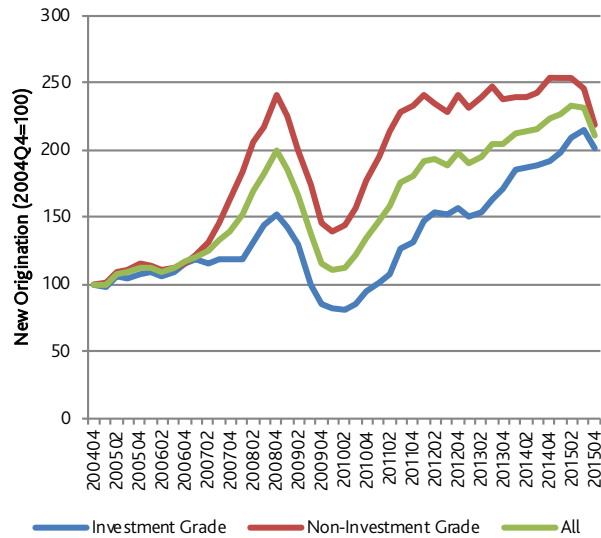
4.6 Segmentation by Both Credit Quality and Tenor

Credit quality and tenor seem to be the main two levers banks use to mitigate risk in their lending. In an effort to separate the effects of these two factors, we segment loans by both credit quality and tenor to examine one while controlling for the other.

Figure 14 plots the indices of new origination segmented by both credit quality and tenor. While both are important, it appears that tenor plays an even bigger role than credit quality in explaining loan origination dynamics. This is likely driven by banks issuing shorter-term loans during downturns as a way to mitigate liquidity and credit risk, and borrowers looking for longer-term loans in good economic times to fund longer-term investments. New origination of long-term loans (both investment grade and non-investment grade) begins to decrease at the beginning of the crisis, while lending to short-term loans declines slightly for two quarters and then drops sharply. The increase in short-term loan origination is especially stark leading into the crisis, as banks try to limit risk. After the crisis, origination of Long-Tenor loans recovers and grows faster than Short-Tenor loans. In fact, it looks like short term lending has not recovered much since the crisis. The dynamics of new origination are more similar across different credit quality segments of the same tenor, than across different tenor segments of the same credit quality. In other words, it appears that segmenting new origination by tenor explains the origination dynamics better than segmenting by credit quality alone.

Figure 14 New Origination Indices for Loans Segmented by Credit Quality and Tenor.
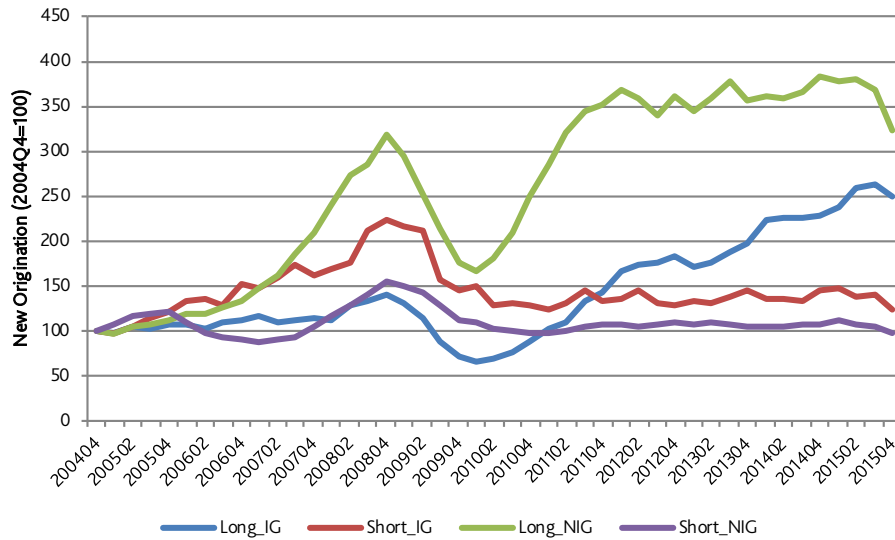
Figure 15 shows the composition of new origination across segments over time. Specifically, we compute the percentage of new origination amount attributed to each individual segment as a portion of the overall origination amount in each quarter. We can see that the portion of new origination of short-tenor loans increases during financial crisis. After the crisis, origination of long-tenor loans begins to dominate, accounting for almost 80% of total origination from Q2 2011–Q4 2015. During this period, the proportion of investment grade loans with long tenor grows from 20% at the end of 2010 to 40% as of the end of 2015. This increase in long-term lending to investment-grade borrowers post crisis is partly due to an improvement in credit quality of many borrowers during this period, increasing the overall size of the investment-grade segment.

Figure 15 Composition of Loan Origination through Time.
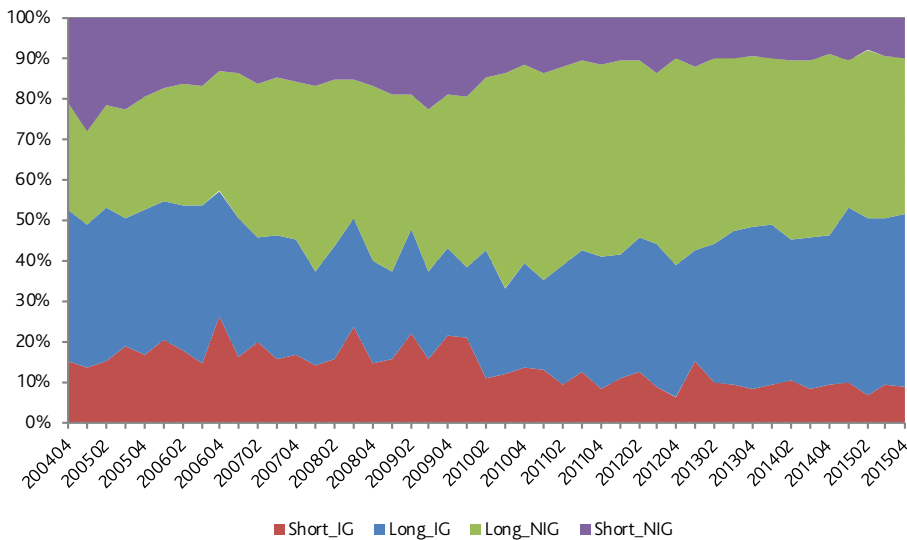
5. Summary

This study describes the Moody's Loan Accounting System (CRD/LAS) dataset and constructing aggregated quarterly indices of loan origination using the data. The CRD/LAS dataset includes loan-level and borrower-level information, which provides flexibility for customized segmentation and building granular models. Fields such as origination date and origination amount allow for better modeling of new origination. This study is the first to provide large sample dynamics of loan origination to U.S. middle-market borrowers.

We validate the contributed C&I portfolio of each bank against its C&I loan balance reported in FR Y-9C. We also validate the aggregate data against FR Y-9C and Call Reports of the contributing banks, CCAR banks, and the entire banking industry (i.e., all depository institutions), and conclude that the CRD/LAS data is highly representative of all these groups.

To illustrate the importance of data segmentation when modeling C&I loan origination, we examine the origination dynamics for a few segments. We look at segmentation by loan types, credit quality, loan tenor, and borrower's industry. In general, new origination exhibits a decline during the financial crisis and recovery after the crisis, in-line with expectations. The severity of the decline and the recovery speed varies by segment, as does the precise timing of the decline and recovery. For example, lending to the Mining, Oil & Gas sector is highly correlated to the price of crude oil, and the sector should be modeled separately from other industries in order to capture that factor.

The study suggests that tenor may be the most important dimension for segmentation. Lending to short-tenor loans does not decline as much at the beginning of the recession when origination of long-tenor loans decreases significantly. After the crisis, long-tenor lending recovers much faster, as borrowers look to fund long-term investments, and banks are less concerned with liquidity and credit risks. When segmenting the data by both credit quality and tenor, we further find that tenor seems to explain origination dynamics better than credit quality. In addition, segmenting by tenor provides better tenor estimates for newly originated loans, which leads to better runoff projections and overall balance projections.

The study clearly shows that different segments of the C&I portfolio respond differently to the macro environment, and it is exceedingly important for loan origination models to capture these differences.

References

Chicago Federal Reserve, BHC Merger Data,

https://www.chicagofed.org/Home/banking/financial-institution-reports/merger-data

Chicago Federal Reserve, FR Y-9C Data,

https://www.chicagofed.org/banking/financial-institution-reports/bhc-data

Federal Deposit Insurance Corporation, SDI data, https://www2.fdic.gov/SDI/