OakNorth
CREDIT INTELLIGENCE

Journal of Digital Banking

OakNorth Credit Intelligence

A more robust alternative to current commercial loan modeling approaches

Credit intelligence: A more robust alternative to current commercial loan modelling approaches

Received (in revised form): 30th March, 2021

Neil Kahrim
Director of Growth & Operations, OakNorth, USA

Neil Kahrim is the Director of Growth & Operations at OakNorth and is based in New York. Neil was formerly at Bank of America Merrill Lynch for almost 15 years, where he focused on leveraged finance and derivatives.

OakNorth, 445 Park Avenue, New York, NY 10022, USA
E-mail: neil.kahrim@oaknorth.com; Tel: +1 (973) 202 5829

Sean Hunter
Chief Information Officer, OakNorth, UK

Sean Hunter is Chief Information Officer at OakNorth, the creator of the OakNorth Credit Intelligence Suite. Prior to joining OakNorth, Sean was one of the first commercial engineers at Palantir Technologies in Europe, where he led trader oversight partnerships with large financial institutions, particularly Credit Suisse, which led to being co-head of the joint venture called Signac. Before Palantir, Sean was a strategist at Goldman Sachs for eight years, working in a host of areas, including equities, fixed income and algorithmic trading. Prior to GS he was IT director at a dot-com start-up, writing their initial systems and managing a growing development team through two initial public offerings (IPOs).

OakNorth, 57 Broadwick Street, London W1F 9QS, UK
Tel: +44 (0)7976 916762; E-mail: Sean.Hunter@oaknorth.com

Abstract Most commercial lending is based on a decision-making process and modelling approach largely unchanged by technology. By adopting a data-driven alternative that takes into account the fundamental differences between businesses, lenders are able to make data-driven decisions that will ultimately lead to better credit outcomes. This paper aims to briefly outline some of the limitations of the current approach to commercial lending and suggest improvements (collectively, 'credit intelligence'), taking specific note of lessons of the current COVID-19 crisis and how this has transformed the economic landscape. It also provides a case study (OakNorth in the UK) where these principles have been implemented and notes the promising results so far.

KEYWORDS: credit intelligence, credit risk, portfolio monitoring, commercial lending, commercial banking

INTRODUCTION

If a person from the 1980s was somehow transported to the present day, a lot about the world would astonish and amaze them. Our current era is nothing like the future that was imagined in popular culture at the time — there are no flying cars, we do not have the ability to teleport, and, sadly, as the last 12 months have demonstrated, we still have significant improvements to make when it comes to healthcare and social equality.

One thing, however, that this time traveller would not find surprising at all is the decision-making process used by banks when a small or medium-sized business applies for a loan.

Indeed, much about that process has changed little (if at all) over the last four decades despite the enormous advances in technology and society in that time. Banks still make credit decisions largely on the basis of intuition and experience of their credit officers, and if they use models for forecasting at all, the forecasts are based on historical data and share fundamentally flawed root assumptions: firstly, that tomorrow will be a lot like yesterday and, secondly, that most businesses are more or less alike. Over the years, this has proved to be good enough for the most part, and, as a result, these models are considered sacrosanct.

The Traditional Approach

Credit modelling in highly commoditised markets such as credit card, mortgage or auto loans is based on large volumes of historical data showing the performance of similar loans. From a pool of loans of a certain origination vintage, it is possible to observe the historical rates at which loans enter various stages of delinquency and eventual default and thereby estimate a multi-state Markov model.[1] This can then be used to simulate the trajectories of loans in a given portfolio and estimate the probability of default for these loans. The stages in the model in this case would be the states of delinquency (current; 30-, 60- and 90-days delinquent; and defaulted, for example), and the model includes a transition matrix that shows the probability of a loan in a given state moving to any other state. It would include the likelihood of a current loan becoming 30 days delinquent or a 60-day delinquent loan becoming current or defaulting, for example. In order to fit such a model, however, it is necessary to have enough loans in the pool to estimate the probability of transition for all of these states in the model. It is also necessary to be able to observe these transitions (from 30 to 60 days delinquent in this example).

For most commercial lending, however, neither of these conditions apply. Each business loan is different enough that direct comparison is challenging, and the conditions on loans do not always allow for the observation of intermediate stages of default. A loan with a bullet payment, for example, would go from being completely current to being in default with no intermediate steps.

For this reason, traditional commercial loan modelling takes a fundamental approach. The analyst constructs a financial model (usually using a spreadsheet programme such as Microsoft Excel) to simulate the cash flow, balance sheet and income statement of the business. They then project these forward for the lifetime of the loan and use assumptions to 'sensitise' or stress this model to observe the performance of the business under adverse circumstances. This allows them to see whether the business will have liquidity to continue operations and generate enough free cash flow to pay back the loan.[2]

This is augmented by peer group analysis, where a prospective borrower is compared with other similar businesses in order to establish reasonable expectations for future performance. If similar businesses have been able to generate a certain amount of profit or pay back a similar loan in the past, it is reasonable to assume that the borrower will also be able to. A bad experience with a particular type of business, however, could put the credit officer, relationship manager or the bank, generally, off the idea of ever lending to a similar business again. As mentioned earlier, however, commercial lending still relies on intuition – even if the data reveals that a business is credit worthy, a bank still may not lend to them because of a negative experience they have had in the past.

The limitations of this approach

The process of constructing both the financial model and the list of peers is subjective, and, therefore, the quality of results can vary depending on the experience and skill of the credit analyst. In many institutions the usual procedure is to take the most recent model for a similar business, copy it and make changes to update it to reflect the new business. There are, however, several challenges with this:

- It could be subject to human error by accidentally leaving in idiosyncratic features of the initial business in subsequent iterations of the model.
- There could be inconsistencies resulting from the application of different models and therefore different standards to different sectors or from using different base models for different businesses in the same sector.
- There could be concept drift, whereby the model does not account for gradual changes in the underlying industry because it is continuously being reset to a common baseline by this copying process and therefore becomes increasingly inaccurate over time.
- It is also clear that historical models can only account for conditions that are seen in the historical data. Unprecedented events, such as the COVID-19 pandemic, challenge these models to the point where their predictions break down entirely.

> *As we entered this crisis, it became clear that everything we thought we knew was proven incorrect.*[3]

What is worse is that even in normal years, these models can lack power in predicting which businesses will default owing to weaknesses in the modelling methodology,[4] leading to them being particularly sensitive to the estimation of defaults that are, by their very nature, rare. In good times, businesses default owing to idiosyncratic features of the business and in bad times because of broader market or economic conditions, leading to default correlation rising. As a result, when these approaches are wrong, they are *very* wrong, but when they are right, they are merely OK.

> *The scarcity of the data required to estimate credit risk models also stems from the infrequent nature of default events and the longer-term time horizons used in measuring credit risk. Thus, in specifying model parameters, credit risk models require the use of simplifying assumptions and proxy data.*[5]

A typical tenure for a small or medium-sized business loan would be 2.5–3.5 years, which is not long enough to capture a full credit cycle. Even for lenders with a substantial portfolio, once they start to discriminate borrowers by sector and vintage, they will find they have relatively few borrowers who are genuinely directly comparable on a like-for-like basis. While they may have a few borrowers in a certain sector, they will be in a variety of sizes and have taken out their loans at different points in their businesses evolution and the credit cycle.

Even before the onset of the COVID-19 crisis, it was clear that corporate credit markets were becoming more volatile and moving into territory that was highly unusual. This meant that models built on historical data — which were already highly sensitive to the estimation of this rare event of default — were challenged like never before.

> *Out of the ten years with the highest default rates on corporate debt since 1960, six have occurred in the new millennium and the other four in the 1990s.*[6]

These weaknesses are compounded by the fact that in the traditional process of forecasting based on historical data, a small set of models is used to provide forecasts for all businesses. The exact practice varies from bank to bank, but in some cases, there may be a common model used for all businesses in a broad sector. There may even be a single set of revenue assumptions used as a common stress scenario across all businesses — for example, a parallel shift downwards in revenue by some fixed percentage. It is easy to see the limitations of this approach, however. In times of stress, the differences between businesses become extremely important — a revenue cut that might be unthinkably severe for one sector will be inadequate to capture the severity of the downturn in another. The average will not model the situation sufficiently accurately to allow the lender to protect itself or help the borrowers most in need.

To take a concrete example, in the current pandemic, a hotel near a convention centre, catering to event attendees, may well find its revenue completely devastated by restrictions on large events, as such gatherings may be prohibited by lockdown restrictions. Conversely, a hotel in a village setting that caters to cyclists on staycations may find business booming as people seek to holiday domestically rather than travel abroad and face possible quarantines and other restrictions. Although both would be classified as 'hotels', their experiences through the pandemic are completely different.

Similarly, a Michelin-star restaurant in a city's financial district that typically caters to investment bankers, insurance companies and pricey PR firms, may well find its revenues depleted as lockdown restrictions force it to close. It is not the type of meal that many would typically order on Deliveroo, so it is unable to make up a portion of revenue online. Meanwhile, a pizza restaurant that has always centred its model around delivery and that may not even have a dine-in option, may, in fact, find it is doing more business than before the pandemic as people are spending more time at home. Again, both are classified as 'restaurants', but the experience of the one is hardly comparable to that of the other.

Now, clearly, a good deal of the outcome for any given business will still be determined by how well they themselves are able to adapt to the change in circumstances, and lenders will need to take these adaptations into account. There is no question, however, that there are structural differences between businesses that many lenders were not fully taking into account because their modelling approach or sector ontology was simply not fine-grained and specific enough to make these distinctions.

The preceding issues are only compounded by the fact that data used by banks for decision-making has an inherent lag. Economic data coming from government agencies or other trustworthy sources is often released on a monthly or quarterly cadence. Data from other providers will also typically have some delay before it becomes available. This means decisions are made on the best available data that may be from a mix of sources with different 'as of' dates. This is equivalent to trying to navigate through oncoming traffic by relying purely on your rear-view mirror.

The problems with the historical approaches can be summarised as follows: they are only using backward-looking analysis that does not account for rapid changes in the economic context; the data used to fit these models is not current; and the modelling approach itself is fragile to the number of defaults.

As these are 'tail events', models built in benign conditions may well underestimate the severity of defaults when they occur. The traditional approach to commercial lending is to assess all businesses by using the same set of simplifying assumptions. Unfortunately, this does not capture enough complexity to distinguish between the businesses that will thrive and those that will struggle to survive in adverse circumstances.

An alternative approach

It is therefore prudent to examine additional techniques to supplement historical modelling with forward-looking approaches.[7] While accepting that these methods will not be perfect (and therefore cannot be the sole basis for decision-making), the foresight gained can help shape general commercial lending policy. It can also help identify potential problems and enable lenders to be smarter in both their decisions about which loans to accept and the structuring of those loans if a decision is made to go ahead.

For example, OakNorth has developed a COVID vulnerability rating framework based on a forward-looking analysis of borrowers' vulnerability to COVID-related stress. The rating is based on a score of 1–5, 5 being the most vulnerable, and enables lenders to:

- Classify their loan book into granular sub-sectors and determine the impact of COVID-19 using 204 sector-specific domain models and forward-looking scenarios;
- Assess each sector through the three stages of the crisis — the initial impact from COVID-19, additional waves with short-term reboots in between and the new normal;
- Explore a range of outcomes based on structural changes in consumer behaviour, the regulatory impact, government fiscal stimulus and the impact of increased digital usage;
- Utilise this sub-sectoral analysis to stress test the entire portfolio simultaneously on a loan-by-loan basis and flag which individual obligors may need closer analysis and support;
- Re-underwrite loans to businesses that are vulnerable, and institute closer monitoring while helping management teams understand the stress scenarios;
- Build trust in the scenarios through regular efficacy tests incorporating concepts of 'nowcasting' and 'back-testing', which we will examine more closely later in this paper.

The resulting ratings can help lenders target specific actions at particular borrowers based on how the crisis is predicted to affect them. Borrowers that have an immediate liquidity need but have good debt capacity and long-term profitability prospects, for example, may be targets for additional lending. This will help them survive the immediate crisis and repay as long as this short-term survival is ensured.[8]

In addition to taking a prospective view, both historical and forecasting models benefit from a wider pool of data. While many banks base their model purely on borrower data, there is good evidence that drawing from a wider set of data — including macroeconomic variables, for example — can help to improve model performance.[9] Macroeconomic data can also provide vital context to credit officers, helping them to better understand the factors that drive market size, costs and revenues. This, in turn, can challenge their inherent assumptions about the overall creditworthiness of a sector.

For relationship managers, this same data can help them to better understand the issues affecting their borrower and enable them to be a better adviser to the borrower, structuring the loan to suit their specific circumstances. Models can be developed to identify headwinds in these factors and provide an early warning signal, prompting a conversation with a borrower to check on the state of their business before formal covenants have been breached. This can lead to both improved underwriting and better ongoing monitoring of loans.[10]

In addition to traditional macroeconomic data, there has been an explosion in the availability of alternative sources of data. Many businesses are seeking to monetise the data they capture in their normal course of business and couple this with data from providers who are providing what was previously unavailable or difficult to obtain. Consequently, it is possible to provide timely data-driven answers to many questions that would previously have been difficult, expensive and slow to address in analysis.

The effects of alternative data on finance were first felt when hedge funds began to use alternative sources of data to gain an edge in quantitative decision-making.[11] For example, it is possible to get restaurant reservation data from OpenTable, and, therefore, rather than wait for confirmation from borrower financials, it is possible to use this data directly to analyse the effect of the COVID-19 pandemic and the resulting lockdowns on the restaurant industry.[12]

This data has the additional benefit of likely being more timely than other sources. Banks may only receive borrower financials annually, and many traditional macroeconomic sources only come out quarterly or with a lag. It also more directly addresses the question at hand rather than having the picture obscured by the effect of other factors. Additionally, it reflects the varying actions being taken by policymakers and governments as countries and states go into different stages of lockdown. The benefit is that the lender can immediately see the impact of the current lockdown in each place. As a result, they do not need to try to keep pace with all the different evolving regulations and attempt to model these effects borrower by borrower. Naturally, it can be difficult for lenders to make use of all these sources of alternative external data when they struggle to even make the best use of their existing borrower data. This evolution is, however, essential for institutions to be able to make more intelligent commercial credit decisions.

> *For too long, many organisations in the industries suffering today have been navel-gazing at their own data with little interest in or ability to ingest the vast array of external data available.*[13]

The use of alternative data sources can also help to address the problem of lag in historical sources and help with credit decision-making. This can be augmented by 'nowcasting', which is the practice of attempting to predict the recent past, the present and the immediate future.[14] This uses the techniques of forecasting to fill in the gap between the last observable historical data points and the present time and then try to predict what is likely to happen next. Alternative data sources often have less lag than traditional sources, with many having real-time or daily updates as opposed to weekly, monthly or sometimes even quarterly for traditional sources. This means they can provide a proxy for slower-moving metrics and give banks a snapshot of the likely current state of their portfolio, which is updated very frequently. This can help to inform credit policy and enable the bank to take timely action to intervene where borrowers are heading into difficulty. It is after all, far better to take action early albeit with an element of uncertainty than to wait until the picture is fully clarified when it is often then too late to do anything useful.

Moving away from an Excel-based to a more technology-led and automated approach gives lenders the opportunity to build models that are far more specific to a given business. This is because they are accurately modelling the conditions of the business plan or capturing the nuances of a granular sub-sector. Additionally, they will be much more consistent and directly comparable across businesses. The consistency is gained because bank policy is configured in a single place and then applied across all loans, which is difficult to ensure using spreadsheets designed to be extremely flexible, at the cost of making it more difficult to apply policy consistently. This allows lenders to take a much more granular and rigorous approach to building stress scenarios, using the data to identify clusters of sectors that respond to similar macroeconomic factors, and then modelling the effects of shocks to these particular factors as the basis of the scenario.

The future of commercial lending

The ON Credit Intelligence Suite developed by OakNorth can apply proprietary stress scenarios to lenders' portfolios, assigning each borrower a vulnerability rating based on factors such as liquidity, debt capacity, funding gap and profitability. By analysing each borrower's data in the context of its geography and sector, while monitoring it against its peers using a bottoms-up approach, the software provides lenders with almost instantaneous stress-testing. This works as a strong independent challenge for their risk models and provisioning levels.

Through the continuous monitoring of active credits, lenders are able to turn monitoring — which is based on proactive alerts — into a real-time process, rather than a manual and reactive one. This not only improves credit outcomes, but also means the lenders' relationship managers can spend more time originating deals, finding the right path to 'yes' with prospective borrowers, and building deeper and more meaningful relationships with clients.

Traditionalist lenders could be gradually influenced to adopt these new methodologies by several factors: Firstly, making use of a wider pool of data sources with timelier updates to perform sectoral analysis should feel familiar insofar as it is an evolution of the approach they have been adopting all along. Secondly, the disruptive nature of the COVID-19 pandemic and the resultant policy interventions force lenders to make changes anyway — if they have to do this anyway, why not embrace the opportunity to update?

We have been deploying the ON Credit Intelligence Suite within our own bank in the UK — OakNorth — since its launch five years ago. It has enabled us to lend several billion pounds to several hundred businesses with only a cumulative ten defaults since inception and no credit losses. While it is still too early to definitively quantify the benefits of the alternative approach to underwriting and monitoring outlined above, the credit performance of OakNorth is very encouraging and suggests that the adoption of a more intelligent approach can give a lender tangible benefits to credit performance through the cycle. In 2020, OakNorth was ranked as the fastest-growing business in Europe by the *Financial Times* (FT 1000) and has performance metrics that place it among the top 1 per cent of banks globally — an ROE of 23 per cent, an efficiency ratio of 27 per cent and a borrower net promoter score of 82.

We are deploying the software not only within our own bank in the UK, but also with several other leading banks globally, including: Capital One, Fifth Third, PNC, SMBC, Customers Bank, Old National Bank, and NIBC.

As lenders embrace an approach involving data-driven decision-making and forward-looking analysis, they will in turn become better able to respond to the individual circumstances of individual business borrowers. The same data that helps them to avoid losses also helps them to lend more intelligently and ultimately serve their customers' needs better.

References

1. Gunnvald, R. (2014) 'Estimating probability of default using rating migrations in discrete and continuous time', available at: https://www.diva-portal.org/smash/get/diva2:747996/FULLTEXT01.pdf (accessed 5th May, 2021).
2. Ziegel, A. (2008) 'Corporate credit analysis', available at: http://mountainmentorsassociates.com/files/Lesson_4_Debt_Capacity_and_Cash_Flow_Analysis.pdf (accessed 5th May, 2021).
3. Demchak, Bill. (2020) 'PNC says fears for US economy prompted sale of BlackRock stake', available at: https://www.ft.com/content/da9798a6-0d53-491d-b3f7-cccfcb3f975f (accessed 5th May, 2021).
4. The problem is discussed, in Calabrese, R., Osmetti, S.A. (2012) 'Modelling SME loan defaults as rare events: The generalized extreme value regression model', *Journal of Applied Statistics*, Vol. 40, No. 6, pp. 1172–1188. https://repository.essex.ac.uk/11179/1/CALABRESE%20OSMETTI.pdf as of (accessed 5th May, 2021).
5. Nouy, D. (1999) 'Credit risk modelling: Current practises and applications', *Basle Committee on Banking*

Supervision, available at: https://www.bis.org/publ/bcbs49.pdf (accessed 5th May, 2021).

6. Becker, B., Ivashina, V. (2019) 'Disruption and credit markets', available at: https://voxeu.org/article/disruption-and-credit-markets (accessed 5th May, 2021).
7. Tapinos, E., Piper, N. (2017) 'Forward looking analysis: Investigating how individuals 'do' foresight and make sense of the future', *Technological Forecasting and Social Change*, Vol. 126, pp. 292–302, available at: https://publications.aston.ac.uk/id/eprint/31554/1/Forward_looking_analysis_investigating_how_individuals_do_foresight_and_make_sense_of_the_future.pdf (accessed 5th May, 2021).
8. Fintech Ranking. (2020) 'OakNorth's CIO Sean Hunter on how banks can use technology and manpower to excel in disbursing government assistance', available at: https://fintechranking.com/2020/04/29/oaknorths-cio-sean-hunter-on-how-banks-can-use-technology-and-manpower-to-excel-in-disbursing-government-assistance/ (accessed 5th May, 2021).
9. Tarashev, N. (2005) 'An empirical evaluation of structural credit-risk models', *BIS*, available at: https://www.bis.org/publ/work179.htm (accessed 5th May, 2021)
10. Malikkidou, D., Bräuning, M., Scricco, G., Scalone, S. (2019) 'A new approach to early warning systems for smaller European banks', available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3507506 (accessed 5th May, 2021, describes one such system for supervision of banks, however the same logic can be applied to monitoring of borrowers
11. Manning, John. (2018) 'Hedge funds see huge potential in alternative data', *International Banker*, available at: https://internationalbanker.com/brokerage/hedge-funds-see-huge-potential-in-alternative-data/ (accessed 5th May, 2021).
12. Bursztynsky, Jessica. (2020) 'OpenTable data shows the absolute devastation in the restaurant industry', *CNBC*, available at: https://www.cnbc.com/2020/04/22/opentable-data-shows-restaurant-industry-devastation.html (accessed 5th May, 2021).
13. Laney, D. B. (2020) 'Unprecedented? With data, analytics and scenario planning, there's no such thing', *Forbes*, https://www.forbes.com/sites/douglaslaney/2020/04/06/unprecedented-with-data-analytics-and-scenario-planning-theres-no-such-thing/ (accessed 5th May, 2021).
14. Science Direct. (n.d.) 'Nowcasting', available at: https://www.sciencedirect.com/topics/social-sciences/nowcasting (accessed 5th May, 2021).