The use of digital innovation for reducing SMEs' credit constraints: Hong Kong SAR's recent experience

Hong Kong Monetary Authority

Abstract

One of the key pain points that is constraining small and medium-sized enterprises' (SMEs') access to credit is their lack of sufficient and readily available financial data. Against this background, this note shares some of Hong Kong SAR's recent experience in promoting the utilisation of alternative data in the banking sector to help alleviate the financial data problem in SME lending. The note further presents some recent digitalisation initiatives undertaken in Hong Kong, including the launch of a new data infrastructure platform by the Hong Kong Monetary Authority (HKMA) that could expedite the adoption of alternative data in Hong Kong and thereby enable banks to enhance their credit underwriting processes for SME borrowers.

JEL classification: G21, G28, G32, L15, O33.

Keywords: alternative data, digital innovation, small and medium-sized enterprises, credit scoring.

1. Introduction

Small and medium-sized enterprises (SMEs) are a key contributor to economic activities. According to some estimates (Cornelli et al (2019); Nemoto and Yoshino (2019)), SMEs are estimated to account for over 95% of the number of firm establishments, contribute to 50–70% of employment and constitute 30–60% of gross domestic product (GDP) in various Asian economies. Despite their importance to the economy, SMEs often face difficulties in obtaining external finance, particularly bank, credit. One of the key challenges facing SMEs is the lack of sufficient credit history and readily available financial records. Without such data, it may be difficult for banks to assess the creditworthiness of SMEs based on a conventional credit scoring approach. As a result, SME loans are often subject to higher interest rates and/or are prone to more stringent collateral requirements.

With the rapid advancement in financial technology and greater adoption of technology around the world in recent years, various forms of digital information pertaining to customers and enterprises have become available. The emergence of these digital data, combined with the enhancement in computing capabilities, could open up new alternative ways for banks to enhance their credit assessment and may consequently help improve the provision of financial services to SMEs.

Against this background, this note presents some of Hong Kong SAR's recent experiences in promoting the utilisation of such alternative data in the banking sector

with the aim of helping to improve SMEs' access to credit. Regarding the structure of this note, it first presents the notion of alternative data for credit scoring and then provides evidence on the technical feasibility of utilising such alternative data in banks' SME lending decisions. Finally, the note further presents the recent new data infrastructure initiative undertaken by the Hong Kong Monetary Authority (HKMA) that aims to enable secure and efficient exchange of owner-consented data between data providers and financial institutions in Hong Kong to facilitate a wider adoption of alternative data in practice.

The discussions of this note draw mainly on the findings from various recent studies, including a report by the Hong Kong Applied Science and Technology Research Institute (ASTRI) and commissioned by the HKMA titled "Alternative credit scoring of micro-, small and medium-sized enterprises" (HKMA and ASTRI (2020)), and an applied research report by the Hong Kong Institute for Monetary and Financial Research (HKIMR) titled "The digitalisation of financial services in Hong Kong: recent experience, regulatory developments and considerations for sustainable innovation and growth" (HKIMR (2023)).

2. Utilising "alternative data" to alleviate the financial data problem in SME lending

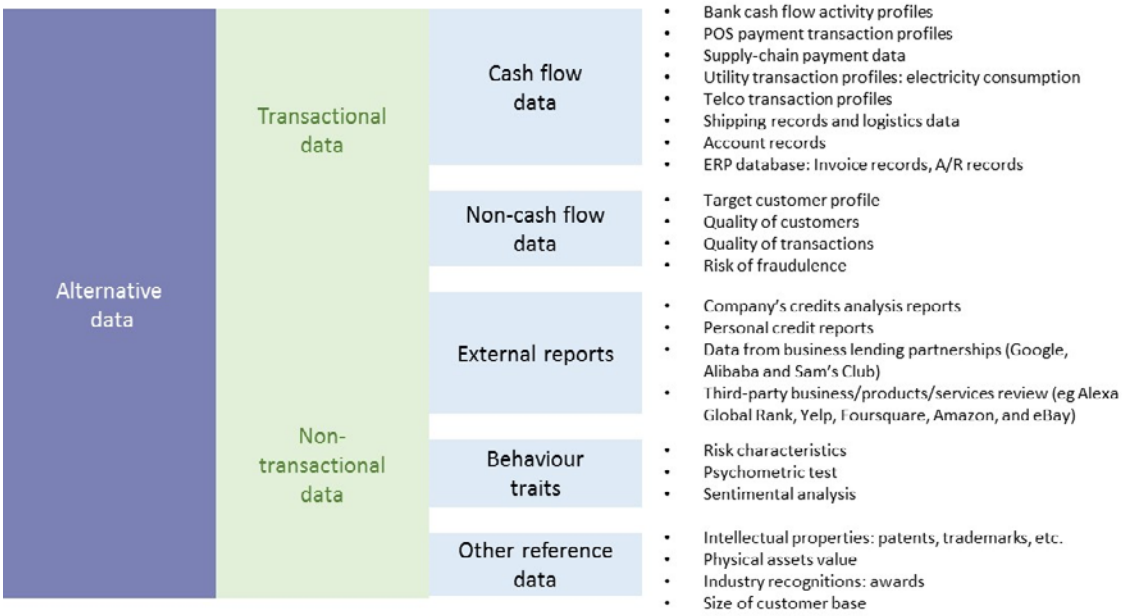
Typically, when a bank underwrites a loan, its lending decision will generally be based on the assessment of borrowers' creditworthiness using a credit scoring approach. Conventionally, such an approach relies on financial ratios or indicators derived from borrowers' audited financial statements. Without sufficient and reliable financial data on SMEs, it becomes more challenging for banks to adequately assess their credit risk under such a conventional approach and this may thus constrain banks' willingness to provide credit to SMEs.

To tackle this challenge, many financial institutions around the world are exploring the use of alternative forms of data to help assess SMEs' loan repayment ability. With the rapid advancement in technology and a wider adoption of digital innovations in the past decade, customers' and firms' digital footprints are becoming more available. The increasing availability of alternative data could provide additional supporting information that may facilitate banks' lending decisions even in the absence of conventional financial data on borrowers.

Before proceeding further, it is useful to first describe the notion of alternative data for credit scoring. These data can be broadly classified into two groups, namely: (a) transactional data and (b) non-transactional data (Figure 1).

Classification of alternative data

Figure 1



Source: HKMA and ASTRI (2020).

"Transactional data" generally refers to the records of business activities between a company and its customers. These usually include revenue-related information (ie cash flow data) and non-monetary-related information (ie non-cash flow data). For the former, these would include firms' customer payment transaction data, which reflect the sales activities of a retail SME; and, for SMEs in trading and logistics industries, supply chain payment profiles, which contain their payment records to their suppliers. For non-cash flow data, these are transactional records, such as payment identifications and transaction patterns of customers with the SMEs. These data can be used to capture information about the quality of their customers and their target customer profiles, which can in turn reflect their business prospects.

"Non-transactional data" generally refers to data from third-party data providers (such as a credit reference agency), which may supply the missing data or add new information for the credit scoring model. For instance, these could include company credit analysis reports as well as the credit history of the SME's owner, etc.

These diverse types and forms of alternative data for credit scoring can provide lenders with more insights into SMEs' creditworthiness and facilitate their continuous monitoring of borrowers. Indeed, with the rapid advancement in computational power and analytical capabilities that help process and analyse these non-standardised data, the value of a non-traditional/alternative approach for evaluating the creditworthiness of SMEs is gaining recognition in both developed and emerging market economies throughout the world. For instance, some major banks in Japan have already taken into account SMEs' cash flow data, such as bank transactional account information, in their credit scoring models (Nemoto and Yoshino (2019); HKMA and ASTRI (2020)). In addition, some mid-tier and cross-regional commercial banks in China are currently using data such as income tax and business tax records

for credit scoring. These examples highlight the potential usefulness of adopting an alternative credit scoring approach for SME loan applications.

3. Some evidence on the feasibility of an alternative credit scoring approach for SME lending

While alternative data are considered to be a useful source of information that helps lenders examine the financial condition of a borrower, it is important to assess the technical feasibility of using such data for assessing the credit risk of SME borrowers and their predictive performance. In this regard, in 2020 the HKMA commissioned ASTRI to conduct a study that explored the use of machine learning algorithms and alternative data for SME lending businesses.[1]

Two proof of concept (PoC) experiments were conducted by a bank and a third-party data provider in Hong Kong, respectively. In each of these experiments, the participants separately developed an alternative credit scoring model based on various types of transactional cash flow and non-cash flow data for SMEs in the retail industry.[2] The objective of these experiments was to examine the technical feasibility of applying SME transactional data to develop machine learning credit scoring models. The key findings and insights from these PoC experiments are summarised below.

A. Transactional cash flow model developed by the participating bank

The first PoC experiment was conducted by a bank. The participating bank used the historical monthly transactional cash flow data (including both inflows and outflows) of its SMEs' bank accounts to develop a transactional cash flow model that aimed to predict whether a default event would occur within a specific period.

The data set contained around 74 million monthly observations of more than 1,000 SMEs for the period from October 2018 to July 2020. Each observation represents the monthly data from an SME's bank account. The data set contained six main variables, including the occurrence of delinquency (ie the target variable), total cash credit/debit amount, number of credit/debit transactions, and number of years of bank-client relationship. Based on these key variables, extra variables were derived using different statistical transformations. To examine the performance of the machine learning models, Table 1 shows the area under curve (AUC) scores of the nine selected machine learning algorithms across different prediction horizons (ie one month ahead, two months ahead and three months ahead).[3]

[1] For details of the commissioned study, see www.astri.org/wp-content/uploads/2016/11/White-paper-on-credit-scoring.pdf.

[2] There was no sharing of data among the participating organisations, and the machine learning models involved in these PoC experiments were developed based on the transactional data of each organisation's data set.

[3] Usually, the AUC is equal to 0.5 for a random model and approaches 1 as a model approaches perfection.

Experiment results of the transactional cash flow model Table 1

Prediction horizon AUC scores	One month ahead	Two months ahead	Three months ahead
Logistic regression	0.9162	0.8098	0.7592
Random forest	0.9249	0.8312	0.7606
Extra trees	0.9198	0.8287	0.7727
LightGBM	0.9141	0.8187	0.7987
CatBoost	0.9339	0.8518	0.7641
XGBoost	0.9368	0.8436	0.7926
KNN	0.7783	0.7699	0.7681
CNN	0.8489	0.777	0.7265
Stacking	0.9227	0.8103	0.7515

Source: HKMA and ASTRI (2020).

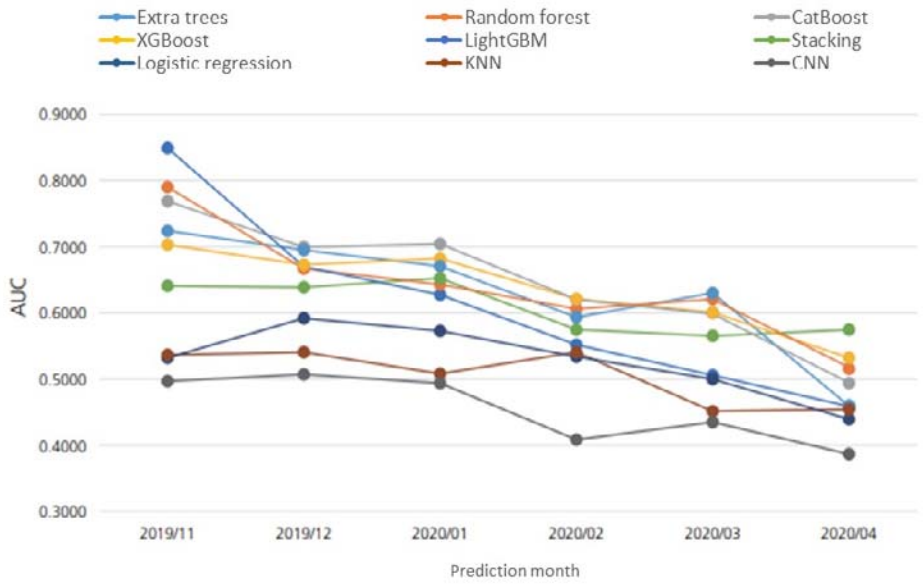
Overall, the results suggest that transactional cash flow models can achieve desirable performance for short-term default prediction. For instance, the AUC scores for most of the models (such as XGBoost, CatBoost and random forest) are found to be over 0.91 for one-month-ahead default prediction. Such a desirable performance suggests that banks can in practice update the credit risk assessment prediction results for SMEs on a regular basis with incoming transactional data (ie bank statements) from loan applicants. This may also enable banks to achieve continuous monitoring of SMEs' creditworthiness.

B. Model developed by a third-party payment data provider

The second PoC experiment was conducted by a third-party point of sale (POS) payment data provider using non-cash flow transactional data from SMEs. The data set contained nearly 1.55 million observations of more than 300 SMEs over the period from April 2019 to April 2020. Each observation represents an SME payment transaction, which includes the occurrence of late payment of any service charges, company type, transaction date and time, currency, payment amount and payment method. As the data set does not contain SME loan default information, information on service charge payment delinquency by SMEs was used to proxy their default status. As such, the machine learning model developed can be treated as a pre-screening model which can identify SMEs with potential loan default problems.

The pre-screening models were developed using the monthly transaction data of the POS payment data provider from April to October 2019. Figure 2 shows the AUC scores of the nine selected machine learning models for predicting the occurrence of late service payments.

Performance of the pre-screening models by POS payment data provider Figure 2



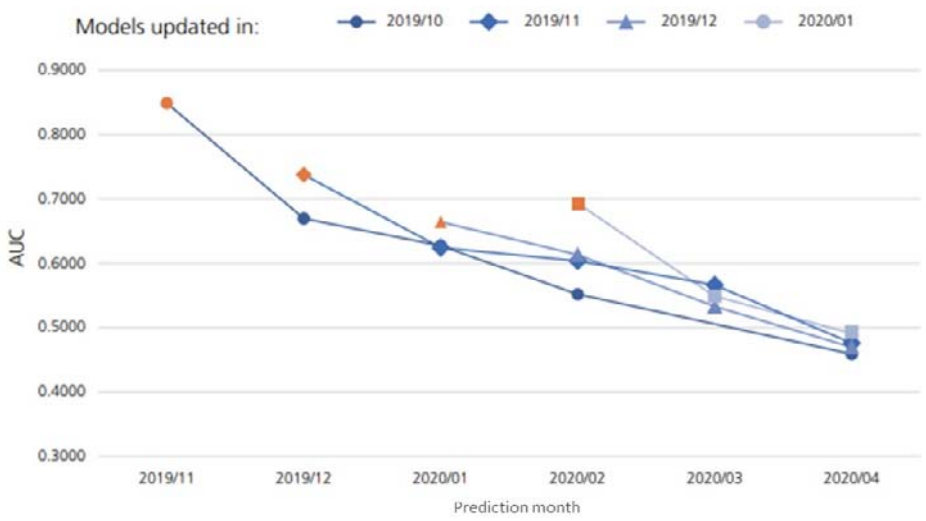
Source: HKMA and ASTRI (2020).

As shown in Figure 2, the top five pre-screening models are found to achieve reasonably good predictive performance for the short-term prediction horizon. Similar to findings in the previous PoC experiment undertaken by the participating bank, the accuracy of the pre-screening models dropped gradually for longer prediction horizons. Nonetheless, the AUC could be improved by retraining the model with the updated information. Taking the "extra trees" pre-screening model as an example, Figure 3 shows the AUC scores of the pre-screening model would be higher after subsequent model retraining.

In summary, given that both the transactional cash flow model and the non-cash flow pre-screening model achieved a reasonably good predictive performance for assessing SME credit risk, the banking industry can leverage on alternative data in developing an effective alternative credit scoring framework for SMEs. This would help address the financial data availability problem in SME lending.

Performance comparison for "extra trees" pre-screening model with and without retraining

Figure 3



Each line in Figure 3 presents the AUC scores of the "extra trees" pre-screening model that was updated (ie underwent model retraining) in a specific month. The four lines show the prediction results of the model after updating in October 2019, November 2019, December 2019 and January 2020.

Source: HKMA and ASTRI (2020).

4. Supporting a wider adoption of alternative data by banks – the implementation of Commercial Data Interchange (CDI)

While the PoC evidence in the previous section illustrates the benefits and the technical feasibility of adopting an alternative credit scoring approach, there are several operational challenges (such as data privacy, a safe and secure environment for data exchange, data protection, etc) that need to be addressed in practice in order to expedite the adoption of alternative data in the banking industry.

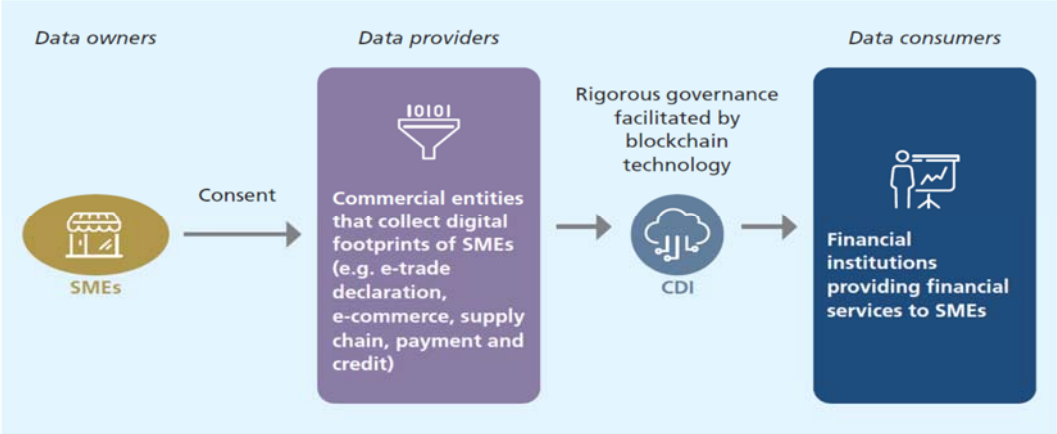
To enable scalable, efficient and secure flows of data between banks and data providers, in October 2022 the HKMA launched a new market development initiative called Commercial Data Interchange (CDI). As one of the initiatives under the HKMA's "Fintech 2025" strategy, CDI is a consent-based data-sharing infrastructure with a standardised and secure technical interface. Banks and data providers can connect to the interoperable platform to share commercial data and use this data to offer better

products for their clients. The CDI infrastructure is designed to foster a vibrant and trusted data-sharing ecosystem in the industry.[4]

A promising CDI use case is to facilitate the credit assessment process for lending to SMEs. Figure 4 presents a brief description of how CDI facilitates SMEs' loan applications. With the consent of the data owner (eg an SME), the CDI platform allows banks to obtain commercial data (such as e-trade declarations, e-commerce, supply chain, payment and credit data) relating to the concerned enterprise from third-party platforms in order to conduct more objective and accurate credit risk analyses such as those mentioned in the previous section. The access to alternative data via CDI may reduce the need for enterprises to provide collateral and effectively enhance their access to bank credit and other financial services.

How CDI facilitates SMEs' loan applications

Figure 4



Source: HKIMR (2023).

In addition, the connection between CDI and the Commercial Credit Reference Agency (CCRA) has recently come online, enhancing the availability of non-transactional SME data for banks. Indeed, several CDI participating banks have successfully connected to the CCRA via CDI and gained access to machine-readable corporate credit reference data. This may enable banks to carry out more automated customer onboarding, credit approval and ongoing credit review and monitoring of their customers.

As it has only been launched recently, it could be premature to thoroughly evaluate the effectiveness of CDI for improving SME credit access. Nonetheless, there is already some anecdotal evidence on its benefits. Since the launch of CDI in October 2022, the number of participating banks has reached 26, and key data providers have doubled, from six to 12. The CDI utilisation rate has also increased nearly ninefold, to 8,900 loan approvals from close to 1,000 during the pilot phase, with cumulated credit approvals exceeding HK$ 8 billion as of end-September 2023. It is also reported that

[4] For details of CDI, see the dedicated CDI website (cdi.hkma.gov.hk). For details related to CDI governance and controls, see (cdi.hkma.gov.hk/wp-content/uploads/2022/10/CDI-Framework-2022-Oct.pdf).

some successful SME loan applicants were able to obtain credit from banks via CDI with less stringent collateral requirements and more competitive financing terms. In light of these experiences, some interviewed market participants felt that CDI could help address the financial data availability problem faced by SMEs (HKIMR (2023)).

Going forward, to further unleash the potential of CDI the HKMA has planned to diversify CDI data sources with the introduction of new data providers covering government, logistics and catering services data.[5] For instance, the Companies Registry became the first government data source to be connected to CDI via the government's Consented Data Exchange Gateway at the end of 2023. The connection with the Companies Registry will provide banks with automated access to key data needed for their know-your-customer (KYC) processes, such as company name, company number, individual directors, shareholders and members, address, country and date of incorporation. Easier access to these data should help banks to further digitalise and streamline their internal risk management processes.

5. Conclusion

This note shares some recent Hong Kong experience on the use of digital innovations in addressing some of the long-standing pain points in SME financing. Specifically, the launch of CDI, a consent-based data infrastructure, is enabling SMEs to share their commercial data from different data sources with banks in a secure and efficient manner. This could enable banks to leverage various sources of alternative data to enhance their credit underwriting processes for SME borrowers, thereby potentially enhancing SMEs' access to financial services.

It should be noted that this is only one of the many plausible examples of how digital innovations can help improve SMEs' access to credit. Besides the launch of CDI, other innovative projects are also under way in Hong Kong and globally that aim to resolve other key pain points in SME finance. For instance, Project Dynamo, a collaborative work between the BIS Innovation Hub and the HKMA, explores the use of blockchain technology and digital trade tokens (DTTs) as ways to encourage institutional investors to finance SMEs. The project has developed a prototype platform which has demonstrated that DTTs can be used as one innovative solution to address some of the common pain points in SME trade finance.[6] Therefore, it is widely believed that there remains much room for financial technology to play in this area going forward. To effectively harness the transformative potential of digital advancements, more research and collaboration between market participants and policymakers are needed, both domestically and internationally.

[5] See Yue (2023).

[6] In brief, the use of DTTs allows the anchor buyer to send a smart contract-backed conditional payment to their SME suppliers. Before the conditions are met, suppliers can pass the DTT to their upstream counterparts to offset their debt, or to institutional investors to obtain working capital. For details, see www.bis.org/about/bisih/topics/open_finance/dynamo.htm.

References

BIS Innovation Hub and Hong Kong Monetary Authority (2023): *Project Dynamo: catalysing innovation for SME growth*, BIS Innovation Hub project report, June.

Cornelli, G, V Davidson, J Frost, L Gambacorta and K Oishi (2019): "SME finance in Asia: recent innovations in fintech credit, trade finance, and beyond", *ADBI Working Paper Series*, no 1027, October.

Hong Kong Institute for Monetary and Financial Research (2023): "The digitalisation of financial services in Hong Kong: recent experience, regulatory developments and considerations for sustainable innovation and growth", *HKIMR Applied Research Report* no 2/2023, July.

Hong Kong Monetary Authority and ASTRI (2020): "Alternative credit scoring of micro-, small and medium-sized enterprises", white paper, November.

Nemoto, N and N Yoshino (2019): *Fintech for Asian SMEs*, ADB Institute, May, ISBN: 978-4-89974-109-1.

Yue, E (2023): "Keynote at the Hong Kong FinTech Week 2023", 2 November, www.hkma.gov.hk/eng/news-and-media/speeches/2023/11/20231102-1/.