UNIVERSITY OF TECHNOLOGY SYDNEY

Sandra Martin
Executive Manager
Research Office
15 Broadway, Ultimo NSW 2007

Email: Sandra.Martin-1@uts.edu.au

PO Box 123
Broadway
NSW 2007 Australia
www.uts.edu.au

UTS CRICOS PROVIDER CODE 00099F

Chris Dalton
Chief Executive Officer
The Australian Securitisation Forum
Level 7, 14 Martin Place
Sydney NSW 2257

27 August 2025

By email: cdalton@securitisation.com.au

Dear Chris,

RE: COMPLETION OF PROJECT

RISK FACTORS AND GROWTH BARRIERS FOR LENDING TO SMALL AND MEDIUM-SIZED ENTERPRISES

On behalf of UTS I confirm that this project has been completed.

I understand our final report will be published on your website and we would respectfully request that you include the following acknowledgement for us and I note it also includes the wording for acknowledgments of our third party data providers.

"This paper is dedicated to the memory of Professor Harald (Harry) Scheule, who served as the project lead and whose invaluable guidance and expertise greatly shaped the direction and impact of this work. The authors are grateful for the financial support of the Australian Securitisation Forum (ASF) and the feedback provided by the ASF Industry Advisory Group. This research acknowledges the data provided by the Australian Taxation Office (ATO), Australian Bureau of Statistics (ABS) as well as the Australian Financial Security Authority (ASFA). We thank Ignatius McBride, Mark Wicht and Lan Dang at AFSA for providing analytics."

Kind regards,

Sandra Martin

Sandra Martin

Executive Manager

Research Office

UTS File No: PRO23-16622

Predicting credit risk for small and medium-sized businesses in Australia

Harald Scheule, Chung Mai, Anh Nguyen[1]

Executive summary

The paper provides descriptive statistics for failure/insolvency rates and resolution times for Australian firms. Models for failure probabilities and resolution times are estimated and fitted for the universe of Australian firms, with a focus on small and medium-sized enterprises, including both corporate and non-corporate entities.

SME risk profiles are examined by assessing failure probability and resolution time through various lenses, including time, firm age, turnover class, industry, firm type, and income level. Notably, resolution time, used as a proxy for recovery, exhibits a negative correlation with recovery, suggesting that longer resolution periods are often associated with higher costs or lower recovery rates. The analysis reveals a general decline in risk for Australian businesses, marked by reduced failure probabilities and shorter resolution times over the period 2002 through 2022, although some spikes were noted during the GFC and COVID-19 pandemic. It was found that younger and smaller firms face a higher likelihood of failure. Furthermore, as firms grow larger and older, the time required to resolve insolvency tends to increase. Among industries, discretionary sectors like Hospitality experience the highest rates of failure, while non-discretionary sectors, such as Healthcare, are less prone to failure. In terms of resolution time, capital-intensive industries like Mining have the longest average durations, whereas human-intensive industries like Hospitality and Arts tend to resolve more quickly. Comparatively, corporates generally exhibit lower failure risks and longer resolution times than sole traders, though this trend does not uniformly apply across different firm sizes. Lastly, firms with lower income levels tend to encounter higher failure risks and slightly shorter resolution times than their higher-income counterparts.

The fluctuations in insolvency risk and resolution time can be explained by various financial and macroeconomic variables. The income ratio lowers the probability of failure and resolution time. In contrast, the interest ratio intensifies both risk measures. Both turnover growth and liquidity have negative and significant impacts in the failure model but are insignificant in the resolution time model. Higher leverage increases the failure risk but reduces the resolution time. Meanwhile, higher CAPEX exerts the opposite effects. Regarding the macroeconomic variables, lower cash rates and changes in the

[1] Finance Discipline Group, UTS Business School, University of Technology Sydney, Ultimo, NSW 2007, Australia.
This paper is dedicated to the memory of Professor Harald (Harry) Scheule, who served as the project lead and whose invaluable guidance and expertise greatly shaped the direction and impact of this work. The authors are grateful for the financial support of the Australian Securitisation Forum (ASF) and the feedback provided by the ASF Industry Advisory Group. This research acknowledges the data provided by the Australian Taxation Office (ATO), Australian Bureau of Statistics (ABS) as well as the Australian Financial Security Authority (AFSA). We thank Ignatius McBride, Mark Wicht and Lan Dang at AFSA for providing analytics.

unemployment rate have positive effects on both risk metrics, while the change in GDP growth lowers failure risk but lengthens the resolution time. Different failure indicators and several machine learning methods have been used in the robustness check and the correlation between the best-performing model and the traditional model is around 60%.

Key takeaways/highlights

- Larger firms tend to have a lower failure risk; however, they often experience longer resolution times.
- Younger firms are more susceptible to failure, while mature firms generally face longer resolution processes when insolvency occurs.
- Discretionary industries, such as Hospitality, exhibit the highest risk of failure, whereas non-discretionary industries, like Healthcare, demonstrate lower risk levels.
- Capital-intensive industries, such as Mining, require longer resolution times, while human-intensive industries, such as Hospitality and Arts, tend to resolve more quickly.
- Sole traders face a higher risk of failure compared to corporations, although this risk difference has narrowed since the Global Financial Crisis (GFC). Additionally, sole traders experience shorter resolution times.
- Firms with higher income levels are associated with lower failure risks and shorter resolution times when insolvencies do occur.

Working group

The paper is based on a collaboration with the workgroup:

- Chris Dalton and Robert Gallimore (Australian Securitisation Forum)
- James Donovan (Challenger)
- Michael Heath (Judo Bank)
- Dennis Koh (Credabl)
- Michael Landgraf (Illion)
- Stephen Maher (Australian Office of Financial Management)
- Josh Manning (Manning Asset Management)

JEL: G01; G20; G21; C51, C55

Keywords: Australia, Credit Risk, Insolvency, Probability of failure, Resolution time, Small and Medium-sized Business (SME)

1. Motivation

Small-to-medium-sized enterprises (SMEs) are reasonably considered the backbone of most economies. The World Bank (2019) estimates that SMEs represent about 90% of businesses and more than 50% of employment worldwide. This study focuses on Australia, leveraging twenty years of firm-level data, encompassing over 57 million observations for almost 9 million businesses, of which over 99.7% are SMEs.[2] The significant contribution of SMEs to economic growth has been recognized (Beck et al., 2005; Cull et al., 2006). As SMEs provide for a significant proportion of economic activity, SMES should be well supported by investors. This paper aims to provide methods and empirical analysis to examine risk profiles and identify risk drivers in SME lending in Australia. Note that companies and sole traders are both referred to as businesses. The key identifiers are the Australian Business Number (ABN) and Australian Company Number (ACN). Corporates have both ABN and ACN numbers whereas individuals (i.e., sole traders) have only ABN numbers.

After loan origination, the risk is realised through insolvency events and loss rates given failure. A failure event is indicated when firms are involved in different stages of the liquidation process. Resolution times—the period between the first date a failure event is reported and the deregistration date—are used as a proxy for loss rates given failure.[3]

Various data sources from the Australian Taxation Office (ATO) and the Australian Securities and Investments Commission (ASIC), accessed through the Australian Bureau of Statistics (ABS), have been utilised to construct a comprehensive modelling population for this research. By combining traditional econometric methods with advanced machine learning techniques, we evaluate models for failure prediction and resolution time.

The contributions of this paper are as follows. First, the paper analyses tax and security filings of all businesses, and insolvency filings to portray the risk profiles of Australian businesses across time, firm ages, turnover classes, industries, firm types, and income levels. Second, factor models for the probability of failure and time to resolution are constructed using traditional econometric methodology, followed by an analysis of model predictions. Third, the robustness tests using various machine learning models are implemented to examine the model performance.

The paper proceeds as follows. Section 2 introduces the research framework and specifies all models. Section 3 presents the data construction process and descriptive statistics. Section 4 provides the empirical analysis including a detailed data description, factor models, a real fit analysis of model predictions, and sensitivity tests. Section 5 reports findings from robustness tests. Section 6 summarises

[2] This number is based on the threshold of $75 million applied by the Australian Prudential Regulation Authority at present. Other countries apply similar thresholds in their local currencies, e.g., European regulators use EUR 50 million. Due to the low number of observations, we define larger firms, i.e., non-SMEs as firms with turnover above $10 million. Note that small SMEs are mostly organised as sole traders, i.e., businesses that are organised as individual business owners and not as corporates.

[3] Banks typically model loss rates given default (LGD), which are not available for the domain of Australian companies. However, prior studies document a positive correlation between resolution time and LGD. For example, Gürtler and Hibbeln (2013) report a correlation of 26% for US corporate loans and Rösch and Scheule (2020) 39% for mortgage loans.

the findings and derives important implications for data collectors, policymakers, and the lending industry.

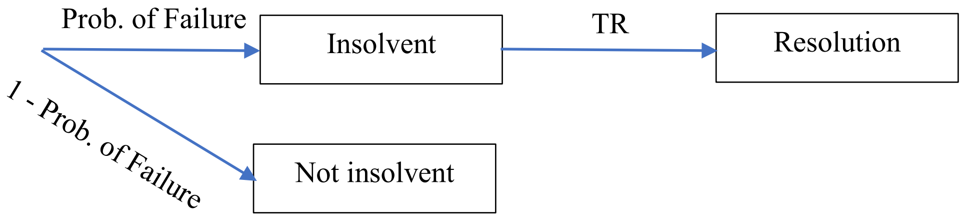
2. Research Framework

2.1. Credit risk assessment

To assess the risk, it is important to understand the process after loan origination. After a loan is originated, it may be insolvent before maturity, or mature. Figure 1 shows the process after loan origination:

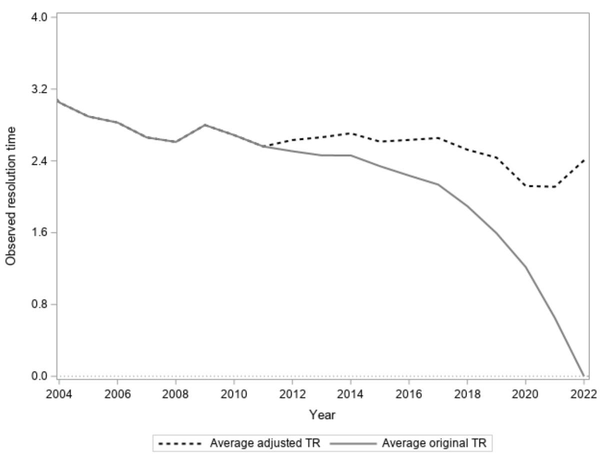
Figure 1: Processes after loan origination

Note: This figure shows the decision process of businesses after loan origination. Businesses may or may not be insolvent. Loans to businesses that fail undergo a liquidation process that may last multiple years. The end of this process is called a resolution.



The respective event probabilities are the probability of failure (PoF). The observed event rates for selected bins are abbreviated as OFR (observed failure rate). The insolvency event is generally followed by a firm deregistration. The difference between the firm deregistration time and insolvency time is the time to resolution (TR). TR is likely downward biased as the recent failure events are likely uncompleted, hence missing in TR values. Hence, TR is biased-corrected with the rolling average value of the previous TR (see, Figure 2). Businesses that are insolvent and then cure, or do not deregister, are treated as missing values. Failure events are recorded if the insolvency event occurs within one year from the measurement of the feature, i.e., features are time-lagged regarding insolvency events and resolution times. This enables models to predict credit risk in the future.

Figure 2: Corrected resolution time over time



2.2. Failure/insolvency definition

This research relies solely on corporate insolvency data, as defined by ASIC. A firm is considered insolvent if it cannot pay all debts when due.[4] Insolvent firms are those unable to pay all debts when they fall due for payment. There are different levels of activity of corporate insolvency in Australia, which can be outlined as follows:

- **Level 1**: First-time entry into external administration or appointment of a controller.
- **Level 2**: Subsequent appointments of controllers or managing controllers.
- **Level 3**: Reports from external administrators or receivers.
- **Level 4**: Appointment of liquidators.
- **Level 5**: Voluntary liquidation initiated by shareholders and/or directors.
- **Level 6**: Court-ordered liquidation following a creditor's application.

A firm is classified as insolvent if it enters any of these levels, with the first recorded event date used as the failure date. During insolvency proceedings, firms may enter a Deed of Company Arrangement—a legally binding agreement between the company and its creditors regarding asset management. Firms may also undergo restructuring if they offer creditors a better return than liquidation. ASIC records these activities, and firms without further liquidation or winding-up records are excluded from the failure sample.

ASIC records approximately 170,000 insolvency events and their corresponding dates over the research period. To simplify, insolvency events are categorised into four types: (1) voluntary liquidation, (2) processing liquidation, (3) receivership, and (4) court liquidation.[5] The higher category number indicates the level of complexity of the insolvency procedure.

The failure flag is indicated if the insolvency date occurs within one year after the most recent financial data. This approach is consistent with Allison (2010) and Kenney et al. (2016). This approach of defining the failure assumes that the insolvent business should have the financial characteristics from the most recent financial reports. However, the number of failure events indicated in the sample is less than 50% of the total number of failure observations available in ASIC data. This issue is possible because the smaller firms do not report to ATO in the lead-up years, and it takes years for ATO to realise and chase these firms. In the robustness test, we allow the failure date to occur within two or three years from the most recent financial report period to increase the matching rate.

Failure indicator	Number of failure events merged to the population
Within one year	72,869

[4] The glossary of insolvency terms can be found here: https://asic.gov.au/regulatory-resources/insolvency/insolvency-information-for-directors-employees-creditors-and-shareholders/insolvency-a-glossary-of-terms/

[5] Creditors voluntary winding up and simplified liquidation are categorised as voluntary liquidation. The appointments of controller, managing controller, or liquidator are categorised as Processing liquidation. The appointments of the receiver or receiver/manager are categorised as Receivership. Court winding up and court liquidation events are categorised as Court liquidation.

Within two years	117,149
Within three years	138,611

Appendix A presents a detailed description of the number of failing firms and the observed insolvency rate using the first definition across the year, firm age, industry, turnover categories, industry x turnover, and firm age x turnover.[6] Using the strictest one-year definition, the overall observed failure rate is 0.49%. This increases to 0.79% when a two-year window is applied and further to 0.93% under the three-year definition.

- By *Year*: Insolvency rates peaked at 0.7% during 2010–11 and 2011–12, reflecting the lingering effects of the Global Financial Crisis.
- By *Firm Age*: The insolvency rate increased steadily over the first seven years of a firm's life and declined thereafter.
- By *Industry*: The Hospitality, Administrative and Support Services, and Construction sectors recorded the highest insolvency rates across the two-decade period.
- By *Turnover Category*: Firms in the smallest turnover bracket consistently faced the highest risk of failure compared to their larger counterparts.
- By *Industry × Turnover*: Small construction firms alone accounted for up to 14% of all insolvency events, followed by small firms in Professional, Scientific, and Technical Services with 6%.
- By *Firm Age × Turnover*: Small firms aged between 3 and 5 years contributed approximately 18% to total insolvency cases in the sample.

2.3. Features

Due to the data limitation, the risk levels will be estimated by two models: Core and extended models. The core model relies on fundamental performance metrics readily available across firms and periods, making it widely applicable for baseline risk assessments. Core features capture essential profitability, cost burden, and growth dynamics, which serve as primary indicators of a firm's financial health and are highly sensitive to changes in operational performance. The extended model builds on this by adding detailed balance sheet metrics like liquidity, leverage, and CAPEX, offering a comprehensive view of financial structure and resource allocation, which aids in assessing long-term stability and growth potential.

In line with current industry practice and prior literature, the following time-varying firm features for the core and extended models are considered:

- Net income over turnover (Income ratio)
- Interest expenses over turnover (Interest ratio)
- Turnover growth (Turnover growth)

[6] The results with the other failure definitions are available on request.

These features are common in credit risk modelling due to their relevance to key financial metrics. All features are turnover-based, as turnover is a critical consideration in tax filings and is consistently observed for most businesses each year. Note that turnover has been inflation-adjusted to the last observation year, 2022, by dividing by the price index of the respective year and multiplying by the price index of 2022, as our outcome variable relates to the year following the feature observation.

The income ratio can indicate profitability and liquidity, and the interest ratio represents leverage, which are essential factors in assessing failure risk. Turnover growth can signal financial health, potentially lowering credit risk, as long as growth is sustainable and not excessively debt-funded. Stable turnover growth predicts stable cash flow.

The following time-varying firm features relate to the balance sheet and Business Activity Statement (BAS) filings and are only considered for the extended models:

- Current assets minus current liabilities over total assets (Liquidity)
- Liabilities over total assets (Leverage)
- CAPEX over turnover (CAPEX)

The features are proxies for liquidity, leverage, and growth and are common in credit risk modelling. They are aligned with Kenney et al. (2016). All continuous variables are winsorised at the 5th and 95th percentiles. Missing values are excluded from the analysis.[7] All firm-level features are one-period lagged to the current period to reflect the fact that information should be available to predict the risk levels.

Fixed effects (effect coding) for the first-level industry classification and the following annual turnover categories are included.[8] In addition, fixed effects for different insolvency types (i.e., voluntary liquidation, processing liquidation, receivership, and court liquidation) are included in the resolution time models.[9] Similarly, the years since the first observation year (Age) and squared age (Age2) are included to reflect the non-linearity of age in the models.

Macroeconomic variables have also been tested. The Australian economy did not experience major downturns despite global episodes such as the Global Financial Crisis or COVID and the macroeconomic variation is limited. Despite that, we consider the impacts of the cash rate, changes in GDP growth, unemployment rate, and low-consumption period. This is because even minor fluctuations in these variables can significantly influence firms' credit risk and operations. By incorporating these

[7] Sometimes missing values may be imputed if data is scarce. Imputation generally averages results as missing values are often tied to non-average observations (e.g., very small businesses).

[8] In effect coding, each category's coefficient represents its deviation from the overall mean of the dependent variable, rather than a specific reference category.

[9] The most common corporate insolvency procedures for insolvent firms are court liquidation, voluntary liquidation and receivership. The court liquidation is indicated by ASIC with the "court winding up" or "court liquidation" events. The voluntary liquidation is indicated by ASIC with the following events: "creditors voluntary winding up", "Creditor voluntary liquidation", "simplified liquidation", or "liquidator of a simplified liquidation". The receivership is indicated by ASIC with the appointment of receiver or receiver/manger. For the insolvency events related to the appointment of a controller, managing controller and provisional liquidation/liquidator, we coded as processing liquidation. For insolvency events related to deed of company arrangement, restructuring plan, or voluntary administration appointment, we do not code as insolvency event.

factors, we aim to ensure our models remain robust and sensitive to potential economic shifts, thereby enhancing the accuracy and reliability of our credit risk assessments.

The cash target rate[10] (cash rate) is included for several reasons: First, the cash rate affects interest expenses, which serves as a proxy for the debt ratio. Since many Australian businesses, especially sole traders, do not report balance sheet information and thus have unobserved leverage, controlling for the cash rate helps address this issue. Second, the lending system is dynamic and tied to the cash rate through exchange settlement accounts. Lending standards have been identified as important factors for the US (see Berger and Udell, 2004; Lown and Morgan, 2006; and Mian and Sufi, 2009) and are also relevant here. Third, a higher cash rate makes debt finance more costly and, hence, has a positive impact on failure. The lagged cash rate is used due to the sticky nature of the interest rate.

The change in GDP growth reflects current economic conditions. Higher GDP growth, resulting in increased consumption and investment, allows businesses to expand and improves their loan serviceability.

The change in the unemployment rate affects both consumer demand and borrower risk profiles. A higher unemployment rate may signal economic distress, reducing disposable income for households and raising failure risks for businesses.

Finally, the Australian economy has experienced a low-consumption period following the GFC and during the COVID-19 period.[11] In response, the Australian government and the RBA implemented emergency measures such as stimulus packages and interest rate cuts to stabilise the economy and mitigate the impacts of these global crises. As recovery takes hold, measures to combat inflation and strengthen the Australian currency often become more pronounced, potentially leading to a low-consumption period. During such periods, failure rates may increase as businesses face higher costs, subdued consumer demand, and restricted financing. To account for these dynamics, we include indicators for the years 2011, 2012, and 2019 to capture the effects of these low-consumption periods in Australia.

3. Data

3.1. Data sources and processing

The research uses firm-level data from the Business Longitudinal Analysis Data Environment (BLADE) provided by ABS. This database contains data on all active businesses from 2001-02 to 2021-22, sourced from the ABS Business Register, Australian Taxation Office (ATO), Australian Securities & Investments Commission (ASIC), etc. The main sources of data from BLADE employed for this project include Business Activity Statements (BAS), Business Income Tax (BIT), and ASIC insolvency data.

[10] The cash target rate is the key monetary policy rate in Australia and applies to exchange settlement accounts held by banks at the Reserve Bank of Australia.

[11] RBA reported the drop in household consumption growth since mid-2019 (see, https://www.rba.gov.au/statistics/tables/csv/h2-data.csv)

BAS and BIT data contain data for four different types of firms that are registered for goods and service tax[12], including corporate, individual, trust, and partnership. However, trusts and partnerships have different tax treatments and hence were removed from the sample.[13] Corporates have both ABN and ACN numbers whereas individuals (i.e., sole traders) have only ABN numbers. The linking tables between ABN and ACN are provided for three financial years: 2019-20, 2020-21, and 2021-22. As suggested by the ABS, the 2019-20 table can be used for previous years' records.

While the main risk factors are constructed from information collected from BAS and BIT, the risk indicators including the failure flag and resolution time are captured in a comprehensive record of insolvency filings provided by ASIC. Note that the ASIC data are only available for corporates, meaning no failure data is available for sole traders, hence failure risk for sole traders can only be obtained through model calibration. The insolvency filings consist of different events and corresponding dates.

Other information such as industry, ANZSIC, and birth date (i.e., the first financial year an ABN reports information through either ABS or ATO) is also collected from the BLADE database. Finally, we manually collect the annual cash rate and GDP growth rate from the Reserve Bank of Australia's website to use as proxies for macroeconomic conditions.

Table 1 summarises the collected information from different sources of data.

Table 1: Source of model features

Data source	**Information**
BAS	Capital expenses, Turnover
BIT	Current assets, current liabilities, firm type, interest expenses, net income, total assets, total liabilities
ASIC	Failure date, failure event type
Other BLADE sources	Birth date, industry, ANZSIC
RBA	GDP growth, cash rate, unemployment rate

Note that the risk models in this project are based on accounting information. Beaver et al. (2012) explore the effect of cross-sectional and time-series differences. They find that the predictive power of the accounting-based model reduces the presence of discretion in financial reporting. Taxpayers have an incentive to submit lower incomes and higher expenses to save taxes. Loan applicants have the opposite incentive: to submit higher incomes and lower expenses to secure a loan. The bias for lender data is likely to be greater than for tax data as the Australian Tax Office actively searches for tax fraud and the penalties for business transactions are likely lower than for tax evasion.

We process data using the following steps. Over the period from 2001-02 to 2021-22, the BAS data has

[12] GST registered companies with turnover in excess of $75,000 per annum provide a Business Activity Statement (BAS) and we additionally collect some information (e.g., capital expenditures-CAPEX) from these submissions.

[13] A partnership has its own ABN but doesn't pay income tax on its profit, but each partner reports their income in their own individual tax returns. A trust must lodge a trust income tax return.

55.47 million observations for 7.69 million businesses, while the BIT data after removing trusts and partnerships has 44.13 million observations for 7.15 million businesses, some entering and some exiting the sample over time.[14] Taking BAS as the base data, BIT and other BLADE data are merged into BAS on each ABN and year. It is noticed that if the firm type is not indicated, all BIT data fields contain missing values. We remove firms where all rows are missing on BIT information.

The data processing includes merging ASIC records using the ACN as the unique identifier. We define the insolvency date as the earliest recorded insolvency event and flag a business as failed if this date falls within one year of its most recent financial report. This definition aligns with the methodologies of Allison (2010) and Kenney et al. (2016), under the assumption that the most recent financials reflect the firm's status leading up to failure. However, fewer than 50% of insolvency cases in the ASIC data are captured under this definition. A likely reason is that smaller firms often stop reporting to the ATO prior to failure, and delays in detection or enforcement may extend over several years. To address this, our robustness checks expand the failure window to two and three years after the latest financial report, improving the matching rate between financial and insolvency data.

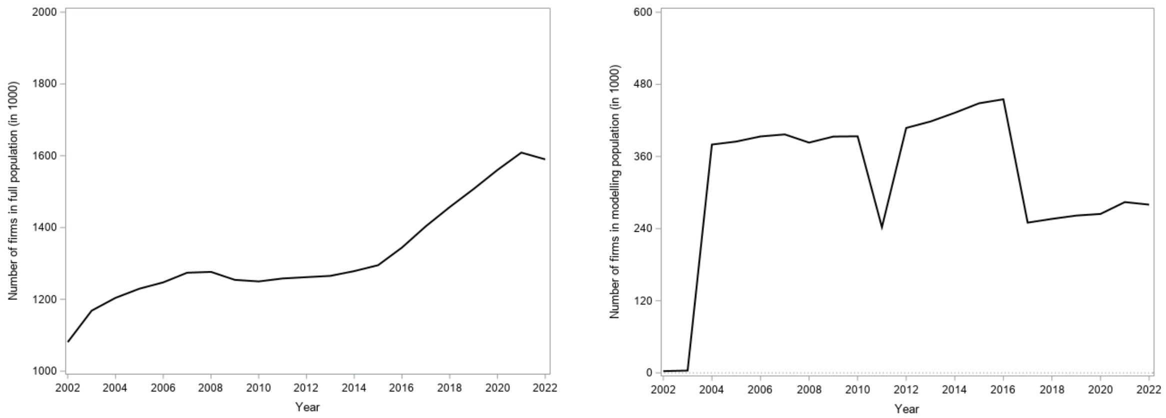
For each identified insolvent business, the resolution time is calculated as the end of the last financial year they submitted the BAS report to ATO minus the first recorded failure event from ASIC data. We took the first failure observation and dropped the remaining for the insolvent businesses. In other words, cures are not considered, which results in a missing resolution time. As a result, the final sample spanning over 21 years contains 27.9 million observations.

Figure 3 illustrates the total number of firms (measured by the number of non-missing ABNs) in the full population over time. The number of firms increased from approximately 1.1 million in 2002 to around 1.6 million in 2022, with an average of 1.3 million firms across the period. In comparison, the modelling population contains a significantly smaller average of about 300 thousand firms. This reduction is primarily due to the application of data lags and the exclusion of observations with missing explanatory variables. The most significant reductions occurred in 2011 and during the COVID-19 period, likely due to an increase in firm exits driven by declines in consumer demand. To account for this, a dummy variable has been introduced in the regression model, taking a value of 1 for these years and 0 for all other years, to control for the impact of these events.

Figure 3: Number of firms in the full and modelling population

Note: This figure shows the number of firms measured by the number of non-missing ABNs in the full (left) and extended modelling (right) populations.

[14] For further details, see https://www.abs.gov.au/statistics/economy/business-indicators/counts-australian-businesses-including-entries-and-exits/latest-release



Resolution time is the time from the failure time to the deregistration time. The deregistration is not directly observed, and it is approximated as the EOFY of the last period a insolvent firm makes its BAS submission. The features are winsorised by the 5th and 95th percentile to mitigate the impact of outliers showing the correlations for all outcome variables and features.

3.2. Descriptive statistics

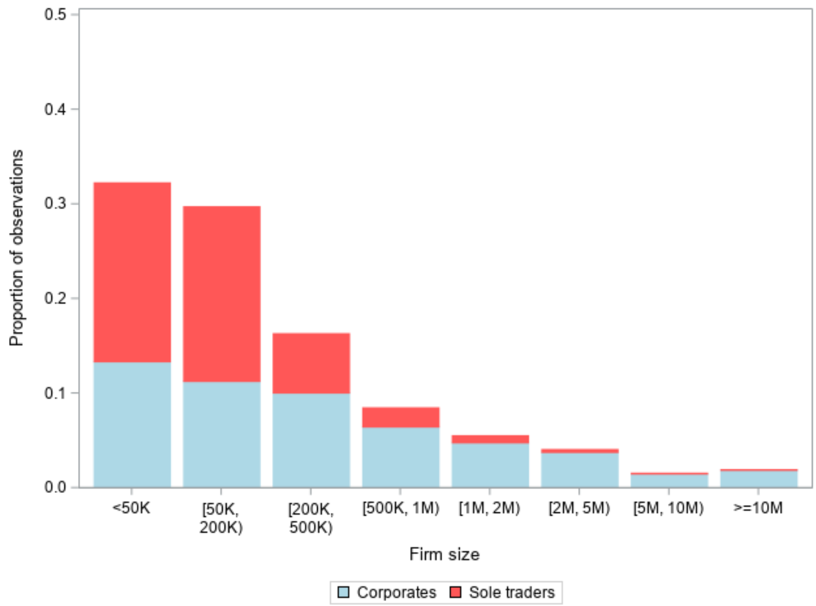
In our analysis, sole traders and turnover classes may indicate SMEs as smaller businesses with smaller turnover and are dominantly organised as individuals rather than corporates for cost and complexity reasons. Figure 4 shows the number of observations for each firm type and firm size.

The turnover category is slightly adjusted from the ABS's turnover size where the group above 200K and less than 5 million is further split. The size classification is used for micro-businesses (up to 50K), very small businesses (from 50K to less than 200K), small businesses (from 200K to less than 500K and from 500K to less than 1 million), medium businesses (from 1 million to less than 2 million and from 2 million to less than 5 million), larger medium businesses (from 5 million to less than 10 million) and larger businesses (from 10 million and more). It is shown that smaller firms have a proportionally larger degree of sole traders than corporates and larger firms do the opposite. For example, the fraction of sole traders is very negligible for businesses with above $10 million in turnover, while around two-thirds of businesses with less than $50K turnover are sole traders.[15]

Figure 4: Frequency distribution by turnover class and sole trader/corporate

Note: This figure shows distribution by turnover category and sole trader/corporate. The following classification is used micro-businesses (up to $50,000), very small businesses (from $50,000 to less than $200,000), small businesses (from $200,000 to less than $500,000, and from $500,000 to less than $1 million), medium businesses (from $1 million to less than $2 million, and from $2 million to less than $5 million), larger medium businesses (from $5 million to less than $10 million) and larger businesses (from $10 million and more).

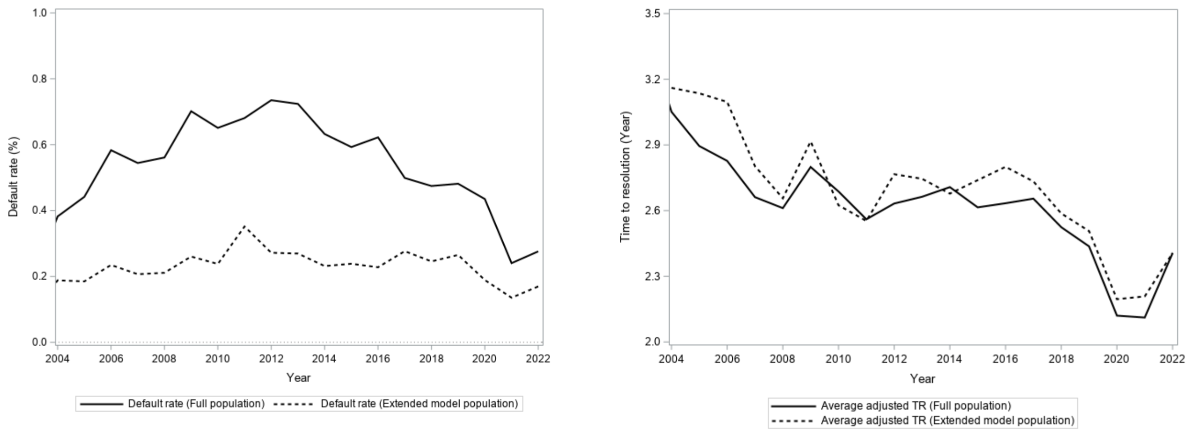
[15] Incorporation includes registration and maintenance fees as well as ongoing accounting and auditing fees.



Despite the number of failure events merged into the sample, the model population is reduced due to the missing values on explanatory variables. This could impact the observed failure rate and resolution time. Figure 5 shows the failure rate and average resolution time over time obtained from the full population (solid lines) and population used for extended models (dash lines).

Figure 5: Average observed outcomes over different populations

Note: This figure shows the failure rates (left), and time to resolution (right) for corporates only. The average failure rate is observed between 2002 and 2022. Similarly, the average time to resolution is observed between 2004 and 2022.[16]



The failure rate from the full population is greatly higher than that from the extended model population, with the former failure rates double the latter. In addition, the failure rate from the full population depicts a hump shape over time, while this pattern is not clearly observed with the extended model population. Despite that, the peak failure rate seems to align between the two populations. The averages obtained from the two populations likely align over time regarding the resolution time. This could be due to the gentle fluctuations in resolution time across firms. In sum, the missing values on explanatory variables

[16] The number of insolvency events in the extended modelling population is fewer than 10 in 2002 and 2003; therefore, the resolution time is being excluded.

lead to the exclusion of a significant proportion of outcomes in the model population, which may affect the model calibration.

Table 2 summarises the outcome events as well as the observations used for model fitting (F) and prediction (P) of expected outcomes for the various outcome variables.

Post-origination process outcomes (insolvency and resolution times) are limited to corporate insolvency filings. In other words, insolvency events and resolution times for sole traders are not observed. The fitted values are computed for all categories, returning consistent results for the categories.

The failure model is based on ATO data and has approximately 7 million observations, and the resolution time model is based on the ATO data, and the observation of insolvency events and has 16,940 observations. The extended models have lower observation numbers as some observations were dropped due to missing features. The number of events in TR models is lower than that of failure models, which is likely to be attributed to the missing insolvency type (i.e., AID type) as this factor is included in the TR models but not in the failure models.

The observed failure rate and average resolution time are demonstrated for different periods, firm age, firm size, industry, and income categories using the extended model population.

Table 2: Sample sizes for various models

Note: This table shows the outcome events, and the observations used for model fitting (F) and prediction (P) of expected outcomes for failure, and time to resolution of core and extended models.

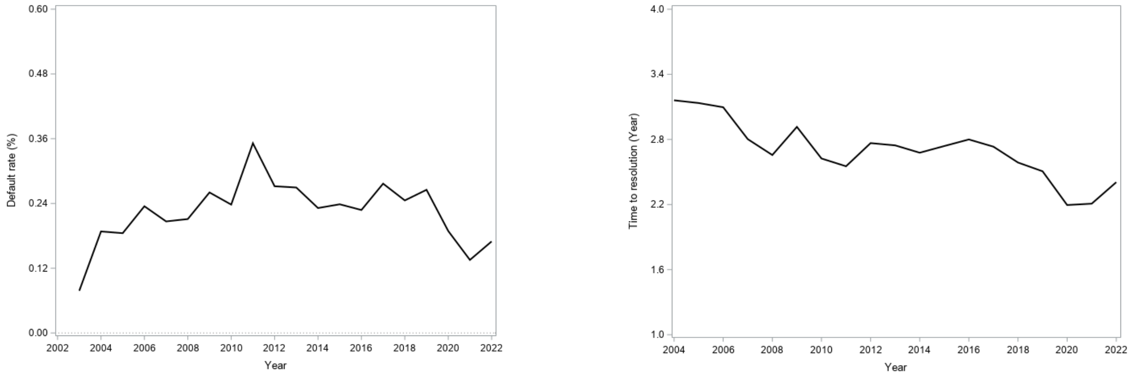
Outcome	**Failure**	**TR**
Method	Logit	OLS
Core		
Events	17,100	16,940
Nobs. F	7,220,362	16,940
Nobs. P	13,712,730	13,712,730
Extension		
Events	15,451	15,231
Nobs. F	6,705,379	15,231
Nobs. P	6,705,379	6,705,379

Figure 6 shows the observed failure rates (left) as well as average resolution time (right) for corporates only over time[17].

Figure 6: Average observed outcomes over time

[17] Note, the figure for resolution time only spans from 2004 to 2022 due to the ABS clearance rule of 10.

Note: This figure shows the failure rates (left), and time to resolution (right) for corporates only. The average failure rate is observed between 2004 and 2022. Similarly, the average time to resolution is observed between 2004 and 2022.



The failure rate peaked in 2011, likely due to the lingering effects of the Global Financial Crisis (GFC). Excluding this peak, the failure rates before and after the GFC were relatively stable, with a slight increase observed in 2019. However, failure rates declined sharply following 2019, potentially driven by the economic stimulus measures such as lower cash rates and stimulus payments implemented during the COVID-19 pandemic. As for resolution times, these decreased consistently over the period, which can be attributed to improvements in the legal framework for handling insolvencies more efficiently. An example of this improvement is the implementation of a simplified liquidity process from 1 January 2021, which allows a streamlined creditors' voluntary winding up for companies having liabilities less than $1 million. The simplified liquidation process is expected to take a few months rather than the potentially longer timelines of traditional liquidations, which can take up to a year or more.

Figure 7: Average observed outcomes by firm age

Note: This figure shows the failure rates (left), and time to resolution (right) for corporates of different ages. The average failure rate is observed between 2002 and 2022. Similarly, the average time to resolution is observed between 2002 and 2022.

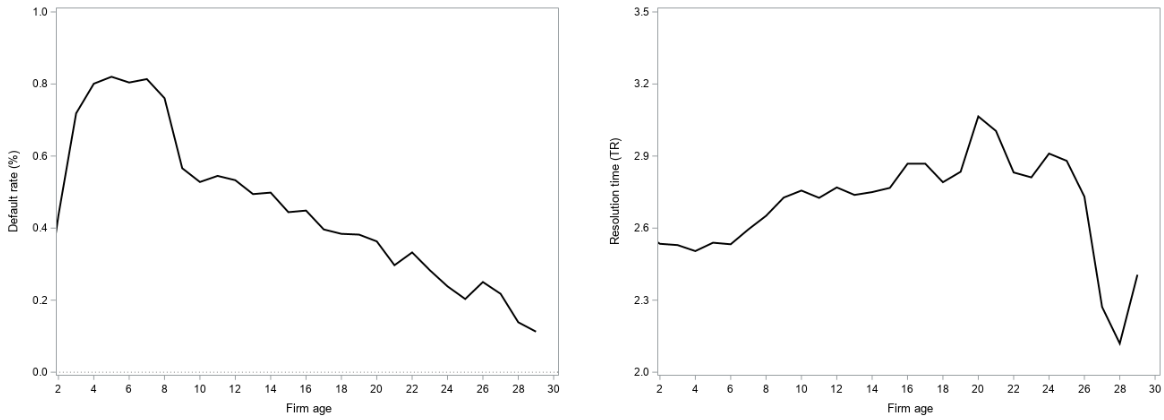


Figure 7 shows that older businesses have significantly lower failure rates than younger ones, likely due to their established market presence and operational stability. However, when older firms do fail, the resolution process takes longer. This is because these firms tend to have more complex financial structures, larger asset portfolios, and long-standing relationships with multiple stakeholders, which can complicate the resolution. Notably, the figure depicts a sharp decrease in resolution time for businesses over 25 years old, which may be an outlier driven by a lack of observations for this age group.[18] In contrast, younger firms, while more prone to insolvency due to limited financial history, experience shorter resolution times because their simpler structures and fewer assets make liquidation or restructuring quicker and less complex.

Figure 8 shows the observed failure rate (left) and average resolution time (right) for corporates across different industries.[19] The industry with the highest failure rate is Hospitality, followed by Support, Construction, and Manufacturing. The elevated failure rates in Hospitality and Support can be attributed to the economic slowdown, particularly during the COVID-19 pandemic. Financial difficulties in Construction and Manufacturing are likely due to the ongoing effects of the pandemic, compounded by a higher interest rate environment, lockdowns at construction sites, and rising material costs.[20] Meanwhile, industries such as Healthcare and Finance experience lower failure rates because they provide essential services, which consumers continue to rely on even during economic slowdowns, such as those caused by the COVID-19 pandemic. Regarding resolution time, the Mining industry has the longest duration for resolving insolvencies, as its heavy, specialised machinery is more challenging to liquidate. In contrast, the Hospitality industry has the shortest resolution time, as buildings like hotels and restaurants, though asset-heavy, can be easily repurposed or renovated for other uses.

Figure 8: Average observed outcomes by industry

Note: This figure shows the failure rates (left), and time to resolution (right) for corporates across different industries. The average failure rate is observed between 2002 and 2022. Similarly, the average time to resolution is observed between 2002 and 2022.

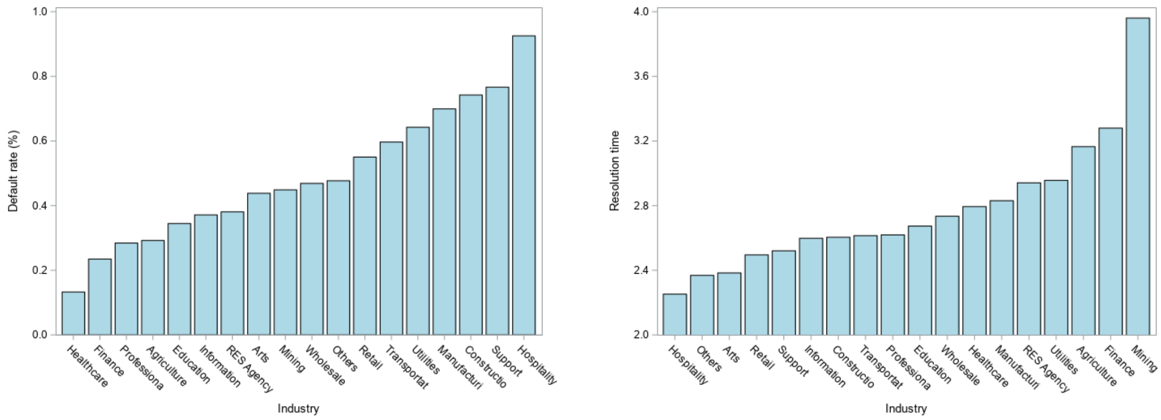


Figure 9 shows the observed failure rate (left) and average resolution time (right) for corporates across different size classes. The two bar charts reveal contrasting trends for failure rates and resolution times across different business sizes. Larger businesses face lower failure risk but have longer resolution times. The lower failure rates are expected, as larger businesses tend to be well-established and operate with greater efficiency. However, their longer resolution times can be attributed to their complex financial structures, which often involve multiple assets and liabilities. Additionally, these businesses

[18] We run the factor models on the sample excluding firms older than 25 years old. The results are consistent with the main analysis. Results are available on request.

[19] We remove Government sector and unknown sector from the sample.

[20] https://www.rsm.global/australia/insights/restructuring-insights/australian-construction-industry-rebuilding-or-are-more-collapses-way

are subject to heightened regulatory and legal scrutiny, which prolongs the resolution or liquidation process.

In contrast, smaller businesses tend to experience higher failure rates but have shorter resolution times. Their higher risk of failure is due to limited resources, lower operational stability, and greater vulnerability to economic fluctuations. However, the resolution process for smaller businesses is often faster, as their financial structures are simpler, with fewer assets and liabilities to manage. This allows for quicker liquidation or restructuring, with fewer legal and regulatory hurdles compared to their larger counterparts.

Figure 9: Average observed outcomes by turnover classes

Note: This figure shows the failure rates (left), and time to resolution (right) for different-sized corporates. The average failure rate is observed between 2002 and 2022. Similarly, the average time to resolution is observed between 2002 and 2022.

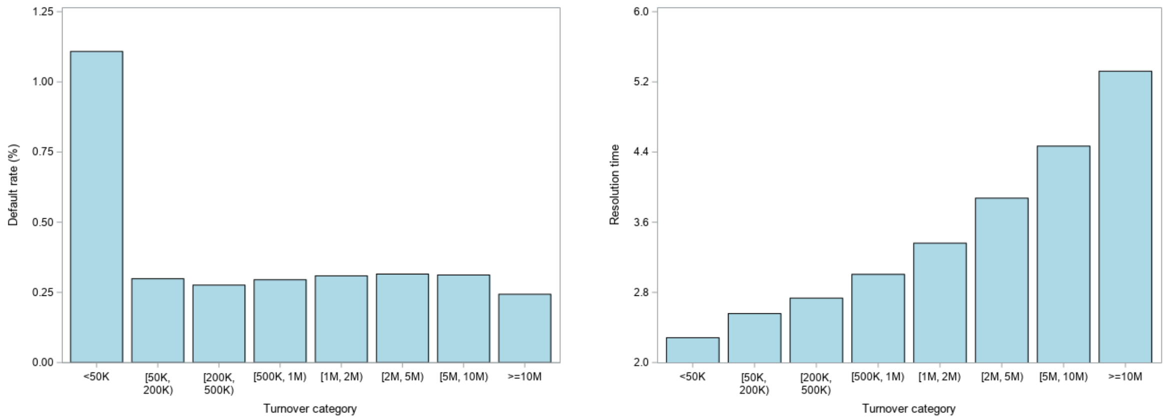


Figure 10: Average observed outcomes by income level

Note: This figure shows the failure rates (left), and time to resolution (right) for corporates of different income levels. The average failure rate is observed between 2002 and 2022. Similarly, the average time to resolution is observed between 2002 and 2022.

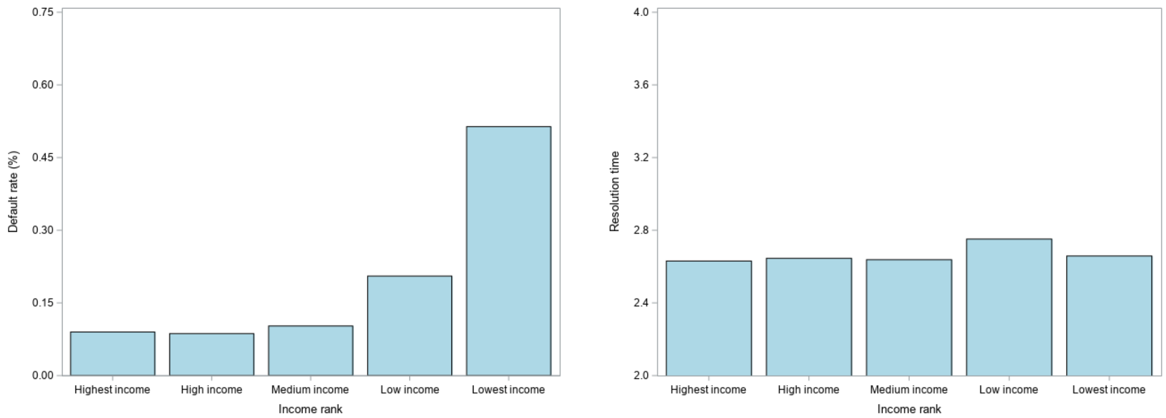


Figure 10 demonstrates a clear inverse relationship between firm income and failure rates, indicating that firms with lower income levels are more prone to insolvency. Lower-income firms are often more vulnerable to economic shocks, as they typically have fewer financial resources and buffers to deal with periods of reduced revenue or increased costs. Furthermore, such firms may face higher borrowing costs or restricted access to credit, exacerbating their financial constraints. Regarding the resolution time, the differences are negligible across different income categories.

Table 3: Descriptive statistics for outcomes and features variables

Note: This table shows the summary statistics for outcome variables (failure and time to resolution), and features (Income, Interest, Turnover growth, Liquidity, Leverage, CAPEX ratio, age, age-squared, and cash rate). Income is defined as Income/Turnover. Interest is calculated as Interest/Turnover. Turnover growth is measured as (Current turnover – previous turnover)/Previous turnover. CAPEX is computed as CAPEX/Turnover. Liquidity and Leverage are defined as (Current assets – Current Liabilities)/Total assets and Total liabilities/Total assets, respectively. The variables are observed between 2002 and 2022. The cash rate is reported as a time series due to reporting requirements from the data provider.

Variable	Nobs	Mean	Std. Dev	Nobs	Mean	Std. Dev	Nobs	Mean	Std. Dev
		All			Corporates			Sole traders	
Output events									
Failure rate (%)	7,221,032	0.236	0.049	7,221,032	0.236	0.049			
Resolution time	16,942	2.715	2.043	16,942	2.715	2.043			
Risk drivers									
Turnover growth	13,712,730	0.212	0.806	7,221,032	0.212	0.806	6,491,698	0.147	0.705
Income ratio	13,712,730	0.032	0.233	7,221,032	0.032	0.233	6,491,698	0.038	0.231
Interest ratio	13,712,730	0.015	0.028	7,221,032	0.015	0.028	6,491,698	0.009	0.026
Liquidity ratio	6,732,232	-0.023	0.678	6,705,965	-0.024	0.678	26,267	0.380	0.547
Leverage	6,783,556	1.020	1.110	6,757,223	1.023	1.110	26,333	0.427	0.709
CAPEX	13,712,730	0.029	0.074	7,221,032	0.029	0.074	6,491,698	0.022	0.067
Age	13,712,730	10.215	6.244	7,221,032	10.224	6.248	6,491,698	7.935	4.484
Age^2	13,712,730	143.329	160.335	7,221,032	143.565	160.513	6,491,698	83.079	86.830
Cash rate	13,712,730	0.036	0.020	7,221,032	0.036	0.020	6,491,698	0.040	0.018
GDP growth pct change	13,712,730	0.041	0.020	7,221,032	0.041	0.020	6,491,698	-0.003	0.015
Unemployment rate change	13,712,730	-0.152	0.541	7,221,032	-0.152	0.541	6,491,698	-0.122	0.426
Low-consumption period	13,712,730	0.135	0.342	7,221,032	0.136	0.342	6,491,698	0.107	0.309

Table 3 shows the descriptive statistics for the core model population and two subsamples for corporates and sole traders for all outcome variables and features. Regarding the outcome variables, *failure* is a binary indicator with an observed mean rate of 0.236%, while *resolution time* has an average duration of 2.715 years. The reported failure rate reflects the aggregate ratio of failed companies to the total number of registered firms. Although the observed failure rate in our sample is somewhat lower than the overall failure rate (OFR) of 0.49%—primarily due to missing data—this estimate remains robust for two key reasons. First, our sample estimate aligns closely with the long-run average failure rate reported by Kenney et al. (2016) in their RBA study, which finds a historical average of around 0.2% using data from 1967 to 2016. Second, our estimate is also broadly consistent with recent statistics from the Australian Securities and Investments Commission (ASIC), which report a national failure rate of 0.33% for the 2023–2024 period. Taken together, the alignment of our findings with both historical and current benchmarks supports the reliability and external validity of our failure rate estimate.[21]

Regarding the risk factors, the income ratio is on average 3.2% with a standard deviation of 23.3%. Turnover has an average growth of 21.2%. The average liquidity ratio is -2.3%. The negative liquidity ratio can be explained by how the BIT filings are recorded. The average interest ratio is 1.5%, the average CAPEX ratio is 2.9%, the average firm age is 10 years, the average cash rate is 3.6%, the average of changes in the unemployment rate is -15.2%, and the average GDP growth is 4.1%.

Table 4 shows the correlation between outcome variables (failure and time to resolution), and features. The correlation coefficients are obtained from the sample used for the extended model, but those obtained from the observations used in the core model remain consistent. The features of the core model (income, interest, and turnover growth) are highly correlated to the extended model counterparts as they are chosen to cover liquidity and leverage and sources of funding supply and need from positive and negative growth. Liquidity and leverage are the key drivers in the double trigger theory of default and insolvency laws where an insolvency is filed if a business is over-indebted (i.e., debt exceeds assets) or unable to service liabilities. The correlations between failure and all examined features are statistically significant. For resolution time, all correlations, except for the income ratio, are also statistically significant.

[21] https://asic.gov.au/about-asic/news-centre/news-items/annual-asic-insolvency-data-reveals-increase-in-companies-failing/

Table 4: Correlation table for outcome and features variables

Note: This table shows the correlation between outcome variables (failure and time to resolution), and features (Income, Interest, Turnover growth, Liquidity, Leverage, CAPEX ratio, age, age-squared, cash rate, GDP growth change, Unemployment rate change, and low-consumption period). The definitions of these features are explained in Section 2.2. The variables are observed between 2002 and 2022.

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
Failure [1]	1.000	.	-0.021	-0.018	0.011	-0.029	0.025	-0.008	-0.001	-0.009	0.013	-0.005	0.013
Resolution time [2]	.	1.000	0.042	-0.005	0.053	0.115	-0.121	0.035	0.055	0.047	0.061	0.000	-0.015
Turnover growth [3]	-0.021	0.042	1.000	0.045	-0.058	0.004	-0.001	0.040	-0.132	-0.104	0.005	-0.010	0.002
Income ratio [4]	-0.018	-0.005	0.045	1.000	-0.197	0.251	-0.340	-0.141	0.002	0.001	-0.083	-0.005	-0.020
Interest ratio [5]	0.011	0.053	-0.058	-0.197	1.000	-0.107	0.097	0.096	0.087	0.081	-0.056	0.021	-0.007
Liquidity [6]	-0.029	0.115	0.004	0.251	-0.107	1.000	-0.650	-0.057	0.087	0.080	-0.048	0.014	-0.028
Leverage [7]	0.025	-0.121	-0.001	-0.340	0.097	-0.650	1.000	0.003	-0.109	-0.094	-0.032	0.003	0.004
CAPEX [8]	-0.008	0.035	0.040	-0.141	0.096	-0.057	0.003	1.000	-0.083	-0.076	0.146	-0.015	0.038
Age [9]	-0.001	0.055	-0.132	0.002	0.087	0.087	-0.109	-0.083	1.000	0.943	-0.245	0.028	0.024
Age^2 [10]	-0.009	0.047	-0.104	0.001	0.081	0.080	-0.094	-0.076	0.943	1.000	-0.296	0.032	-0.010
Cash rate [11]	0.013	0.061	0.005	-0.083	-0.056	-0.048	-0.032	0.146	-0.245	-0.296	1.000	-0.144	0.216
GDP growth pct change [12]	-0.005	0.000	-0.010	-0.005	0.021	0.014	0.003	-0.015	0.028	0.032	-0.144	1.000	-0.572
Unemployment rate change [13]	0.013	-0.015	0.002	-0.020	-0.007	-0.028	0.004	0.038	0.024	-0.010	0.216	-0.572	1.000
Low-consumption period [14]	0.007	-0.022	0.002	0.003	0.032	-0.004	0.004	-0.014	0.058	0.041	0.031	0.006	0.108

4. Empirical analysis

4.1. Factor models

Table 5 shows the parameter estimates, number of observations, Pseudo R-square, and AUC measure for the core and the extended models for the probability of failure (logit model) and resolution time (linear model) respectively.[22] All models include industry and turnover fixed effects (FE).[23] Sole trader dummies are not included in failure and resolution time models as insolvency events are only observed for corporates. Note that resolution times are continuous, hence AUC ratios cannot be directly computed. We create a binary indicator taking the value of 1 if TR is greater than the median value and 0 otherwise and later regress on the fitted value of TR to obtain the AUC of TR models.

Table 5: Factor Models

Note: This table shows the parameter estimates for the core model and the extended model for the insolvency/failure events and resolution time. It includes parameter estimates, standard errors (in brackets), and significance levels. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level. The event models show pseudo r-squares and AUC as performance ratios and the resolution time model R-square and the AUC ratio are based on a binary indicator that is one if an observed value is above the median.

		Core model		Extended model	
Feature		**Failure**	**Resolution time**	**Failure**	**Resolution time**
Income ratio		-0.877*** (0.025)	-0.147** (0.07)	-0.574*** (0.028)	-0.324*** (0.078)
Interest ratio		4.347*** (0.209)	4.35*** (0.576)	4.553*** (0.22)	2.894*** (0.606)
Turnover growth		-0.022** (0.009)	0.003 (0.018)	-0.018* (0.009)	-0.016 (0.019)
Liquidity				-0.56*** (0.013)	0.035 (0.024)
Leverage				0.014* (0.007)	-0.057*** (0.013)
CAPEX				-1.53*** (0.118)	0.186 (0.278)
Insolvency type	Processing liquidation				0.471*** (0.102)
	Receivership				0.802*** (0.107)
	Court liquation				0.642*** (0.041)
Age		-0.019*** (0.005)	0.042*** (0.009)	-0.019*** (0.005)	0.031*** (0.01)
Age2		-0.001*** (0.000)	-0.001*** (0)	-0.001*** (0)	-0.001** (0)
Cash rate		0.026 (0.411)	7.687*** (0.787)	0.477 (0.437)	6.592*** (0.821)
GDP growth change		-0.990** (0.468)	-0.737 (0.763)	-1.168** (0.493)	-0.933 (0.81)

[22] The impact of fluctuations of HPI and unemployment to control for potential economic recessions in all four models. The signs of coefficients are as expected but the increase in AUC is negligible, indicating their insignificant predictabilities. This is because the Australian economy did not experience a financial crisis during the GFC thanks to the limited exposure to the US market. The results are available on request.

[23] The coefficients on different industries and size categories are provided in Appendix B for future model calibration. Models without fixed effects produce consistent coefficients on risk drivers in terms of sign and magnitude. The results are available on request.

Unemployment rate change	0.137*** (0.017)	-0.144*** (0.029)	0.126*** (0.018)	-0.162*** (0.03)
Low-consumption period	0.270*** (0.021)	-0.176*** (0.035)	0.287*** (0.022)	-0.166*** (0.036)
Intercept	-6.153*** (0.031)	1.788*** (0.152)	-6.250*** (0.034)	1.783*** (0.159)
Industry FE	Yes	Yes	Yes	Yes
Turnover FE	Yes	Yes	Yes	Yes
Nobs	7,220,362	16,940	6,705,379	15,231
(Pseudo) r-square	0.055	0.146	0.072	0.175
AUC	0.705	0.656	0.735	0.672

The parameter estimates can only be interpreted for the linear resolution time model in terms of its impact on resolution time. For the logistic regressions for failure, the sign of the parameters can be interpreted. The magnitude depends on the realisation of all feature models (hence the risk level) as the weighted parameters are transformed non-linearly to obtain probabilities. Logistic regressions for binary dependent variables are the standard in industrial applications and are preferred as they demonstrate fitted values between zero and one, convex relation between features (e.g., leverage) and outcomes (e.g., failure), simplicity of re-calibrations to target levels, weight of evidence to name a few (for details refer to Rösch & Scheule, 2020). All models are calibrated in-sample and the fitted event probabilities, and resolution times equal the observed averages.

The probability of failure (stand-alone model) decreases with the net income, turnover growth, liquidity CAPEX, firm age, firm age square, and GDP growth and increases for interest, leverage, the cash rate, unemployment rate, and low-consumption periods. These variables make sense in economic terms. For example, liquidity and leverage have opposite effects: liquidity decreases failure risk whilst leverage increases it and is hence in line with the so-called double trigger theory of default. Firms with higher incomes tend to be less risky due to their greater financial sources to serve the loans, while firms paying higher interest expenses are likely to suffer the financial burden which can elevate failure risk. For macroeconomic factors, a higher cash rate or higher unemployment rate heightens failure risks, while an improvement in GDP growth benefits firms through increased demand for goods and services and enhances their loan serviceability.

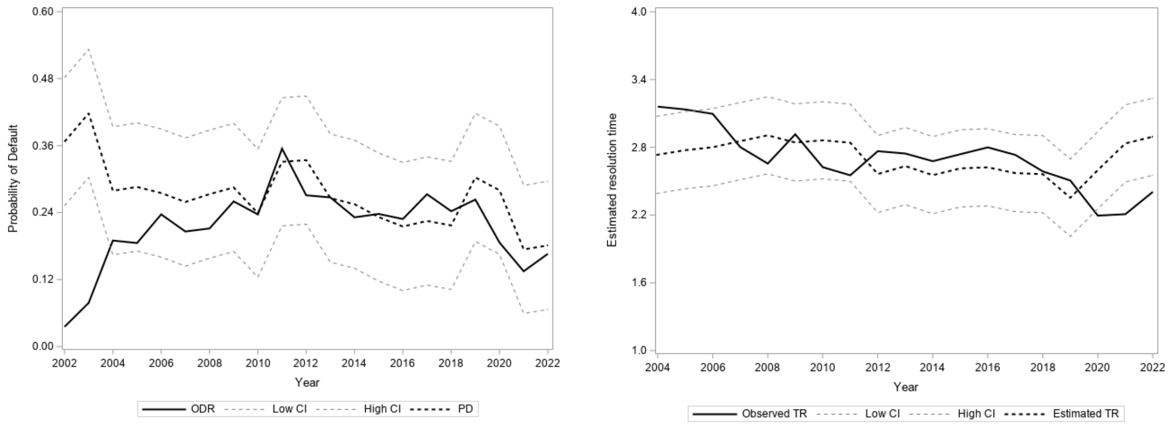
The resolution time (stand-alone model) decreases with the net income, turnover growth (insignificant), leverage, GDP growth, unemployment rate, and during low consumption periods and increases for interest, liquidity (insignificant), CAPEX (insignificant), insolvency severity, and cash rate. Most variables make sense in economic terms. For example, the resolution time is lower for greater profitability, indicating greater interest in the liquidated parts of a business. Leverage has a negative sign as greater leverage may indicate a lower number of assets to liquidate. Interestingly, firm age has a positive relationship with resolution time, while firm age squared has a negative relationship, indicating a nonlinear pattern in which resolution time initially increases with firm age but eventually decreases as firms become older. The impacts of macroeconomic factors can also be explained. An increase in cash rate may lengthen the resolution process as higher interests may reduce the stakeholder's willingness to engage in asset sales, while the increase in GDP growth creates more opportunities for firms to sell assets efficiently, hence shortening the resolution process. Interestingly, higher unemployment or low-consumption periods tend to accelerate resolution times, perhaps because firms in distress seek quicker resolutions during challenging economic conditions to conserve resources. All models control for industries and turnover classes. This is to discern the behaviour between different industries where possible. For example, the mining and farming industries may by nature take longer to resolve due to infrequent revenue, whereas food and retail should be faster to resolve.

The AUC is generally between 70 and 73% and pseudo-R-squares are around 5-7% for the failure model. The R-square is close to 17.5 for the resolution time which is comparable to other regression models for loss rates or recovery rates given failure. These values are in line with other empirical international studies. The number of observations explained in detail in Table 2 is due to the data sources relating to the population of Australian businesses.

Calibrated values from the extended model supersede those from the core model when choosing the fitted values. This is because the extended model incorporates more information, hence the prediction is likely more accurate. Despite that, core and extended models overlap to a large degree and the general economic findings are not sensitive to the results.

Figure 11 demonstrates the alignment between observed and fitted values regarding failure indicator and resolution time over time. The dashed lines show the 95% confidence intervals regarding the predicted values. Most of the observed failure rate points fall into the confidence intervals except the 2002 – 2003 PD. This could be due to the limited observations of this period. The number of observations in 2002 and 2003 is only a fraction of those in other years. The observed TR lies within the confidence intervals based on predicted TR nearly every year. Along with AUC and R-square, Figure 11 confirms the effectiveness of both failure and TR models.

Figure 11: Accuracy of PD and predicted TR over time



The comparison of observed event rates and average even probabilities over time shows that the probabilities are reasonably calibrated. Note that the Australian economy has experienced limited variation relative to other developed economies, particularly European and North American economies.

The event rates by industry and turnover categories are calibrated to the average event probabilities as a fixed effect has been included. The effect shifts the average event probability per category to the observed event rate per category (see e.g. Hamerle et al., 2006). There are some economic reasons for a mismatch, such as policy interventions or differences in economic resilience across industries, which may not be fully captured by the fixed effects or the overall model structure.

4.2. Calibration analysis

The analysis framework is based on a skeleton built by various categorical variables including year, firm age, industry, firm size, firm type, and income level. While the univariate analysis consecutively investigates the model calibration based on a single categorical variable, the multivariate analysis provides the examination of a combination of firm type and another categorical variable.

4.2.1. Univariate analysis

This analysis demonstrates the fluctuations in PD and estimated TR across various classes based on a single categorical variable, aiming to provide a comparison between the observed and predicted results.

Figure 12: PD and predicted TR over time

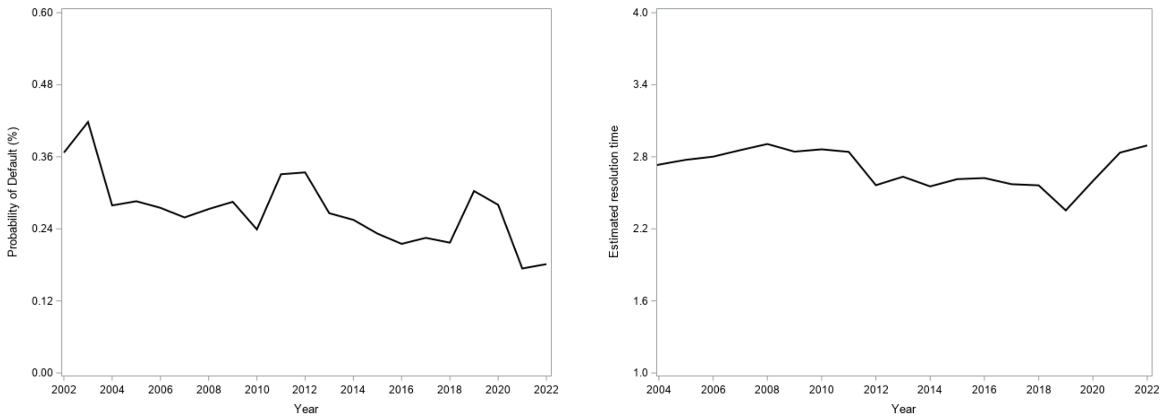


Figure 12 shows the fluctuations of PD and predicted TR over time. While the PD model captures the general trend and stays within the confidence intervals of the observed failure rate, it overestimates the failure rate before the GFC. However, it successfully predicts the increase in failure risk post-GFC and during the COVID-19 period, indicating the model's strength in capturing failure risk trends during economic downturns. For the resolution time model, the predicted values align well with the observed averages, both around 2.8 years, demonstrating that the TR model effectively estimates the duration of resolutions.

Figure 13: PD and predicted TR over firm age

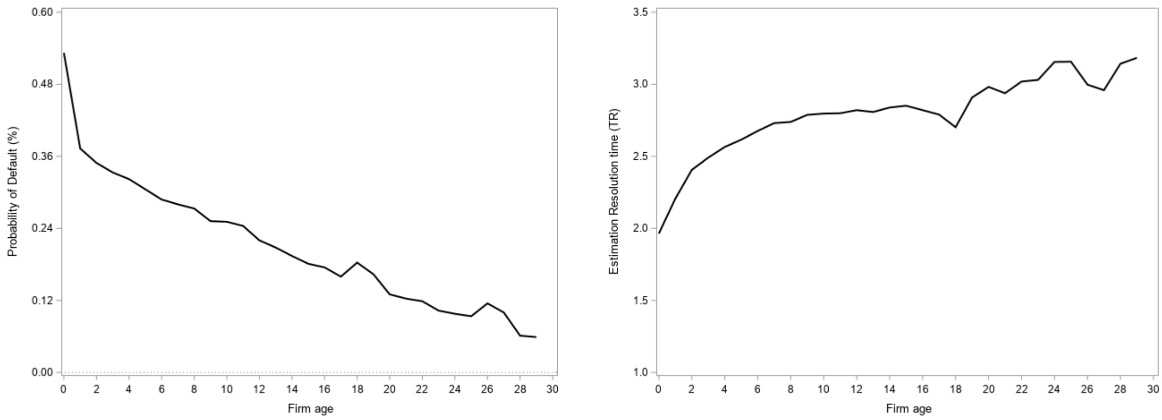


Figure 13 shows the fluctuations of PD and predicted TR across different firm age groups, with patterns consistent with the observed outcomes. Specifically, long-established businesses exhibit lower PDs but higher TRs in the event of failure. Over time, companies accumulate financial reserves and retained earnings, which help them navigate financial challenges. Additionally, a more established customer base, strong relationships with partners and suppliers, and a developed risk management system contribute to a lower failure risk. Conversely, resolution time slightly increases with firm age, which can be attributed to a larger asset base—including real estate, equipment, and intellectual property—and more complex financial obligations, all of which contribute to lengthier resolution processes.

Figure 14: PD and predicted TR over different industries

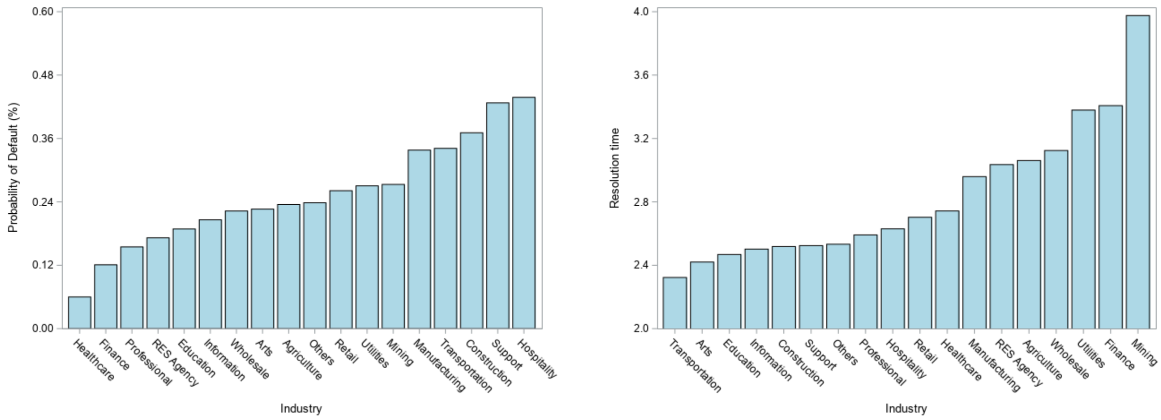


Figure 14 shows the changes in PD and predicted TR across different industries. The failure model's predictions align with the observed outcomes, with the Hospitality industry experiencing the highest PD and the Healthcare industry the lowest. Regarding the resolution time model, its prediction that businesses in the Mining industry have the longest TR is consistent with the observed data. However, at the lower end, the predictions do not align well with the observed values, suggesting that the TR model may have limitations in accurately capturing resolution time for industries with shorter TRs.

Figure 15: PD and predicted TR over different firm sizes

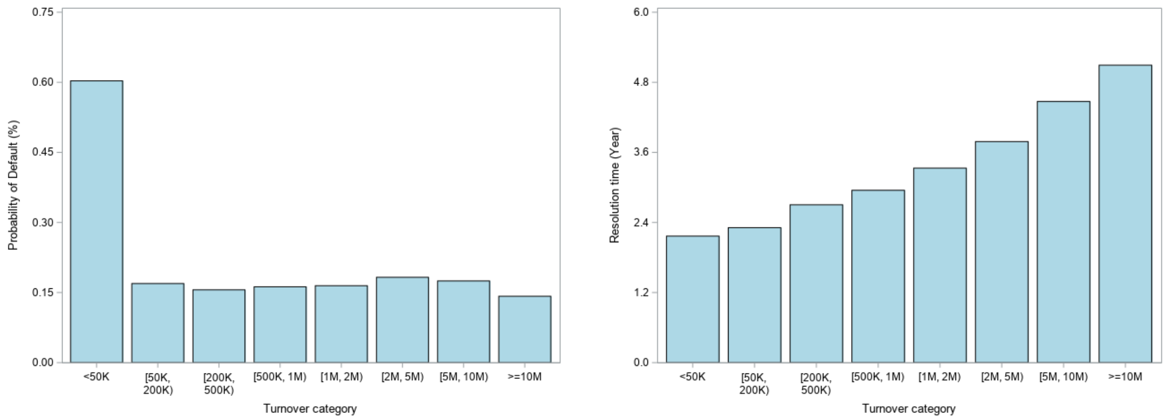


Figure 15 captures the fluctuations of PD and estimated TR across different firm sizes. The patterns derived from the predicted values are highly consistent with those observed in the actual data. This indicates that the models effectively capture the relationship between firm size and both failure risk and resolution time.

Figure 16: PD and predicted TR over two different firm types

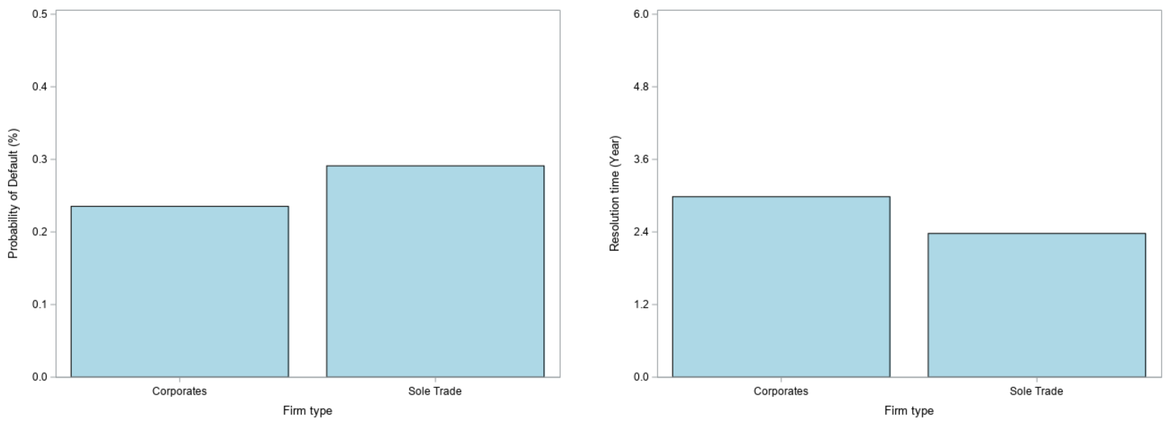


Figure 16 displays the variations in PD and predicted TR for two firm types—Corporates and Sole Traders. Since observed failure and resolution time data are unavailable for Sole Traders, a direct comparison with observed outcomes is not possible. Instead, we could investigate the risk profiles between sole traders and corporates using their fitted values. Calibrated results suggest that Sole Traders generally have higher PDs and shorter resolution times compared to Corporates. These findings are expected as large businesses tend to have a higher volume of assets. Moreover, they are also subject to stricter regulatory requirements and compliance measures. A more detailed analysis of the differences in failure risk between these two firm types will be provided in the following multivariate analysis section, examining the issue from multiple perspectives.

Figure 17: PD and predicted TR over income ranks

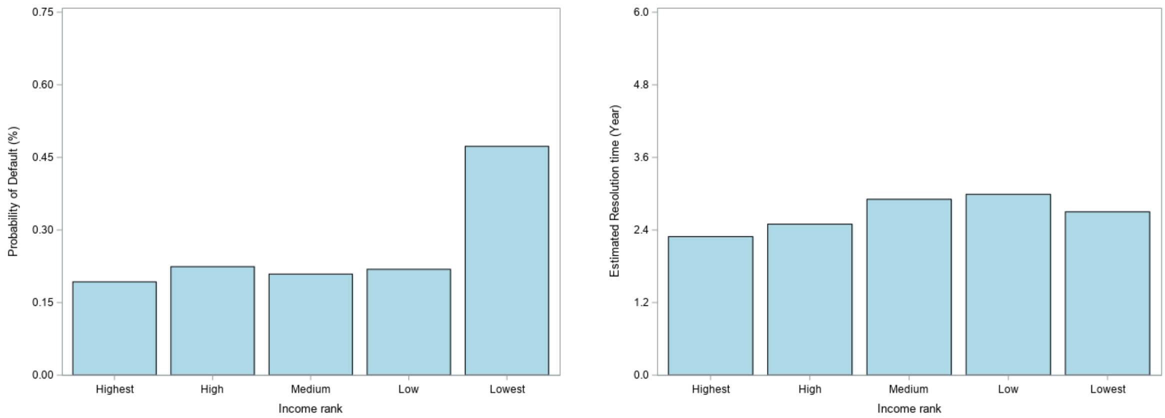


Figure 17 presents the differences in PD and estimated TR across various income groups. Businesses in the lowest income group are estimated to have the highest PD levels, which aligns with expectations since higher income typically indicates better loan serviceability and liquidity, thereby reducing failure risk. In terms of resolution time, the variations across income groups are relatively small. However, medium- and low-income businesses tend to have longer resolution times compared to those with higher incomes, possibly due to fewer resources to expedite the resolution process.

4.2.2. Multivariate analysis

The calibration helps to obtain the estimation of failure probability and resolution time of sole traders, allowing for the comparison of risk levels across different categories between corporates and sole traders.

Figure 18: PD and estimated TR over time for corporates and sole traders

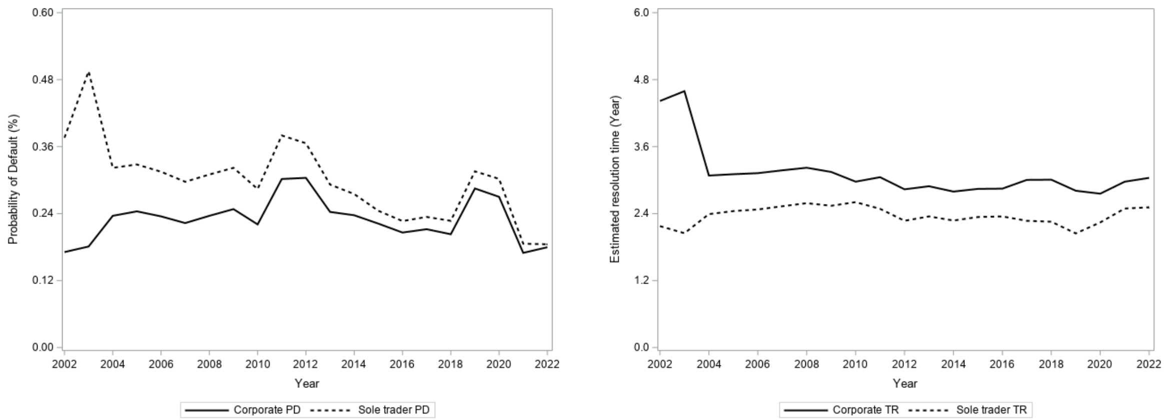


Figure 18 shows the PD and estimated TR over time for corporates (solid line) and sole traders (dashed line). Sole traders consistently exhibit higher PDs compared to corporates, reflecting their greater vulnerability to financial stress due to limited resources and less diversified income streams. Before the GFC, the gap between the PDs of sole traders and corporates was larger. However, after the GFC, this difference narrowed, suggesting that corporates faced greater challenges post-crisis, perhaps due to tighter credit conditions, regulatory changes, or shifts in economic conditions that impacted larger businesses more heavily. While still at a higher risk, sole traders may have become less disproportionately affected as the market adjusted to new financial conditions.

An opposite pattern is observed for resolution time, where corporates consistently have longer resolution times than sole traders. This is likely due to the more complex business structures of corporates, which involve multiple assets, liabilities, and legal obligations, making the resolution process more time-consuming. In contrast, sole traders have simpler operations and fewer assets, allowing for quicker liquidation when they fail. Note that the significant differences in PD and resolution time observed in 2002 and 2003 may be influenced by the limited number of observations during those years.

Figure 19: PD and estimated TR over firm age for sole traders and corporates

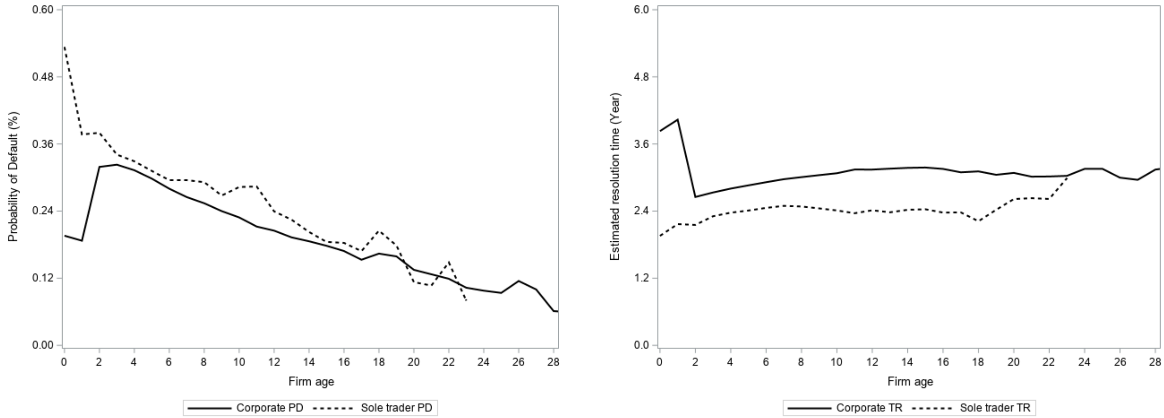


Figure 19 illustrates the PD and estimated TR by firm age for both corporates and sole traders. Regardless of various firm age groups, sole traders exhibit higher PDs, but lower resolution times compared to corporates. These differences again lie in the financial stability, complexity of business structures, and operations between the two types of firms.

Figure 20: PD and estimated TR over the industry for sole traders and corporates

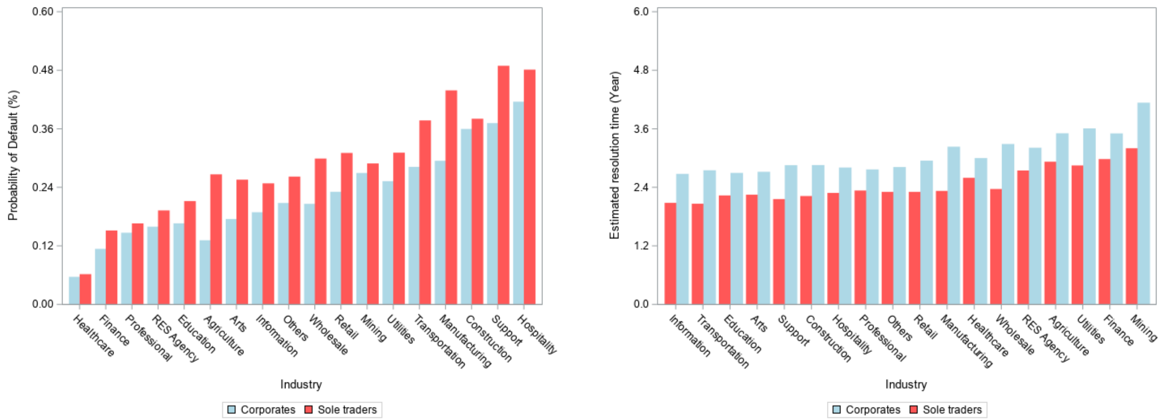


Figure 20 highlights the differences in PD and estimated TR between corporates and sole traders across various industries. The findings show that the industries with the highest and lowest PDs are consistent between the two types of firms. Specifically, the Hospitality industry, regardless of business structure, experiences the highest PD, while the Healthcare industry shows the lowest PD in both corporates and sole traders. In terms of resolution time, variations across industries are relatively small. However, businesses in the Mining industry face the longest resolution times due to their specialised machinery, which is difficult to liquidate during resolution. Notably, sole traders consistently exhibit higher PDs and shorter resolution times than corporates across all industries.

Figure 21: PD and estimated TR over firm size for sole traders and corporates

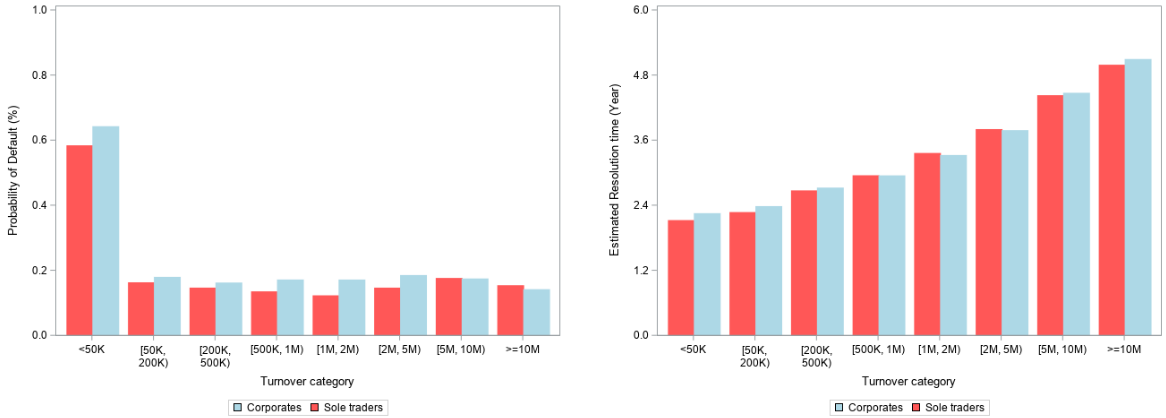


Figure 21 presents the differences in PD and estimated TR between corporates and sole traders across various firm sizes. Interestingly, these differences are less pronounced compared to other categories previously discussed. For smaller firms, corporates have a slightly higher PD than sole traders. In contrast, large sole traders experience higher PD than large corporates. In terms of resolution time, corporates and sole traders show similar liquidation times across all firm sizes, suggesting that firm size has less impact on resolution time for both business structures.

Figure 22: PD and estimated TR over income class for sole traders and corporates

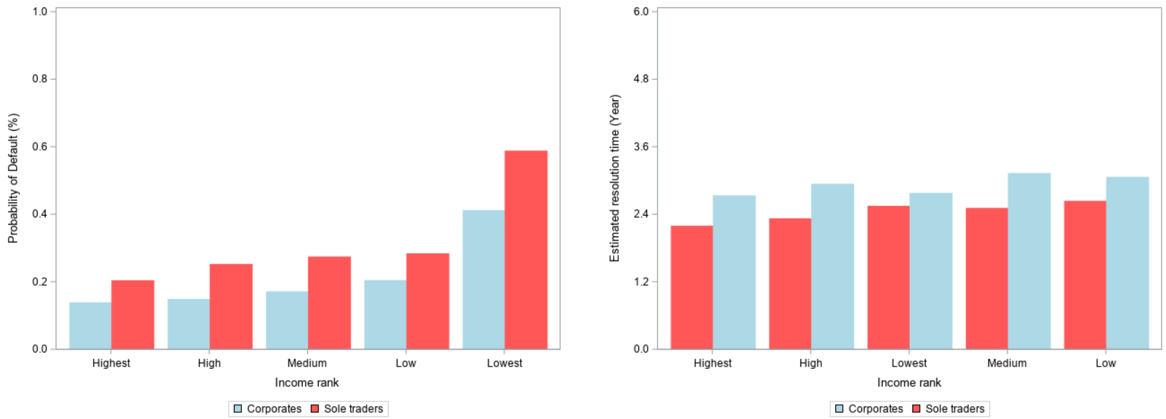


Figure 22 presents the PD and estimated TR for corporates and sole traders across various income groups. The findings indicate that sole traders consistently have higher PDs but lower resolution times compared to corporates. The lower PDs for corporates are expected, as their larger operational scope enables them to generate higher income, improving loan serviceability and reducing failure risk. Additionally, higher income allows corporates to grow in size and complexity, which likely explains the longer resolution times when they become insolvent.

4.2.3. Sensitivity analysis

We investigate how changes in risk factors drive the business risk levels by stressing each factor to increase from 1% to 10% and recalibrating the probability of failure and resolution time.

Table 6: Sensitivity tests

	Percentage change									
Risk factor	**1%**	**2%**	**3%**	**4%**	**5%**	**6%**	**7%**	**8%**	**9%**	**10%**
Panel A: Absolute changes in Probability of failure (%)										
Turnover growth	0.122	0.122	0.122	0.122	0.122	0.122	0.122	0.122	0.122	0.122
Income ratio	0.144	0.166	0.188	0.210	0.233	0.255	0.278	0.302	0.325	0.349
Interest ratio	0.130	0.138	0.146	0.154	0.163	0.171	0.179	0.188	0.196	0.205
Liquidity ratio	0.155	0.187	0.220	0.253	0.287	0.321	0.356	0.391	0.427	0.463
Leverage	0.123	0.124	0.126	0.127	0.128	0.129	0.130	0.131	0.132	0.133
CAPEX	0.124	0.126	0.127	0.129	0.131	0.132	0.134	0.136	0.138	0.139
Cash rate	0.123	0.123	0.123	0.123	0.124	0.124	0.124	0.124	0.125	0.125
GDP growth change	0.126	0.129	0.133	0.136	0.140	0.143	0.147	0.150	0.154	0.157
UER change	0.123	0.123	0.123	0.123	0.124	0.124	0.124	0.125	0.125	0.125
Panel B: Absolute changes in Resolution time (months)										
Turnover growth	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045
Income ratio	0.044	0.044	0.044	0.044	0.043	0.043	0.043	0.043	0.043	0.042
Interest ratio	0.044	0.044	0.044	0.044	0.043	0.043	0.043	0.043	0.042	0.042
Liquidity ratio	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045
Leverage	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045
CAPEX	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045
Cash rate	0.041	0.038	0.035	0.031	0.028	0.025	0.021	0.018	0.014	0.011
GDP growth change	0.046	0.047	0.048	0.049	0.050	0.051	0.052	0.054	0.055	0.056
UER change	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.044

Table 6 presents the results of sensitivity tests, with absolute changes in PD (Panel A) measured in basis points and absolute changes in resolution time (Panel B) in months. In the PD sensitivity test, firm-level factors with the strongest influence on PD are the liquidity ratio and income ratio. An increase in the liquidity ratio from 1% to 10% raises PD from 16 bps to 46 bps, while a similar rise in the income ratio escalates PD from 14 bps to 35 bps. These findings underscore the critical roles of liquidity and loan serviceability in determining a business's risk level. In comparison, other firm-level factors, such as the interest ratio, CAPEX, and leverage, exert less influence. A 10% increase in the interest ratio raises PD by 21 bps, while CAPEX and leverage contribute increases of 2 bps and 1 bp, respectively. Turnover growth has a negligible effect. Regarding macroeconomic factors, GDP growth significantly lowers PD by 3 bps to 34 bps, while an increase in the cash rate drives PD higher by 2 bps to 15 bps, confirming the importance of economic conditions in shaping failure risk.

For the TR sensitivity test, changes in firm-level factors appear to have a similar and minimal influence on resolution time. The maximum recorded change in TR is only 0.05 months, confirming that internal financial factors play a minor role in determining resolution time. Among macroeconomic variables, a 10% increase in the cash rate extends resolution time by approximately 0.01 months, while a similar increase in GDP growth and the unemployment rate results in changes to TR of roughly 0.06 months and 0.04 months, respectively. These findings suggest that the impacts of both firm-specific financial factors and macroeconomic conditions on resolution time are modest and largely equivalent, indicating that neither set of variables exerts a dominant influence. The limited number of observations for resolution time, coupled with the absence of precise resolution completion dates, restricts the model's ability to capture variations effectively.

5. Robustness checks

5.1. Failure 1 year and 3 years (vs. 2 years)

Alternative indicators define failure flags when any insolvency event occurs within 2 or 3 years after the last reporting period. The failure and TR models are re-estimated using these new failure and TR outcomes, and the results show that all coefficients remain highly consistent with those from the main models. Table 7 outlines the performance of these alternative models.

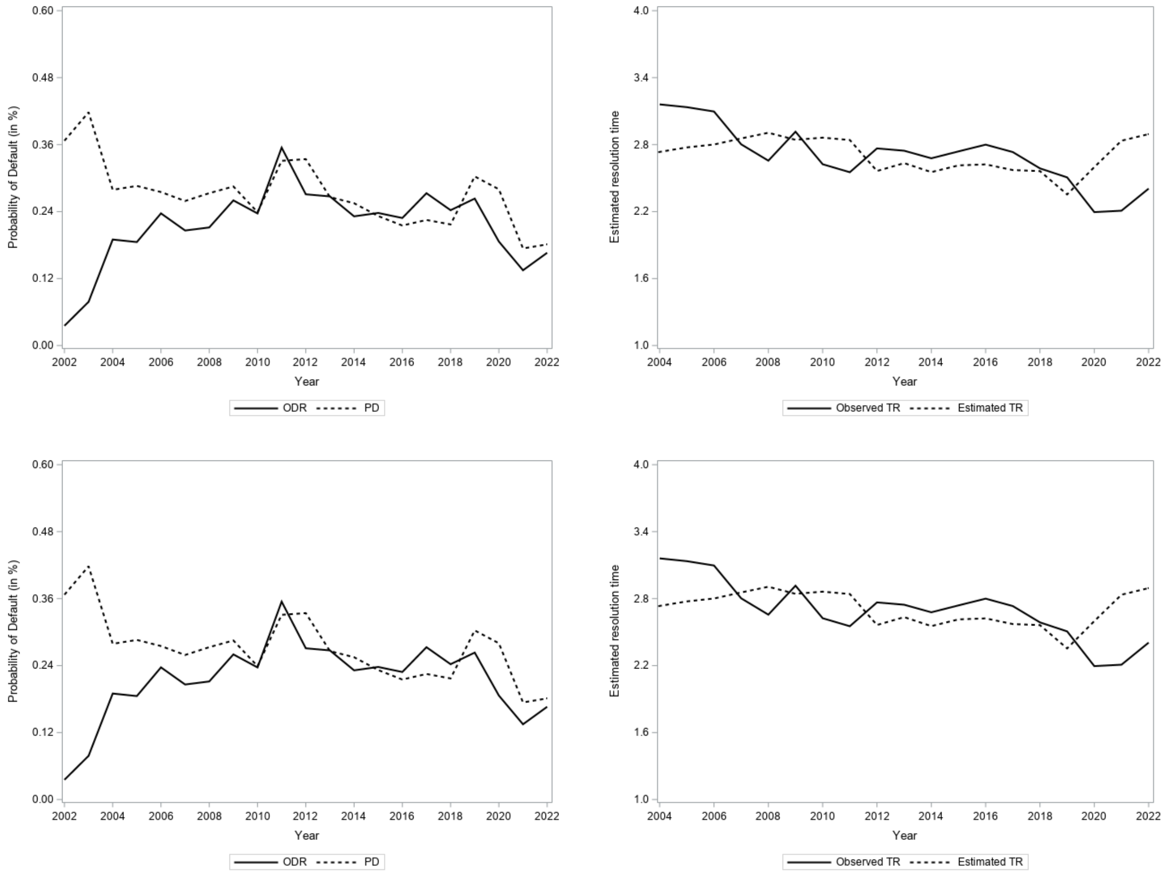
Table 7: Failure and TR model performances regarding alternative failure indicators

Model	**No of obs.**	**AUC**	**R**
Failure 2 years			
Failure - Core	7,221,668	0.713	0.058
Failure - Extended	6,706,440	0.742	0.075
TR – Core	18,901	0.652	0.148
TR – Extended	16,853	0.674	0.182
Failure 3 years			
Failure - Core	7,222,665	0.713	0.056
Failure - Extended	6,707,307	0.741	0.073
TR – Core	19,790	0.669	0.156
TR – Extended	17,586	0.695	0.198

Models using alternative failure indicators show slightly better performance compared to the original model, though the improvement is negligible. There is only a minimal increase in AUC and R-square compared to the results presented in Table 7. We also plotted the observed and predicted values for both failure and resolution time over time to assess the calibration of the alternative models (see, Figure 23). Similar alignments were observed, demonstrating the robustness of both models when using the alternative outcomes.

Figure 23: Observed vs. Predicted failure and resolution over time

Note: This figure shows the observed and fitted failure (top left) and time to resolution (top right) where failure occurs within 2 years of the last reporting period; and the observed and fitted failure (bottom left) and time to resolution (bottom right) where failure occurs within 3 years of the last reporting period.



5.2. Machine learning methodology

There are some further modelling complexities that the area of machine learning has recently highlighted. Interactions between features and non-linear relations between the features and outcome variables may exist. For example, it has been shown that the observed failure rate increases at an increasing rate when features indicate a greater level of financial distress (e.g., leverage of a firm). Linear models do not model such as convex relationships, or non-linear logistic regressions to some degree, and machine learning methods can apply these as they bin features (e.g., using a decision tree) and feature combinations and compute the observed average outcome variable per bin.[24]

In this paper, a random forest is used, then simulated over randomisation for bins. As several iterations are aggregated, this method is also called an ensemble. The reliance on the average of observed outcomes, rather than weighting using fitted parameters as in regressions, implies that random forests and some other machine learning methods are non-parametric. Here is the fitted probability of an event E for a random forest:

[24] IRB banks apply a reduction for SMEs based on turnover (categories 1-8). This extension has been tested in unreported results. It has no impact on our results and would imply a monotone transformation for almost all businesses. It is not included in the paper as it introduces unnecessary complexity and it only applies to IRB lenders.

$$PE_{it} = \sum_{k=1}^{K} \frac{1}{n_s} \sum_{s=1}^{S} E_{ksit} \quad (1)$$

with iteration k, k = 1, . . . , K and feature based sub-sample s, s = 1, . . . ,S.

Random forests have hyperparameters that can be tuned to add a greater level of complexity to the models (see e.g., Couronné et al. 2018). Two random forests are analysed: first with the following hyperparameter: maximum depth is 5, the maximum number of features is 25, the number of trees is 100, and the minimum number of sample splits is 2; second, hyperparameters tuning is performed using a grid search method to find the optimal parameters.[25]

We compare linear (OLS) models, non-linear models (logistic regressions for event data and loglinear model for resolution times), a random forest, and a random forest with hyperparameter tuning to include interactions between features and also non-linear relations between features and outcome variables.

Further, models may be too specific and not accurate with regard to the data-generating process and hence not able to predict future outcomes. Thus, the following analysis tests whether the results hold up for 50% of the data (test sample) whilst using the complementary data as a training sample. For simplicity, we have halved all data sets reported in Table 2.[26]

Table 8 shows the Area under the Receiver Operating Characteristics Curve (AUC, closer to one=better) as a measure of discrimination, the square root of the Brier score (SQR-BS, lower=better) as a measure for calibration for the failure model, R-square (R2, closer to one=better), and the root mean squared error (RMSE, lower better) for resolution time. These measures are common and well explained in the credit risk literature (see e.g., Rösch & Scheule, 2020).

Table 8: Analysing linear, non-linear, and machine learning models for failure probabilities, in-sample and out-of-sample

Note: This table shows the performance of in-sample and out-of-sample estimates of four models: a linear model, a non-linear model (logistic regression for insolvency events and a log-linear model of resolution times), and a random forest with and without hyperparameter tunings. We highlight the best-performing model for the out-of-sample test and in the following analysis have compared this model to the probabilities to failure and fitted resolution times.

	In-sample		**Out-of-sample**	
Model	**AUC/R2**	**SQR-BS/ RMSE**	**AUC/R2**	**SQR-BS/RMSE**
Failure				
Linear	0.749	0.045	0.750	0.045
Non-linear	0.757	0.045	0.756	0.045
Random Forest	0.764	0.045	0.771	0.045
Random Forest (CV)	0.787	0.045	0.776	0.045
TR				
Linear	0.131	2.064	0.138	2.048
Non-linear	0.118	3.003	0.127	2.981
Random Forest	0.124	2.073	0.155	2.030
Random Forest (CV)	0.118	2.085	0.069	2.137

There are two main differences: Firstly, the out-of-sample models may be less accurate as they are based

[25] The search grid includes maximum depth: 2, 10, 5, maximum number of features: 11, 8, 33, 4, the number of trees: 167, 168, 14, 14, the minimum number of sample splits: 3, 6, 7, 5.

[26] Note that for the sample for resolution times an additional filter had to be imposed to ensure values greater than zero so that natural logarithm of resolution times could be calculated for the log-linear model. This results in a drop in observation count from 14,595 to 14,422 by 1.1%.

on test data that did not go into model estimation. Secondly, some models may be more accurate as they consider convexity and interactions.

The models have been re-run for the best fit and our results are consistent. The results of the factor models and the following analyses are consistent. The correlations between the best-performing model and the presented model are 53.6% for failure and 64.9% for resolution time.

6. Industry impact and discussion

The paper investigates failure risk by analysing failure probability and resolution time from various perspectives, including time, firm age, firm size, industry, firm type, and income level. The findings reveal that businesses in Australia have experienced a decrease in risk, characterized by lower failure probabilities and shorter resolution times over time, despite minor spikes during the COVID-19 pandemic. Younger and smaller firms exhibit a higher risk of failure. Additionally, the time taken to resolve insolvency increases with firm size and age. Among industries, discretionary sectors such as Hospitality show the highest failure rates, while non-discretionary sectors like Healthcare industry demonstrate a lower likelihood of failure. In terms of resolution time, capital-intensive industries, such as Mining, have the longest average resolution times, whereas human-intensive industries, such as Arts, have the shortest. Corporates generally have lower failure risks and longer resolution times compared to sole traders; however, these patterns do not consistently apply across different firm sizes. Furthermore, firms with lower income levels face higher failure risks and slightly shorter resolution times compared to their higher-income counterparts.

The probability of failure decreases with the net income, turnover growth, liquidity CAPEX, firm age, and GDP growth and increases for interest, leverage, cash rate, unemployment rate, and during low-consumption periods. The resolution time decreases with the net income, turnover growth, leverage, and firm age squared and increases for interest, liquidity, CAPEX, insolvency severity, cash rate, and firm age. The probability of failure and time to resolution move in parallel after controlling for industries, age, and turnover classes.

Government datasets on tax filings, business registrations, insolvencies, and security registrations are valuable sources of information. Over time, broader public access to these datasets could enable more detailed analyses beyond the scope of this paper. The research team identified several areas where existing data collection could be expanded to support future research. One key area is the inclusion of non-corporate personal insolvency data—particularly those involving sole traders—which is currently not captured in the dataset used in this study. This information can be sourced from the Australian Financial Security Authority (AFSA). It is recommended that the ABS consider working with AFSA to incorporate these insolvency records into the broader sample, which would improve the model's ability to capture risk dynamics across the full spectrum of business types. Additional enhancements include expanding business characteristics surveys in both coverage and content and collecting data on loan amounts sought and approved. While the current survey sample of 10,000 businesses aligns with international standards, there remains scope for deeper insights. Furthermore, as tax filings primarily detail income and expenditure, they provide limited visibility into business assets and liabilities. Requiring businesses to report on their assets and wealth could strengthen data quality, not only by enriching the information available for analysis, but also through the potential disciplining effect such transparency may bring.

This paper has focused on business features from tax filings. Future research may look into quantifying the lending opportunity on a national level in terms of needed requested and approved credit amounts after the above-discussed data limitations are overcome. Also, the comprehensive collection and analysis of other features may be of interest to future research. This may include non-financial information such as demographics (e.g., year in operation, director tenure, etc.) and validated behavioural (e.g., business or customer transactions) available through digital footprints and open banking.

APPENDIX

Appendix A: Detailed description of insolvency rate

Table A1: Observed insolvency rate by year

FY	Number of insolvency events	Number of firms	Insolvency rate (%)
2001-02	259	538,230	0.048
2002-03	968	571,710	0.169
2003-04	2,225	591,065	0.376
2004-05	2,649	609,135	0.435
2005-06	3,562	619,961	0.575
2006-07	3,408	632,837	0.539
2007-08	3,548	642,226	0.552
2008-09	4,466	641,780	0.696
2009-10	4,165	645,004	0.646
2010-11	4,436	658,457	0.674
2011-12	4,916	674,029	0.729
2012-13	4,952	690,036	0.718
2013-14	4,450	711,738	0.625
2014-15	4,276	731,433	0.585
2015-16	4,646	754,112	0.616
2016-17	3,861	780,469	0.495
2017-18	3,802	809,959	0.469
2018-19	3,991	836,731	0.477
2019-20	3,699	856,172	0.432
2020-21	2,105	882,496	0.239
2021-22	2,485	904,998	0.275
Total	**72,869**	**14,782,578**	**0.493**

Table A2: Observed insolvency rate by firm age

Firm age	No of insolvency events	Number of firms	Insolvency rate (%)	Firm age	Number of insolvency events	Number of firms	Insolvency rate (%)
0	63	1,079,208	0.006	**15**	1,713	387,619	0.442
1	1,197	1,288,514	0.093	**16**	1,561	348,983	0.447
2	5,144	1,191,562	0.432	**17**	1,246	314,475	0.396
3	7,533	1,064,614	0.708	**18**	1,086	283,320	0.383
4	7,748	980,765	0.790	**19**	967	254,149	0.380
5	7,241	892,934	0.811	**20**	827	227,861	0.363
6	6,174	778,198	0.793	**21**	609	204,539	0.298
7	5,479	682,406	0.803	**22**	583	176,423	0.330
8	4,625	615,575	0.751	**23**	434	153,434	0.283
9	3,965	708,285	0.560	**24**	321	134,211	0.239
10	3,342	638,441	0.523	**25**	212	104,927	0.202
11	3,103	573,816	0.541	**26**	204	82,017	0.249
12	2,739	517,904	0.529	**27**	163	75,552	0.216
13	2,314	469,147	0.493	**28**	95	68,331	0.139
14	2,117	427,820	0.495	**29**	64	57,548	0.111

Table A3: Observed insolvency rate by industry from 2001-02 to 2021-22

Industry	No of insolvency	Number of firms	Insolvency rate (%)
Agriculture	1,008	344,950	0.292
Mining	464	103,384	0.449
Manufacturing	6,424	918,591	0.699
Utilities	369	57,429	0.643
Construction	18,475	2,488,727	0.742
Wholesale	4,554	971,540	0.469
Retail	5,986	1,088,066	0.550
Hospitality	5,833	630,459	0.925

Transportation	4,834	810,079	0.597
Information	880	237,006	0.371
Finance	2,330	992,994	0.235
RES Agency	3,459	908,659	0.381
Professional	7,183	2,528,319	0.284
Support	4,773	622,722	0.766
Education	721	209,204	0.345
Healthcare	910	685,388	0.133
Arts	728	166,117	0.438
Others	3,082	646,268	0.477
Missing	856	372,676	0.230
Total	**72,869**	**14,782,578**	**0.493**

Table A4: Observed insolvency rate by turnover category from 2001-02 to 2021-22

Turnover category	No of insolvency	Number of firms	Insolvency rate (%)
<50k	40,726	3,718,672	1.095
50k to <200k	9,300	3,137,780	0.296
200k to <500k	7,667	2,798,845	0.274
500k to <1M	5,249	1,798,592	0.292
1M to <2M	4,071	1,332,522	0.306
2M to <5M	3,290	1,051,209	0.313
5M to <10M	1,289	417,786	0.309
>= 10M	1,277	527,172	0.242
Total	**72,869**	**14,782,578**	**0.493**

Table A5: Proportion of insolvency events across Industry x Turnover categories from 2001-02 to 2021-22

Industry	<50k	50k to <200k	200k to <500k	500k to <1M	1M to <2M	2M to <5M	5M to <10M	>= 10M	Total
Agriculture	514	123	114	84	73	56	22	22	1,008
	0.705	*0.169*	*0.156*	*0.115*	*0.100*	*0.077*	*0.030*	*0.030*	*1.383*
Mining	215	48	42	33	28	32	21	45	464
	0.295	*0.066*	*0.058*	*0.045*	*0.038*	*0.044*	*0.029*	*0.062*	*0.637*
Manufacturing	2,916	888	788	589	471	432	167	173	6,424
	4.002	*1.219*	*1.081*	*0.808*	*0.646*	*0.593*	*0.229*	*0.237*	*8.816*
Utilities	206	41	40	22	20		40		369
	0.283	*0.056*	*0.055*	*0.030*	*0.027*		*0.055*		*0.506*
Construction	10,787	2,166	1,712	1,232	1,070	851	353	304	18,475
	14.803	*2.972*	*2.349*	*1.691*	*1.468*	*1.168*	*0.484*	*0.417*	*25.354*
Wholesale	2,302	543	483	365	296	298	141	126	4,554
	3.159	*0.745*	*0.663*	*0.501*	*0.406*	*0.409*	*0.193*	*0.173*	*6.250*
Retail	3,132	744	690	480	376	286	138	140	5,986
	4.298	*1.021*	*0.947*	*0.659*	*0.516*	*0.392*	*0.189*	*0.192*	*8.215*
Hospitality	3,275	762	764	489	298	172	46	27	5,833
	4.494	*1.046*	*1.048*	*0.671*	*0.409*	*0.236*	*0.063*	*0.037*	*8.005*
Transportation	2,696	654	494	338	265	221	63	103	4,834
	3.700	*0.898*	*0.678*	*0.464*	*0.364*	*0.303*	*0.086*	*0.141*	*6.634*
Information	511	122	91	54	43	31	14	14	880
	0.701	*0.167*	*0.125*	*0.074*	*0.059*	*0.043*	*0.019*	*0.019*	*1.208*
Finance	1,422	313	189	135	81	97	33	60	2,330
	1.951	*0.430*	*0.259*	*0.185*	*0.111*	*0.133*	*0.045*	*0.082*	*3.198*
RES Agency	1,961	416	374	250	167	171	57	63	3,459
	2.691	*0.571*	*0.513*	*0.343*	*0.229*	*0.235*	*0.078*	*0.086*	*4.747*
Professional	4,325	1,002	710	433	335	228	77	73	7,183
	5.935	*1.375*	*0.974*	*0.594*	*0.460*	*0.313*	*0.106*	*0.100*	*9.857*
Support	2,828	564	448	325	248	213	80	67	4,773
	3.881	*0.774*	*0.615*	*0.446*	*0.340*	*0.292*	*0.110*	*0.092*	*6.550*
Education	381	110	83	51	44	29	13	10	721
	0.523	0.151	0.114	0.070	0.060	0.040	0.018	0.014	0.989

Healthcare	521	111	102	72	56	28	20		910
	0.715	*0.152*	*0.140*	*0.099*	*0.077*	*0.038*	*0.027*		*1.249*
Arts	440	121	85	37	28	17			728
	0.604	*0.166*	*0.117*	*0.051*	*0.038*	*0.023*			*0.999*
Others	1,783	463	372	211	132	73	28	20	3,082
	2.447	*0.635*	*0.511*	*0.290*	*0.181*	*0.100*	*0.038*	*0.027*	*4.230*
Missing	511	109	86	49	40	39	11	11	856
	0.701	*0.150*	*0.118*	*0.067*	*0.055*	*0.054*	*0.015*	*0.015*	*1.175*
Total	**40,726**	**9,300**	**7,667**	**5,249**	**4,071**	**3,290**	**1,289**	**1,277**	**72,869**
	55.889	***12.763***	***10.522***	***7.203***	***5.587***	***4.515***	***1.769***	***1.752***	***100***

The first rows show the number of insolvency events and the second rows (italic) shows the proportion (in %) across industry x turnover group. The merged cells are subject to the clearance rule of 10 of ABS requiring all cells to have at least 10 observations.

Table A6: Proportion of insolvency events across firm age x turnover categories

	<50k	**50k to <200k**	**200k to <500k**	**500k to <1M**	**1M to <2M**	**2M to <5M**	**5M to <10M**	**>= 10M**	**Total**
0	30	18				15			63
	0.041	*0.025*				*0.021*			*0.086*
1	289	220	258	169	135	87	23	16	1,197
	0.397	*0.302*	*0.354*	*0.232*	*0.185*	*0.119*	*0.032*	*0.022*	*1.643*
2	2,494	723	683	487	361	246	81	69	5,144
	3.423	*0.992*	*0.937*	*0.668*	*0.495*	*0.338*	*0.111*	*0.095*	*7.059*
3	4,215	1,011	842	568	418	278	91	110	7,533
	5.784	*1.387*	*1.155*	*0.779*	*0.574*	*0.382*	*0.125*	*0.151*	*10.338*
4	4,632	969	801	506	384	282	102	72	7,748
	6.357	*1.330*	*1.099*	*0.694*	*0.527*	*0.387*	*0.140*	*0.099*	*10.633*
5	4,362	895	719	488	349	251	99	78	7,241
	5.986	*1.228*	*0.987*	*0.670*	*0.479*	*0.344*	*0.136*	*0.107*	*9.937*
6	3,686	753	632	415	285	230	95	78	6,174
	5.058	*1.033*	*0.867*	*0.570*	*0.391*	*0.316*	*0.130*	*0.107*	*8.473*
7	3,303	671	495	356	271	223	97	63	5,479
	4.533	*0.921*	*0.679*	*0.489*	*0.372*	*0.306*	*0.133*	*0.086*	*7.519*
8	2,700	575	449	304	239	198	87	73	4,625
	3.705	*0.789*	*0.616*	*0.417*	*0.328*	*0.272*	*0.119*	*0.100*	*6.347*
9	2,328	507	362	264	201	180	54	69	3,965
	3.195	*0.696*	*0.497*	*0.362*	*0.276*	*0.247*	*0.074*	*0.095*	*5.441*
10	1,932	422	320	212	171	162	65	58	3,342
	2.651	*0.579*	*0.439*	*0.291*	*0.235*	*0.222*	*0.089*	*0.080*	*4.586*
11	1,681	386	319	221	191	157	77	71	3,103
	2.307	*0.530*	*0.438*	*0.303*	*0.262*	*0.215*	*0.106*	*0.097*	*4.258*
12	1,474	357	273	201	171	144	58	61	2,739
	2.023	*0.490*	*0.375*	*0.276*	*0.235*	*0.198*	*0.080*	*0.084*	*3.759*
13	1,246	299	255	161	143	109	48	53	2,314
	1.710	*0.410*	*0.350*	*0.221*	*0.196*	*0.150*	*0.066*	*0.073*	*3.176*
14	1,126	293	244	135	134	99	33	53	2,117
	1.545	*0.402*	*0.335*	*0.185*	*0.184*	*0.136*	*0.045*	*0.073*	*2.905*
15	965	215	154	99	95	90	45	50	1,713
	1.324	*0.295*	*0.211*	*0.136*	*0.130*	*0.124*	*0.062*	*0.069*	*2.351*
16	836	182	149	116	97	91	39	51	1,561
	1.147	*0.250*	*0.204*	*0.159*	*0.133*	*0.125*	*0.054*	*0.070*	*2.142*
17	647	154	114	98	70	76	27	60	1,246
	0.888	*0.211*	*0.156*	*0.134*	*0.096*	*0.104*	*0.037*	*0.082*	*1.710*
18	580	124	117	65	62	72	31	35	1,086
	0.796	*0.170*	*0.161*	*0.089*	*0.085*	*0.099*	*0.043*	*0.048*	*1.490*
19	498	119	111	69	57	52	31	30	967
	0.683	*0.163*	*0.152*	*0.095*	*0.078*	*0.071*	*0.043*	*0.041*	*1.327*
20	396	98	99	58	52	63	18	43	827
	0.543	*0.134*	*0.136*	*0.080*	*0.071*	*0.086*	*0.025*	*0.059*	*1.135*

21	289 *0.397*	73 *0.100*	57 *0.078*	56 *0.077*	47 *0.064*	45 *0.062*	18 *0.025*	24 *0.033*	609 *0.836*
22	286 *0.392*	65 *0.089*	63 *0.086*	58 *0.080*	34 *0.047*	42 *0.058*	21 *0.029*	14 *0.019*	583 *0.800*
23	238 *0.327*	45 *0.062*	39 *0.054*	31 *0.043*	38 *0.052*	20 *0.027*	23 *0.032*		434 *0.596*
24	165 *0.226*	39 *0.054*	28 *0.038*	31 *0.043*	16 *0.022*	28 *0.038*	14 *0.019*		321 *0.441*
25	86 *0.118*	32 *0.044*	19 *0.026*	23 *0.032*	18 *0.025*	19 *0.026*	15 *0.021*		212 *0.291*
26	89 *0.122*	27 *0.037*	19 *0.026*	20 *0.027*	10 *0.014*	19 *0.026*	20 *0.027*		204 *0.280*
27	69 *0.095*	16 *0.022*	22 *0.030*	18 *0.025*	14 *0.019*	12 *0.016*	12 *0.016*		163 *0.224*
28	50 *0.069*	13 *0.018*	11 *0.015*	10 *0.014*	11 *0.015*				95 *0.130*
29	34 *0.047*	30 *0.041*							64 *0.088*
Total	**40,726** ***55.889***	**9,300** ***12.763***	**7,667** ***10.522***	**5,249** ***7.203***	**4,071** ***5.587***	**3,290** ***4.515***	**1,289** ***1.769***	**1,277** ***1.752***	**72,869** ***100.000***

The first rows show the number of insolvency events and the second rows (italic) shows the proportion (in %) across industry x turnover group. The merged cells are subject to the clearance rule of 10 of ABS requiring all cells to have at least 10 observations.

Appendix B: Factor models with fixed effects

Note: This table is the extension of those presented in Table 5 by showing the coefficients on all fixed effects related to turnover categories and industries.

		Core model		Extended model	
Feature		**Failure**	**Resolution time**	**Failure**	**Resolution time**
Income ratio		-0.877*** (0.025)	-0.147** (0.07)	-0.574*** (0.028)	-0.324*** (0.078)
Interest ratio		4.347*** (0.209)	4.35*** (0.576)	4.553*** (0.22)	2.894*** (0.606)
Turnover growth		-0.022** (0.009)	0.003 (0.018)	-0.018* (0.009)	-0.016 (0.019)
Liquidity				-0.56*** (0.013)	0.035 (0.024)
Leverage				0.014* (0.007)	-0.057*** (0.013)
CAPEX				-1.53*** (0.118)	0.186 (0.278)
	Processing liquidation				0.471*** (0.102)
Insolvency type	Receivership				0.802*** (0.107)
	Court liquation				0.642*** (0.041)
Age		-0.019*** (0.005)	0.042*** (0.009)	-0.019*** (0.005)	0.031*** (0.01)
Age2		-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001** (0.000)
Cash rate		0.026 (0.411)	7.687*** (0.787)	0.477 (0.437)	6.592*** (0.821)

GDP growth change		-0.990** (0.468)	-0.737 (0.763)	-1.168** (0.493)	-0.933 (0.81)
Unemployment rate change		0.137*** (0.017)	-0.144*** (0.029)	0.126*** (0.018)	-0.162*** (0.03)
Low-consumption period		0.270*** (0.021)	-0.176*** (0.035)	0.287*** (0.022)	-0.166*** (0.036)
Turnover category	[50K, 200K)	-0.133*** (0.021)	0.184*** (0.04)	-0.268*** (0.022)	0.211*** (0.041)
Turnover category	[200K, 500K)	-0.219*** (0.021)	0.565*** (0.046)	-0.273*** (0.022)	0.596*** (0.048)
Turnover category	[500K, 1M)	-0.165*** (0.023)	0.816*** (0.052)	-0.158*** (0.024)	0.802*** (0.052)
Turnover category	[1M, 2M)	-0.169*** (0.025)	1.211*** (0.062)	-0.108*** (0.026)	1.207*** (0.064)
Turnover category	[2M, 5M)	-0.08*** (0.026)	1.69*** (0.065)	0.024 (0.027)	1.658*** (0.068)
Turnover category	[5M, 10M)	-0.088** (0.04)	2.381*** (0.109)	0.014 (0.041)	2.344*** (0.114)
Turnover category	>=10M	-0.234*** (0.041)	2.972*** (0.12)	-0.116*** (0.043)	2.886*** (0.125)
Industry	A - Agriculture	-0.459*** (0.06)	0.26 (0.207)	-0.427*** (0.064)	0.095 (0.215)
Industry	B - Mining	0.019 (0.092)	0.634** (0.316)	0.22** (0.095)	0.511 (0.331)
Industry	C - Manufacturing	0.498*** (0.028)	-0.235 (0.15)	0.502*** (0.029)	-0.203 (0.154)
Industry	D - Utilities	0.237** (0.111)	0.301 (0.34)	0.238** (0.119)	0.278 (0.353)
Industry	E - Construction	0.617*** (0.02)	-0.307** (0.144)	0.63*** (0.021)	-0.296** (0.148)
Industry	F - Wholesale	0.025 (0.033)	-0.267* (0.155)	0.064* (0.034)	-0.295* (0.159)
Industry	G - Retail	0.17*** (0.03)	-0.478*** (0.149)	0.148*** (0.031)	-0.446*** (0.153)
Industry	H - Hospitality	0.72*** (0.031)	-0.411*** (0.151)	0.581*** (0.033)	-0.366** (0.156)
Industry	I - Transportation	0.336*** (0.031)	-0.338** (0.152)	0.353*** (0.033)	-0.403** (0.157)
Industry	J - Information	-0.257*** (0.07)	-0.347* (0.185)	-0.303*** (0.073)	-0.252 (0.193)
Industry	K - Finance	-0.662*** (0.046)	0.436** (0.187)	-0.614*** (0.049)	0.45** (0.194)
Industry	L - RES Agency	-0.329*** (0.039)	0.113 (0.164)	-0.273*** (0.041)	0.057 (0.169)
Industry	M - Professional	-0.409*** (0.027)	-0.204 (0.148)	-0.366*** (0.028)	-0.161 (0.152)
Industry	N - Support	0.599*** (0.032)	-0.28* (0.152)	0.565*** (0.034)	-0.251 (0.156)
Industry	P - Education	-0.3*** (0.078)	-0.238 (0.2)	-0.353*** (0.083)	-0.162 (0.211)
Industry	Q - Healthcare	-1.051*** (0.068)	-0.105 (0.196)	-1.077*** (0.072)	-0.205 (0.201)
Industry	R - Arts	-0.179** (0.079)	-0.247 (0.227)	-0.311*** (0.085)	-0.213 (0.234)

Industry	S - Others	0.128***	-0.258	0.076*	-0.25
		(0.039)	(0.158)	(0.041)	(0.161)
Intercept		-6.153***	1.788***	-6.250***	1.783***
		(0.031)	(0.152)	(0.034)	(0.159)

References

Australian Financial Security Authority (2024). Treatment of property in bankruptcy. https://www.afsa.gov.au/resource-hub/practices/practice-guidance/treatment-property-bankruptcy#35

Australian Taxation Office (2024) Unfair preference payments. https://www.ato.gov.au/tax-and-super-professionals/for-tax-professionals/your-practice/insolvency-practitioners/unfair-preference-payments#ato-Unfairpreferencepaymentsforindividuals

Beaver, W. H., Correia, M., & McNichols, M. F. (2012). Do differences in financial reporting attributes impair the predictive ability of financial ratios for bankruptcy?. Review of Accounting Studies, 17, 969-1010.

Beck, T., Demirguc-Kunt, A., & Levine, R. (2005). SMEs, growth, and poverty: Cross-country evidence. Journal of Economic Growth, 10, 199-229.

Berger, A. N., & Udell, G. F. (2004). The institutional memory hypothesis and the procyclicality of bank lending behavior. *Journal of financial intermediation*, *13*(4), 458-495.

Cull, R., Davis, L. E., Lamoreaux, N. R., & Rosenthal, J. L. (2006). Historical financing of small-and medium-size firms. Journal of Banking & Finance, 30(11), 3017-3042.

Gürtler, M., & Hibbeln, M. (2013). Improvements in loss given default forecasts for bank loans. Journal of Banking & Finance, 37(7), 2354-2366.

Hamerle, A., Liebig, T., & Scheule, H. (2006). Forecasting credit event frequency-empirical evidence for West German firms. Journal of Risk, 9(1), 75-98.

Kenney, R., La Cava, G., & Rodgers, D. (2016). Why Do Companies Fail? Reserve Bank of Australia Research Discussion Papers, retrieved from https://www.rba.gov.au/publications/rdp/2016/2016-09/full.html

Lown, C., & Morgan, D. P. (2006). The credit cycle and the business cycle: new findings using the loan officer opinion survey. *Journal of Money, Credit and Banking*, 1575-1597.

Mian, A., & Sufi, A. (2009). The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis. *The Quarterly journal of economics*, *124*(4), 1449-1496.

Productivity Commission. (2021). Small business access to finance: The evolving lending market.

Rösch, D., & Scheule, H. (2020). Deep Credit Risk. Machine Learning in Python. Kindle Direct Publishing. URL: www.deepcreditrisk.com.

Scheule, H., & Jortzik, S. (2019). Benchmarking loss given default discount rates. Journal of Risk Model Validation, 14(3).

World Bank. (2019, October 19). Small and Medium Firms (SMEs) Finance. https://www.worldbank.org/en/topic/smefinance#:~:text=SMEs%20account%20for%20the%20majority,(GDP)%20in%20emerging%20economies.