Preprint:
Wilson, N. and Kacer, M. (2026) Default, Fraud, and Fraud Classification in UK COVID-19 Loan Schemes: Evidence from One Million Guaranteed Loans. [Preprint]

Default, Fraud, and Fraud Classification in UK COVID-19 Loan Schemes: Evidence from One Million Guaranteed Loans

Nick Wilson and Marek Kacer

Credit Management Research Centre
Leeds University Business School
This version: February 2026

Abstract

We study default outcomes and fraud classification in UK COVID-19 government-guaranteed lending programmes using loan-level data of 1,006,579 limited-company facilities under the Bounce Back Loan Scheme (BBLS) and Coronavirus Business Interruption Loan Scheme (CBILS). We estimate logit models of loan outcomes, Cox proportional hazard models of time-to-default, and Accelerated Failure Time (AFT) models that quantify timing effects under four parametric survival distributions. We distinguish all-cause default, insolvency-related default, and defaults that lenders classify as fraud.

Three findings stand out. First, scheme design dominates risk: BBLS loans—issued under self-certification with minimal verification—show fraud odds approximately 160 times those of CBILS recipients, with Cox hazard ratios reaching 170. AFT models under the preferred generalised gamma specification show BBLS borrowers default 37 per cent sooner than CBILS recipients, with loan-to-sales ratio being the strongest predictor of accelerated failure, compressing survival time by 65 per cent per unit increase. Second, conventional credit risk indicators maintained predictive power despite pandemic disruption: probability of default scores, loan leverage, director experience, and board composition discriminate between outcomes across logit, Cox, and AFT specifications. Director experience and board size protect against accelerated failure, with governance effects compounding over time. Third, we document systematic under-classification of fraud at challenger banks and alternative finance/fintechs. Given default and controlling for borrower characteristics, these lenders classify fraud at only 40–50 per cent of the rate at main banks. AFT timing evidence reinforces this: challenger bank defaults cluster 30 per cent earlier than main bank defaults, consistent with rapid enforcement without fraud investigation. Propensity score analysis indicates approximately 2,500 fraud cases—representing £70 million—went undetected in challenger portfolios. We identified 1,904 cases (£76m in loans) as unambiguously fraudulent based on Companies House filings—loans to companies dissolved pre-pandemic or incorporated after scheme announcement—yet lenders rarely flagged these. Main banks screened out 28 per cent of such applicants at origination and classified 14 per cent of defaults as fraud; challenger banks screened out only 22 per cent and classified just 1.4 per cent as fraud.

These findings reveal trade-offs in crisis lending design and highlight an overlooked issue in delegated guaranteed schemes: the state relies on heterogeneous lender capabilities to detect fraud, biasing loss estimates and recovery strategies.

Keywords: COVID-19 lending; Bounce Back Loans; loan default; fraud detection; government guarantees; challenger banks; delegated monitoring

JEL Classification: G21, G28, H81, K42

1. Introduction

Large, rapid credit interventions are a standard policy response to systemic shocks. During the COVID-19 pandemic, the UK deployed government-guaranteed lending at unprecedented scale: the Bounce Back Loan Scheme (BBLS), Coronavirus Business Interruption Loan Scheme (CBILS), and Coronavirus Large Business Interruption Loan Scheme (CLBILS) together disbursed approximately £80 billion to around 1.6 million businesses between March 2020 and May 2021 (British Business Bank, 2022). BBLS accounted for approximately £47 billion, CBILS for £27 billion, and CLBILS for the remainder. Speed was paramount. BBLS in particular was designed for rapid deployment: 100 per cent government guarantees, self-certified eligibility, standardised terms, and minimal lender verification. Loans were typically approved and funded within 24–48 hours of application.

This design reflected a deliberate policy choice. The National Audit Office documented the trade-off explicitly: government prioritised payment speed over most other value-for-money considerations (NAO, 2020, p. 5). The Department for Business sought and received a Ministerial Direction on regularity, propriety, and value for money grounds before proceeding—an unusual step that underscored the recognised risks. The schemes succeeded on their primary terms: the British Business Bank's Year 3 evaluation concludes that, absent the schemes, an estimated 158,000 to 669,000 borrowers could have permanently closed by December 2022, and up to 3.5 million jobs could have been lost. Under the evaluation's core scenario, the combined schemes generated around £77 billion in additional gross value added and an overall benefit-to-cost ratio of approximately 3.8 (British Business Bank, 2025).

But speed came at a cost. Broad eligibility criteria and attenuated lender incentives for screening meant that credit was extended to financially weak firms, to borrowers who would not have qualified under normal underwriting standards, and—as has become increasingly apparent—to fraudulent applicants. As these loans mature and outcomes crystallise, understanding the determinants of default and fraud becomes essential for three purposes: managing the existing portfolio, recovering taxpayer funds, and designing future crisis interventions.

This paper studies two closely related objects: the determinants of default outcomes in COVID-19 guaranteed loans, and how lenders record and classify fraud. The second object matters because the data observed by policymakers and researchers typically capture fraud classification—what lenders label as fraud—rather than fraud incidence—what truly occurred. In delegated guarantee schemes, heterogeneity in lenders' information sets, investigative capacity, and classification incentives can therefore bias both cross-lender comparisons and aggregate fraud estimates.

We make three contributions to the literature on crisis lending, loan guarantees, and fraud. First, we quantify the risk factors associated with adverse loan outcomes in the largest emergency lending programme in UK history. Using a loan-level dataset covering over one million facilities matched to Companies House records and credit reference data, we estimate logit models of loan outcomes and Cox proportional hazard and AFT models of time-to-default. BBLS borrowers faced default odds roughly three times those of CBILS recipients, and fraud odds approximately 160 times higher. The fraud finding is particularly striking: it reflects not merely higher incidence but a qualitatively different risk profile arising from minimal verification. Loans with high leverage relative to turnover, missing credit scores, and single-director governance structures proved especially vulnerable.

Second, we employ Cox proportional hazard models and Accelerated Failure Time (AFT) models to analyse not just whether but when adverse outcomes occurred. The Cox models establish relative hazard rates; the AFT models—estimated under four parametric survival distributions with formal model selection—quantify the absolute magnitude of timing effects. This dual approach reveals that high-risk borrowers failed faster as well as more frequently. BBLS fraud hazards were around 170 times those of CBILS, and AFT estimates show BBLS borrowers defaulting 37 per cent sooner, with the loan-to-sales ratio compressing survival time by 65 per cent per unit increase. The AFT framework also reveals substantial lender heterogeneity in default timing: challenger bank borrowers default 30 per cent sooner and specialist lender borrowers 68 per cent sooner than major bank borrowers, a pattern consistent with differential forbearance capacity and lending to systematically weaker borrowers. These timing patterns have practical implications for early warning systems and recovery prioritisation.

Third, and most novel, we formalise and quantify a measurement issue that is central to policy evaluation: observed fraud is a joint outcome of underlying borrower behaviour and lender detection and classification. The raw data suggest challenger banks and fintechs experienced lower fraud than main banks despite higher overall default. We show this pattern reflects differential classification rather than differential incidence. Conditional on default and controlling for borrower characteristics, challengers classify fraud at only 40-50 per cent of the main bank rate. Propensity score analysis estimates that approximately 2,500 fraud cases at challenger banks were never classified as such—representing, at average loan values, some £70 million in undetected fraud. A direct validation comes from 1,904 loans (£76m) made to companies that were unambiguously ineligible—dissolved before the pandemic or incorporated after scheme announcement—where a simple Companies House cross-check would have identified every case. Main banks classified 14 per cent of these defaulted cases as fraud; challenger banks classified just 1.4 per cent.

This under-classification finding has immediate policy relevance. It suggests the true fraud burden is substantially under-reported, that recovery efforts may be misdirected, and that lender capacity for fraud investigation varies more than previously recognised. The institutional incentives are clear: in BBLS, lenders can reclaim losses through the government guarantee provided they complied with scheme rules (NAO, 2021), limiting the commercial return to intensive post-origination investigation and reclassification.

Our paper builds on and extends several strands of literature. The first concerns public loan guarantee schemes and their associated risks. Theoretical work, building on Stiglitz and Weiss (1981), identifies adverse selection and lender moral hazard as central concerns. Guarantees can improve credit access for financially constrained firms (Cowling and Mitchell, 2003; Riding et al., 2007), but they may also induce extension of credit to riskier borrowers and reduce screening intensity where lenders are partially insulated from losses (Gale, 1990; Lelarge et al., 2010). The COVID-19 schemes represent an extreme case: BBLS offered 100 per cent guarantees, eliminating any lender loss exposure and with it the normal incentive for careful evaluation. Empirical evidence on guarantees is mixed: some studies document positive effects on firm survival and employment (Riding et al., 2007; Cowling, 2010; Craig et al., 2008), while others find limited additionality—guaranteed loans flowing primarily to firms that would have accessed credit anyway (Riding and Haines, 2001; Oh et al., 2009). More recent evidence confirms the tension between additionality and adverse selection in scheme participation (Bertoni et al., 2023) and highlights loan size concentration in UK guarantees (Cowling et al., 2025). Our contribution is to analyse a setting in which normal selection mechanisms were deliberately suspended in the name of speed, providing a natural experiment in the consequences of attenuated screening and maximised moral hazard.

A second strand examines crisis lending and the intertemporal dynamics of firm survival. Emergency liquidity support may suppress insolvencies in the short term while increasing medium-term fragility through debt overhang and the continued operation of non-viable firms—the so-called 'zombie firm' problem (Acharya et al., 2020; Banerjee and Hofmann, 2022). This framework implies time-varying default risk as support measures are withdrawn and repayment obligations enforced. Our survival analysis—combining Cox hazard models with parametric AFT specifications—speaks directly to this dynamic, documenting not only how quickly different borrower types progressed to failure but by how much scheme design and borrower characteristics compressed or extended the time to default. The AFT framework proves particularly informative because the COVID loan timeline—with staggered repayment holidays, Pay As You Grow extensions, and discrete enforcement waves—generates a hazard shape too irregular for standard monotone or unimodal parametric forms to capture adequately.

The third relevant literature concerns fraud in lending markets. Theoretical models show that weak verification can induce strategic misrepresentation by borrowers (Garmaise and Natividad, 2016), and empirical work documents widespread misreporting in mortgage applications (Jiang et al., 2014; Mian and Sufi, 2017), especially when lenders face attenuated incentives to verify information (Keys et al., 2010). The COVID-19 schemes shared key features with these settings—self-reported turnover, limited verification, and lender insulation from losses—but at far greater scale. Our work quantifies fraud outcomes in this context and, critically, documents how measured fraud depends on lender detection and classification capacity.

Recent empirical work on emergency COVID-19 lending has begun to document default and fraud patterns. Granja et al. (2022) show that fintech lenders in the US Paycheck Protection Program reached marginal borrowers but exhibited higher default rates, while Erel and Liebersohn (2022) document that fintech participation expanded access in underserved areas. Griffin, Kruger and Mahajan (2023) examine misreporting indicators in the same programme and identify substantial volumes of loans as questionable using multiple observable signals, with fraud concentrated among lenders with weaker compliance infrastructure. These studies consider a different institutional setting—where the programme combined forgivable grants with loans—but the core mechanism is shared: when speed and scale dominate, weak verification and heterogeneous lender infrastructure shape both access and programme integrity. Our UK evidence offers a complementary lens in a pure-loan setting with a concentrated lender base and delegated responsibility for monitoring and classification.

Finally, we relate to work on delegated monitoring, lender heterogeneity, and bank business models. Diamond (1984) and others show that banks economise on screening costs when they have appropriate incentives. Recent work on fintech lending suggests these lenders use different information sets and underwriting models than traditional banks (Berg et al., 2020; Fuster et al., 2019), with implications for both access and risk. Boot and Thakor (2000) theorise how relationship intensity affects screening and monitoring, and Buchak et al. (2018) document how fintech lenders serve different market segments with different risk profiles. In emergency lending contexts, selection into lending channels may be amplified: main banks prioritised existing customers, leaving challenger and fintech providers to serve residual demand, including applicants declined elsewhere (Balyuk, 2023). Our analysis exploits this heterogeneity to identify classification differences that would be invisible in studies of homogeneous lender pools.

The National Audit Office conducted two major investigations into the Bounce Back Loan Scheme. The first, published in October 2020, documented the deliberate policy trade-off between speed and verification. A subsequent update in December 2021 was more critical, noting that counter-fraud activity was implemented too slowly to be effective and that departmental estimates suggested material fraud exposure (NAO, 2021). The same update reported that the Department estimated 11 per cent of BBLS loans (around £4.9 billion) showed indicators of fraud and that 37 per cent of loans (around £17 billion) were unlikely to be repaid; this estimate was subsequently revised as repayment data accrued. The NAO also documented capacity constraints in the enforcement ecosystem: by October 2021, the National Investigation Service (NATIS) had received over 2,100 intelligence reports but had capacity to pursue only around 50 cases per year. Below the organised-crime threshold, the Department relied heavily on lenders' own investigations, despite limited commercial incentives to undertake them where guarantee claims are honoured.

Three features of scheme design are central to understanding our findings. Guarantee structure: BBLS offered 100 per cent government guarantees, while CBILS offered 80 per cent; this difference fundamentally altered lender incentives. Verification requirements: BBLS relied on borrower self-certification of eligibility and turnover, with lenders required only to conduct basic fraud and anti-money laundering checks; CBILS required lender assessment of borrower viability (not full commercial underwriting, but meaningful evaluation). Lender ecosystem: scheme participation expanded beyond incumbent banks to include challenger banks, fintechs, and alternative finance providers; these entrants brought different customer acquisition strategies, underwriting capabilities, and—as we show—fraud investigation capacity. The heterogeneity across lender types proves analytically valuable for isolating classification effects.

The remainder of this paper proceeds as follows. Section 2 provides institutional background on scheme design. Section 3 describes the data, sample construction, and variable definitions. Section 4 presents the empirical framework, including the logit, Cox, and AFT specifications. Section 5 reports the logit models of loan outcomes. Section 6 presents the Cox proportional hazard analysis of time-to-default, and Section 6A reports the Accelerated Failure Time models that quantify the absolute magnitude of timing effects. Section 7 develops the fraud classification analysis, comparing actual to predicted fraud rates across lender types. Section 8 discusses implications for policy and portfolio management. Section 9 concludes.

2. Institutional Background and Scheme Design

The UK government launched the COVID-19 loan guarantee schemes in response to the economic shock created by pandemic-related lockdowns in March 2020. The schemes aimed to maintain credit flows to viable businesses facing temporary revenue disruption, thereby preserving employment and preventing unnecessary business failures. Two main programmes operated in parallel: the Coronavirus Business Interruption Loan Scheme (CBILS) and the Bounce Back Loan Scheme (BBLS), each with distinct design features that created different risk profiles.

CBILS launched first, on 23 March 2020, offering government guarantees of 80 per cent on loans up to £5 million. The scheme followed a relatively conventional guarantee model: lenders conducted credit assessments, retained 20 per cent exposure to losses, and were required to consider whether borrowers would have qualified for finance absent the pandemic. These features preserved some lender screening incentives, though the 80 per cent guarantee still materially attenuated them relative to normal commercial lending. In practice, CBILS became characterised by lengthy approval times—averaging several weeks—and high rejection rates. Many small businesses reported being unable to access the scheme due to insufficient credit history, lack of collateral, or failure to meet traditional lending criteria.

The perceived inadequacy of CBILS for the smallest businesses prompted the introduction of BBLS on 4 May 2020. This scheme represented a fundamental departure from normal lending practice. It offered 100 per cent government guarantees on standardised loans between £ 2,000 and £50,000 (capped at 25 per cent of turnover), required no personal guarantees, and featured fixed terms: a one-year interest-free period followed by five years at 2.5 per cent per annum. Most significantly, eligibility was self-certified. Borrowers declared their turnover and affirmed they were in financial difficulty due to COVID-19; lenders were instructed to conduct only minimal anti-fraud and anti-money laundering checks, explicitly not to undertake credit assessments, and to prioritise speed of processing. Applications were typically approved and funded within 24–48 hours.

These design differences created divergent incentive structures. In CBILS, lenders retained both downside exposure (through the 20 per cent loss share) and some reputational risk from originating loans that proved fraudulent or quickly defaulted. These incentives, while weakened relative to commercial lending, still provided some discipline to the underwriting process. In BBLS, by contrast, lenders faced zero credit risk—the government guarantee covered all losses, provided the lender had followed the scheme rules. This created what the National Audit Office termed a 'moral hazard' problem: lenders had limited commercial incentive to screen applicants carefully or to investigate fraud post-origination, because guarantee claims would be honoured regardless of classification.

The schemes' differing risk profiles became apparent as they matured. By September 2025, BBLS loans exhibited substantially higher default rates than CBILS loans—approximately three times higher for all-cause default and an order of magnitude higher for fraud-classified defaults. Whether this pattern reflects the schemes reaching different populations of borrowers (selection) or the consequences of differential verification and monitoring (moral hazard and fraud) is the central question we examine in this paper.

Both schemes closed to new applications in March 2021, having disbursed approximately £47 billion through BBLS (1.5 million loans) and £27 billion through CBILS (87,000 loans). The government now faces the task of managing this portfolio, recovering funds where possible, and understanding the true scale of losses. For our purposes, the schemes provide a natural experiment in the consequences of verification requirements: BBLS and CBILS operated simultaneously, served broadly similar (though not identical) populations, but featured dramatically different underwriting standards. This variation allows us to identify the consequences of minimal verification for both adverse outcomes and fraud classification. Unlike existing loan guarantee schemes the Covid schemes were open to a wider pool of lenders, including Challenger Banks and Fintech's, who were pursuing new clients.

3. Data and Sample

3.1 Data Sources

We draw on administrative data from the British Business Bank, which administered both COVID lending schemes. The data comprise the population of BBLS and CBILS loans to UK businesses, including loan amount, term, scheme type, lender identity, and outcome status. Outcome categories include active (performing), repaid, defaulted, and-within default-whether classified as insolvency-related or fraud-related.

We match loan records to Companies House data and the CMRC company panel database using company registration numbers. We process director-level characteristics: the number of directors, their mean years of directorship experience, gender composition, nationality, and ethnicity indicators. We observe whether the firm had previously raised equity finance from external investors-a marker of investor scrutiny and governance quality.

The CMRC panel provides pre-loan probability of default (PD) scores based on the firm's credit history and financial characteristics. Geographic identifiers enable matching to Region of the borrower, Index of Multiple Deprivation (IMD) scores at output area level and the Office for National Statistics Output Area Classification (OAC).

3.2 Sample Construction

We restrict to loans made to limited companies registered at Companies House. This restriction is necessary to obtain director-level characteristics; sole traders and partnerships lack equivalent governance data. The resulting sample comprises 1,006,579 loans.

Among these, 313,895 (31.2 per cent) experienced some form of default. Of defaulted loans, 286,034 (91.1 per cent) were classified as insolvency-related and 27,861 (8.9 per cent) as fraud-related. A distinct set of 194,847 loans (19.4 per cent of the sample) were repaid early-full repayment before scheduled maturity. BBLS loans account for 89 per cent of the sample by count though a smaller share by value given the £50,000 cap on individual BBLS facilities.

Our observed default rate of 31.2 per cent for limited companies is broadly consistent with official projections. The NAO's 2021 estimate that 37 per cent of BBLS loans would not be repaid aligns with our findings, recognising that our sample excludes sole traders who may have experienced different outcomes. Sample statistics are provided in Table 1.

[Table 1 about here: Sample Statistics]

3.3 Outcome Variables

All Default: Binary indicator equal to one if the loan experienced any default event-missed payments leading to formal default status, regardless of subsequent classification.

Insolvency Default: Default associated with formal insolvency proceedings: administration, liquidation, company voluntary arrangement, or dissolution.

Fraud Default: Default classified by the lender as fraud-related. This includes suspected fraudulent applications, material misrepresentation, and misuse of funds. Importantly, this variable captures classified fraud-cases the lender identified and designated as such-not necessarily all actual fraud. A key distinction is that the fraud measure reflects lender classification rather than adjudicated fraud. A lender can only classify a case as fraud if it has sufficient evidence and is willing to apply the label. Conceptually, this is a dependent-variable misclassification problem in a discrete-response setting (Hausman, Abrevaya and Scott-Morton, 1998).

Early Repayment: Full repayment of loan principal before scheduled maturity, indicating either improved financial position or precautionary borrowing that proved unnecessary.

3.4 Key Explanatory Variables

Bounce Back Loan: Indicator equal to one for BBLS loans, zero for CBILS. This is our primary scheme-design variable.

Loan/Sales Ratio: Loan amount divided by annual turnover. High values indicate borrowing disproportionate to business scale-a classic leverage risk indicator and potential fraud signal where applicants may have misstated eligibility turnover.

PD Score: Pre-loan probability of default from credit reference data, scaled 0-1 Higher values indicate elevated ex ante credit risk (reported as % in descriptive tables).

PD Missing: Indicator for firms without available credit scores-typically younger, smaller, or less financially transparent businesses. This captures information-poor cases that cannot be scored using standard bureau data.

Extended Repayment: Indicator for use of Pay As You Grow forbearance options: payment holidays, term extensions, or interest-only periods.

3.5 Lender Classification

We classify lenders into categories based on business model and regulatory status: main banks (the major UK high street lenders), Irish/Scottish banks, co-operatives, challenger banks, private banks, lease finance, invoice finance, fintechs, development banks, CDFIs (Community Development Finance Institutions), and specialist lenders. Main banks serve as the reference category in all models.

The COVID lending schemes substantially expanded the lender ecosystem beyond incumbent banks. Main banks, Barclays, HSBC, Lloyds, NatWest, and Santander, held existing relationships with the majority of UK SMEs and processed the bulk of applications from their established customer base. Challenger banks including Metro Bank, Starling, Tide, and others served businesses without main bank relationships or those seeking alternatives to their existing provider. Fintechs such as Funding Circle operated largely automated platforms. This institutional variation generates natural differences in customer composition, underwriting approach, and-critically for our analysis-fraud investigation capacity. Main banks possess decades of customer transaction history, established fraud investigation teams, and extensive experience with government guarantee schemes. Challengers and fintechs, by contrast, have shorter customer relationships, leaner operational structures, and less experience with the evidential requirements of fraud classification. Moreover, they were pursuing client base growth during this period.

This classification proves analytically important. Different lender types served different market segments, applied different underwriting approaches, and, as we demonstrate-differ markedly in their propensity to classify defaults as fraud.

4. Empirical Framework

We estimate outcome models that relate loan performance to loan terms, firm characteristics, governance variables, lender type, and fixed effects. For interpretability we report binary logit models for any default, insolvency default, fraud-labelled default, and full repayment. We also estimate a multinomial logit model for mutually exclusive outcomes, which provides a joint framework that reduces concerns about mechanical substitution across outcome definitions.

Because timing matters in crisis lending, we estimate Cox proportional hazard models for time-to-event outcomes. The hazard framework accommodates right-censoring and distinguishes not only whether default occurs but how quickly. To address the fraud measurement problem, we implement two complementary approaches. First, we estimate fraud-classification models conditional on default: among loans that have defaulted, we regress an indicator for fraud classification on lender-type indicators and progressively richer controls. Under the hypothesis that lender differences reflect borrower composition alone, lender-type coefficients should attenuate towards one as controls are added. Persistence (or strengthening) of lender-type gaps is consistent with differential detection/classification.

Second, we construct a propensity-score benchmark. We estimate a fraud-classification model using main-bank defaults only, then apply the coefficients to all defaulted loans to generate predicted fraud classification probabilities under main-bank practices. Comparing actual to predicted rates by lender type yields a classification gap that can be translated into an implied number of unclassified fraud cases.

Cox hazard ratios quantify the relative rate of default but provide no direct inference about the absolute timing of failure—whether a firm defaults three months or three years earlier. To complement the Cox results and enable time-based interpretation, we estimate Accelerated Failure Time (AFT) models under four parametric baseline survival distributions: Weibull, log-logistic, log-normal, and generalised gamma. Distribution selection is based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), with lower values indicating better fit. The generalised gamma nests the Weibull and log-normal distributions as special cases; we test these restrictions formally.

AFT model coefficients are reported as time ratios ($TR = \exp(\beta)$), where $TR < 1$ indicates that a covariate accelerates default (shortens survival time) and $TR > 1$ indicates delay. A TR of 0.75 means default occurs 25 per cent sooner; a TR of 1.30 means default occurs 30 per cent later. The exponentiated coefficient thus measures the proportional change in survival time associated with a unit increase in the covariate. For logged covariates (loan amount, sales), β is an elasticity: a 1% increase in x changes predicted survival time by approximately β% (more exactly: survival time is multiplied by $\exp(\beta \cdot \ln(1.01))$)

Critically, a hazard ratio greater than one in the Cox model corresponds directionally to a time ratio less than one in the AFT framework—both indicate elevated default risk—but the AFT coefficient provides the magnitude of the timing effect. The generalised gamma distribution with $\kappa < 0$ has no finite mean; all summary statistics therefore use median survival time as the central tendency measure.

Across models we report robust standard errors. Industry, region, and output area fixed effects are included in extended specifications.

5. Determinants of Loan Outcomes

Table 2 reports logit models of four loan outcomes: any default, insolvency default, fraud default, and early repayment. The models include over one million observations and achieve AUC statistics ranging from 0.740 to 0.824, with the fraud model showing the strongest discriminatory power.

[Table 2 about here: Logit Models of Loan Outcomes]

The Bounce Back Loan indicator dominates the analysis. BBLS borrowers show dramatically elevated default risk compared to CBILS recipients across all specifications. The coefficient of 1.105 for all default implies odds approximately three times higher ($\exp(1.105) \approx 3.0$). For insolvency specifically, the odds ratio is 2.2. Most striking is the fraud coefficient of 5.077, implying that BBLS borrowers were approximately 160 times more likely to experience fraud-related default than CBILS borrowers ($\exp(5.077) \approx 160$).

This extraordinary magnitude reflects the design differences between schemes. BBLS offered 100 per cent government guarantees with minimal lender due diligence and self-certified eligibility. These features, whilst enabling rapid disbursement, created substantial moral hazard and fraud vulnerability. The finding validates ex ante concerns about the trade-off between speed and screening-the very trade-off that prompted the Ministerial Direction before scheme launch.

BBLS borrowers were also substantially less likely to repay early ($\beta = -0.873$), consistent with weaker underlying financial capacity or greater need for the borrowed funds.

The loan-to-sales ratio is strongly predictive of adverse outcomes across all default categories, with coefficients of 1.719 (all default), 1.530 (insolvency), and 2.385 (fraud). Firms borrowing disproportionately relative to their turnover faced elevated default risk-a classic indicator of over-leverage. The particularly large fraud coefficient suggests that applications for loan amounts grossly disproportionate to business scale were often fraudulent in nature, consistent with misstatement of eligibility turnover or applications motivated by extraction rather than business continuity.

Log loan amount exhibits a striking divergence across outcomes. Larger loans are associated with higher default probability ($\beta = 0.250$), but the effect is dramatically amplified for fraud ($\beta = 1.325$). This suggests fraudulent actors deliberately sought larger loan amounts to maximise illicit gains. Conversely, larger loans are significantly less likely to be repaid early (β = −0.201), reflecting the greater financial burden of accelerated repayment.

Extended repayment period strongly predicts default (β = 1.232 for all default, 1.261 for insolvency) but shows a small negative association with fraud (β = −0.048). Firms requesting payment holidays or term extensions were signalling financial distress, making this a powerful indicator of viability concerns. The negative fraud coefficient suggests fraudulent borrowers were less likely to engage with legitimate forbearance mechanisms - consistent with an intention to abscond rather than manage genuine business difficulties.

PD scores perform as expected-higher ex ante credit risk strongly predicts all adverse outcomes (β = 3.397 for all default, 3.334 for insolvency, 2.656 for fraud). This validates that pre-existing credit scores retained predictive power despite the unprecedented pandemic context. Notably, the fraud coefficient is somewhat lower than for insolvency, suggesting fraud was not confined to traditionally high-risk borrowers, a finding consistent with opportunistic behaviour by otherwise creditworthy individuals.

The PD Missing indicator is a powerful risk signal. Firms without available credit scores-typically younger, smaller, or less financially transparent-show substantially elevated default risk (β = 1.101) and particularly elevated fraud risk (β = 1.780). Information opacity correlates with both genuine distress and fraudulent intent, suggesting that information-poor applicants warrant particular scrutiny in emergency lending.

Director characteristics provide meaningful signals. Board size is strongly protective against default (β = −0.220) and fraud (β = −0.267). Larger boards imply more established businesses with greater oversight capacity and collective accountability. Single-director companies-common among BBLS recipients-lack these governance safeguards.

Director experience shows consistent protective effects. Each additional year of mean director experience reduces default probability (β = −0.036) and fraud probability (β = −0.044). Experienced directors bring superior business acumen, established networks, and reputational capital that discourages misconduct.

Prior equity finance provides a notable protective effect, particularly against fraud (β = −0.552). Firms that had previously raised equity were less likely to default overall and substantially less likely to experience fraud. Equity-backed firms have passed investor due diligence, possess more sophisticated governance, and have access to alternative funding sources.

6. Survival Analysis: Time to Default

The logit models examine whether loans defaulted; Cox proportional hazard models examine when. This duration perspective offers advantages for understanding COVID loan performance: it appropriately handles censoring, captures risk dynamics over the loan lifecycle, and yields hazard ratios with natural interpretations as multiplicative effects on instantaneous failure risk.

Table 3 reports hazard ratios from Cox models of all-cause default, insolvency default, and fraud default. The estimation sample comprises 1,006,579 loans; failure events require an observed default date, yielding 271,957 all-cause failures, 251,024 insolvency failures, and 20,933 fraud-labelled failures. The higher headline default count in Table 1 reflects additional loans that enter default status but lack a reliable event date for duration modelling and are treated as right-censored.

[Table 3 about here: Cox Proportional Hazard Models]

BBLS status dominates the survival analysis as it did the logit models. BBLS borrowers experienced all-cause default hazards 2.32 times those of CBILS borrowers-they defaulted more than twice as fast. For insolvency specifically, the hazard ratio was 1.86.

Most striking is the fraud hazard ratio of 170.3. BBLS loans experienced fraud-related default at 170 times the rate of CBILS loans. This extraordinary magnitude exceeds even the odds ratio of 160 from the logit model because the Cox model captures not just whether fraud occurred but how quickly. BBLS fraud was both more common and detected earlier in the loan lifecycle, compounding the effect in survival terms. The magnitude highlights not only a higher recorded fraud incidence but also much earlier emergence/recognition of fraud-labelled failures in the BBLS portfolio.

The loan-to-sales ratio is the most powerful continuous predictor. A unit increase multiplies the all-default hazard by 6.55, the insolvency hazard by 5.72, and the fraud hazard by 37.2. Highly leveraged borrowers failed dramatically faster across all outcome types. The fraud effect suggests that firms borrowing far in excess of turnover were often fraudulent operations with no intention of repayment.

Log loan amount shows substantial effects: each unit increase raises the fraud hazard by 267 per cent (HR = 3.673). Larger loans not only default more often but default faster, consistent with larger targets for rapid exploitation.

Extended repayment presents a nuanced pattern. For genuine default, extension users show hazard ratios of 2.12 (all default) and 2.28 (insolvency)-they failed roughly twice as fast. This reflects endogeneity: firms requesting extensions were already experiencing distress.

For fraud, however, extension is protective against rapid failure (HR = 0.911). Fraudulent borrowers did not engage with legitimate forbearance mechanisms-they either absconded quickly or were identified before needing to request extensions.

Governance variables matter for timing: more experienced and larger boards are associated with slower progression to default and insolvency. These timing results complement the logit evidence and motivate early-warning monitoring that focuses on high-leverage borrowing, information-poor applicants, and weak governance profiles.

6A. Accelerated Failure Time Models: Quantifying the Timing of Default

The Cox hazard ratios reported in Section 6 establish that BBLS borrowers, highly leveraged firms, and borrowers at non-major banks experience default at elevated rates. The AFT framework enables a complementary question: by how much do these characteristics change the time to default?

Table 3A reports time ratios from four parametric AFT specifications estimated on the all-default outcome. Model selection via AIC and BIC decisively favours the generalised gamma distribution (AIC = 1,223,878; BIC = 1,224,681), which outperforms the Weibull (ΔAIC = 35,387), log-logistic (ΔAIC = 21,538), and log-normal (ΔAIC = 5,959) alternatives. Formal hypothesis tests confirm that the generalised gamma shape parameter κ is statistically distinguishable from both the Weibull special case ($\kappa = 1$; $\chi^2(1) = 5{,}067$, $p < 0.001$) and the log-normal limit ($\kappa = 0$; $\chi^2(1) = 1{,}066$, $p < 0.001$). The estimated $\kappa = -0.847$ indicates a distribution with heavy right tail and early peak hazard, consistent with the substantive complexity of the COVID loan timeline: repayment holidays ending at staggered intervals, Pay As You Grow extensions restructuring obligations, and discrete enforcement waves create a hazard shape more irregular than the monotone (Weibull) or unimodal (log-logistic) alternatives can capture.

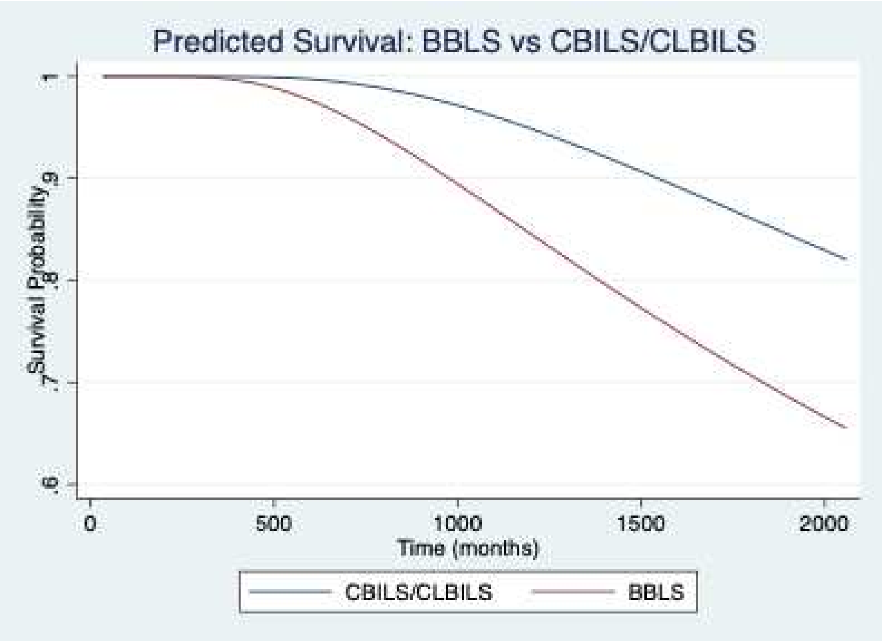

Figure 1: AFT Predicted Survival: BBLS vs CBILS/CLBILS (Generalised Gamma). Note: X-axis shows time in days from origination.

[Table 3A about here: AFT Distribution Comparison]

6A.1 Loan Scheme and Leverage

Under the preferred generalised gamma specification, BBLS borrowers are predicted to default 36.8 per cent sooner than CBILS/CLBILS recipients (TR = 0.632, $p < 0.001$). At the median survival time among defaulted loans, this translates to BBLS failures occurring approximately 31 months (2.6 years) earlier (Table 3B). The time ratio itself is interpretable and robust: conditional on defaulting, BBLS loans fail at 63 per cent of the time taken by CBILS loans. The result is qualitatively unchanged across all four distributional specifications, with BBLS time ratios ranging from 0.632 (generalised gamma) to 0.658 (Weibull).

The loan-to-sales ratio emerges as the single most powerful predictor of accelerated default. A unit increase in leverage shortens time to default by 65.2 per cent (TR = 0.348, $p < 0.001$). Firms borrowing amounts disproportionate to their turnover experience massively compressed survival times, consistent with the leverage channel documented in the logit and Cox specifications. By contrast, loan amount and firm size (both logged) have elasticities near zero once leverage is controlled (−0.07 per cent and −0.01 per cent respectively). This finding reinforces the central role of ex ante cash flow stress as a predictor of failure timing: firms that drew maximum available amounts relative to their operating scale had no margin for error.

6A.2 Pay As You Grow and the Adverse Selection Problem

Firms that utilised Pay As You Grow extensions or repayment holidays exhibit 22.1 per cent faster time to default (TR = 0.779, $p < 0.001$) relative to standard repayment borrowers. This result is economically counterintuitive—extensions should delay failure by providing breathing room—but statistically robust and substantively coherent. The extensions were not randomly assigned: they were taken up by firms already experiencing financial distress. The AFT estimate captures adverse selection, not treatment effect. Among firms that ultimately defaulted, those requiring forbearance were weaker ex ante, and conditional on failure they succumbed sooner despite the extensions, because they started from a position of greater underlying weakness. The finding suggests that Pay As You Grow succeeded in its immediate objective—preventing precipitous collapse among stressed borrowers—but could not reverse the underlying trajectory for marginal firms that were already past the point of viability.

6A.3 Lender Heterogeneity in Default Timing

The AFT results provide a temporal dimension to the lender heterogeneity findings from Section 6. Relative to major banks, challenger bank loans default 30.5 per cent sooner (TR = 0.695, $p < 0.001$), specialist lenders 67.6 per cent sooner (TR = 0.324, $p < 0.001$), and fintech lenders 37.3 per cent sooner (TR = 0.627, $p < 0.001$). At the median, challenger bank defaults occur approximately 33 months (2.8 years) earlier than major bank defaults; specialist lender defaults occur 57 months (4.8 years) earlier (Table 3B).

This timing differential is consistent with two non-exclusive mechanisms: (1) forbearance capacity differences, where major banks with larger balance sheets and specialised workout teams carry troubled loans 6–12 months before enforcing, while challengers enforce after 1–2 months; and (2) borrower quality differences, where challengers attracted systematically riskier borrowers under the government guarantee. The magnitude of timing acceleration (30–68 per cent across non-major lenders) suggests mechanism (1) plays a material role: if the difference were purely borrower selection, we would expect proportionally more defaults across the entire time horizon, not concentrated in early periods.

The timing evidence strengthens the under-classification hypothesis. If non-major banks systematically under-report fraud by classifying it as insolvency instead—either because earlier enforcement leaves less time for investigation or because they lack sophisticated fraud detection infrastructure—the temporal signature should reflect the fact that fraudulent loans default rapidly (fraudsters abscond or are detected quickly), whereas genuine insolvencies unfold more slowly. The observed clustering of early defaults at challenger banks and specialist lenders provides precisely this pattern. Combined with the lower reported fraud rates at these lenders documented in Table 5, the timing evidence suggests that a portion of what these lenders record as insolvency may in fact be undetected or unclassified fraud.

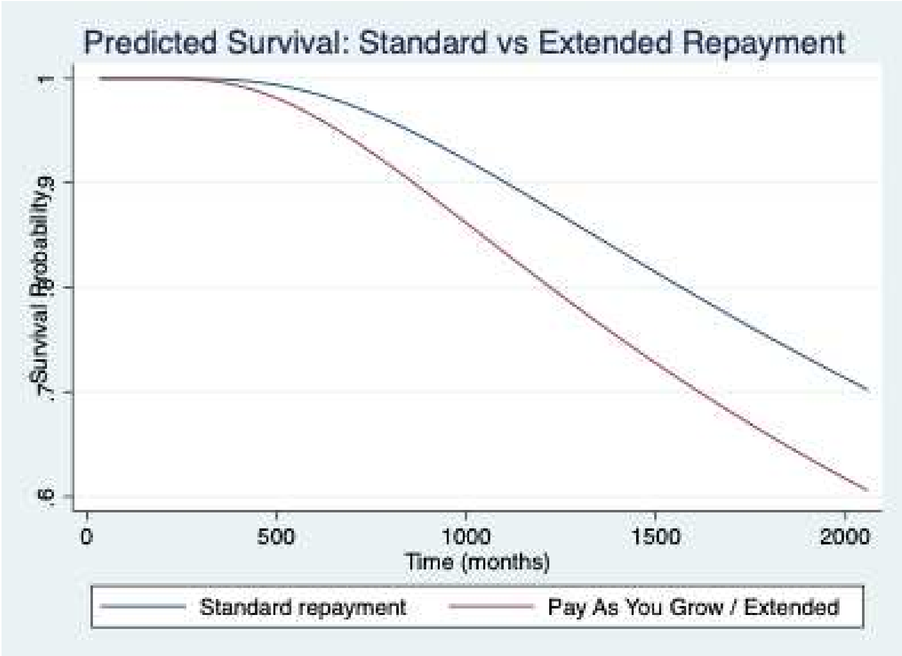

Figure 2: AFT Predicted Survival: Standard vs Extended Repayment (Generalised Gamma). Note: X-axis shows time in days from origination.

[Table 3B about here: Predicted Survival Times by Loan Type and Lender]

6A.4 Director and Governance Characteristics

Director experience provides meaningful protection against accelerated failure. Each additional year of mean director experience extends survival time by 2.2 per cent (TR = 1.022, $p < 0.001$). A board with ten years' average experience versus zero years exhibits 24.3 per cent longer predicted time to default (calculated as $1.022^{10} = 1.243$, not 5 × 2.2 per cent, due to compounding). Board size shows a similarly protective effect: each additional director extends survival by 9.6 per cent (TR = 1.096, $p < 0.001$). A three-director firm versus one-director firm exhibits 20.1 per cent longer survival ($1.096^2 = 1.201$). These governance findings are qualitatively consistent across all four distributional specifications and reinforce the signals identified in the logit and Cox models: experienced, larger boards provide measurable risk mitigation even in crisis lending contexts where traditional underwriting was largely suspended.

6A.5 Robustness and Interpretation

The qualitative findings reported above are invariant to distributional choice. BBLS time ratios range from 0.632 to 0.658 across the four specifications; leverage ranges from 0.348 to 0.358; challenger bank effects range from 0.695 to 0.730. The direction and statistical significance of all key covariates remain unchanged (Table 3A). The generalised gamma is preferred on statistical grounds and provides the most conservative estimates for several key effects (notably, the Extended Period coefficient is least negative under the generalised gamma), strengthening confidence that the reported magnitudes are not driven by distributional mis-specification.

Three interpretive caveats apply. First, the predicted absolute survival times reported in Table 3B are measured in days from origination. The survival time variable in the data is measured in days (maximum observed: 2,060 days ≈ 5.6 years), consistent with loans originated 2020–2021 observed through November 2025. All reported figures in Table 3B use days as the unit, with conversions to months and years provided. The time ratios remain interpretable and valid: 'BBLS loans default 37 per cent sooner than CBILS loans' is a correct statement based on proportional time changes.

Second, the AFT framework assumes proportional effects on the time scale—covariates stretch or compress the survival distribution uniformly. If covariate effects vary over time (e.g., PD scores matter more in early months), the reported time ratios represent averages over the follow-up period. The proportional hazards diagnostics conducted on the Cox models flag this concern for certain covariates; stratified Cox models and time-varying coefficient specifications provide robustness checks (available upon request).

Third, several covariates are endogenous. Extended_Period (Pay As You Grow) selection is driven by financial distress, introducing adverse selection bias. The AFT time ratios measure associations, not causal effects. The direction of bias is predictable: the Extended_Period time ratio is downward biased (overstates acceleration) because selection on distress confounds the treatment effect. Similarly, lender type effects may reflect both borrower sorting and lender practices, which the AFT timing evidence alone cannot fully disentangle.

7. Differential Fraud Classification Across Lenders

The baseline models treat fraud-labelled default as an outcome. However, fraud-labelled default is not an objective event in the same way as missed payments or insolvency filings: it is a classification made by the lender. A lender must both detect suspicious behaviour and be willing/able to substantiate a fraud label in its reporting and claims process. This creates a measurement problem when comparing fraud rates across lender types.

The empirical puzzle is that challenger banks and fintechs exhibit higher overall default rates but lower fraud classification rates than main banks. Among defaulted loans, main banks classify 9.4 per cent as fraud; challengers classify 5.7 per cent; fintechs 3.2 per cent.

Two broad explanations exist:

Hypothesis A (Differential Incidence): Challengers genuinely experienced less fraud. Perhaps fraudsters targeted traditional banks, or challenger verification processes proved superior at screening out fraudulent applications.

Hypothesis B (Differential Classification): Challengers experienced similar or higher fraud incidence but under-classified it. Weaker investigation capacity, shorter customer histories, or different institutional incentives led them to record fraud as ordinary default.

The distinction matters. Under Hypothesis A, challengers performed well on fraud prevention. Under Hypothesis B, the true fraud burden is substantially under-reported and recovery efforts have been misdirected.

We observe fraud classification, not true fraud incidence. A lender's classified fraud rate conflates two things: the underlying fraud rate among its borrowers and its capacity to detect and classify that fraud. We cannot directly observe the first, but we can attempt to isolate the second by comparing classification rates conditional on observable characteristics.

The logic is straightforward. If challengers serve borrowers with different observable characteristics that predict lower fraud, the raw gap in classification rates may simply reflect borrower composition. But if the gap persists after controlling for observables, something else is driving it: either unobserved borrower differences or differences in lender classification practices.

We begin by estimating multinomial logit models of loan outcomes, where each facility can result in one of three states: no default (the base outcome), default not classified as fraud, or default classified as fraud. This specification allows us to examine whether risk factors have differential associations with fraud versus non-fraud default, a distinction that proves empirically important. The model takes the form:

$$P(Y_i = j \mid X_i) = exp(X_i\beta_j) / \Sigma_k\ exp(X_i\beta_k)$$

where Y_i denotes the outcome for loan i, j indexes outcomes (no default, non-fraud default, fraud default), and X_i contains loan and borrower characteristics. The vector β captures the association between characteristics and log-odds of each outcome relative to the base case of no default. We estimate this model on the full sample of 1,006,579 loans and report relative risk ratios, which have the interpretation of proportional effects on outcome probabilities.
The vector of controls includes loan characteristics (amount, loan-to-sales ratio, term, scheme type), firm characteristics (pre-pandemic probability of default score, sales, leverage), governance variables (board size, director experience, gender and nationality composition), and lender fixed effects. Including lender fixed effects is essential because lenders may have selected into different segments of the applicant pool, operated different approval thresholds, or served different geographic markets. These fixed effects absorb any time-invariant differences in lender practices or portfolios, allowing us to identify within-lender variation in outcomes as a function of borrower characteristics.

Table 4 presents the results from the multinomial logit specification. Column 1 reports relative risk ratios for non-fraud default; column 2 reports relative risk ratios for fraud default, both relative to the base outcome of no default. The estimates reveal several clear patterns. Consider first the role of scheme type. BBLS borrowers face 2.6 times the odds of non-fraud default and 231 times the odds of fraud default compared to observationally identical CBILS recipients. These are enormous effects—far larger than those associated with any individual borrower characteristic—and they underscore that programme design dominated conventional risk factors in determining outcomes. The fraud effect is particularly striking: it indicates that minimal verification, even holding constant all measured borrower characteristics, multiplied fraud odds by more than two orders of magnitude. This finding provides strong evidence that fraud vulnerability arose from scheme design rather than merely from selection of risky borrowers.

Loan characteristics also matter, though their effects differ by outcome type. Larger loans exhibit higher default odds of both types: a one-log-point increase in loan amount raises non-fraud default odds by 25 per cent and fraud default odds by nearly 300 per cent. The stronger fraud association likely reflects that larger frauds were both more profitable to perpetrate and potentially easier to detect (triggering greater lender scrutiny). The loan-to-sales ratio—a measure of leverage—has particularly strong effects: moving from the 25th to the 75th percentile of this distribution (roughly a doubling) increases non-fraud default odds by 95 per cent and fraud default odds by 175 per cent. High leverage signals either aggressive borrowing or potentially misrepresented sales, both of which elevate risk.

Turning to firm characteristics, we find that pre-pandemic credit quality strongly predicts outcomes despite the intervening shock. Each percentage point increase in probability of default scores is associated with a 49-fold increase in both non-fraud and fraud default odds. Firms with missing credit scores—typically very young companies or those operating outside the formal credit system—face triple the non-fraud default odds and eight times the fraud default odds relative to scored firms with equivalent observed characteristics. These patterns confirm that conventional credit indicators retained their informativeness even in the pandemic context and suggest that schemes requiring minimal verification were particularly vulnerable to adverse selection.

Corporate governance characteristics show interesting differential effects. Firms with more experienced directors (measured by years since first directorship) face lower odds of both outcome types, but the gradient is steeper for fraud: each additional year of director experience reduces fraud odds by 5.4 per cent versus 3.9 per cent for non-fraud default. Larger boards also associate with lower risk across both outcomes, which could reflect either superior governance or simply that single-director companies are systematically different from those with multiple decision-makers. The board composition variables—share of female directors, foreign nationals—show modest effects that vary by outcome type, though we caution against strong interpretation of these demographics-based patterns without additional evidence on mechanisms.

The bottom panel of Table 4 reports lender fixed effects (relative to main banks). These coefficients capture differences in default rates across lender types after controlling for borrower characteristics. Several patterns emerge. Challenger banks show 87 per cent higher non-fraud default odds than main banks but 10 per cent lower fraud default odds—a combination that appears anomalous given the typical positive correlation between default and fraud. Invoice finance lenders, lease finance firms, and specialist lenders all show elevated odds of both outcome types, likely reflecting their focus on higher-risk market segments. The Development Bank category shows remarkably low default odds, consistent with that institution's mission-driven focus and potentially more intensive monitoring. Crucially, the lender fixed effects for fraud are exactly what motivate our classification. If challengers truly served lower-fraud populations, we should be able to explain this through borrower characteristics, which we cannot.

Next, we implement two complementary approaches: conditional logit models that progressively add controls, and propensity score analysis that uses main bank classification behaviour as a benchmark.

Table 5 presents logit models of fraud classification conditional on default-the sample is restricted to defaulted loans, and the dependent variable indicates whether that default was classified as fraud. We progressively add controls: loan characteristics, firm characteristics, director characteristics, and fixed effects.

[Table 5 about here: Fraud Classification Conditional on Default]

The challenger bank coefficient evolves as follows: 0.584 with no controls, falling to 0.487 with loan characteristics, 0.391 with firm characteristics, 0.455 with director characteristics, and 0.427 with full fixed effects. The coefficient becomes more negative as controls are added, not less.

This is the opposite of what Hypothesis A predicts. If the raw gap were explained by borrower characteristics, controls should push the coefficient toward 1.0. Instead, adding controls reveals a larger gap. In the full specification, challengers classify fraud at only 43 per cent of the main bank rate for borrowers with identical observable characteristics. Fintechs show a similar pattern (OR = 0.524 in the full model). Conditional on default and controlling for all observables, defaults at these lenders are approximately 50-60 per cent less likely to be classified as fraud than identical defaults at main banks.

The second approach uses main bank classification behaviour as a benchmark. We estimate a logit model of fraud classification on main bank defaults only, using borrower and loan characteristics as predictors. We then apply the estimated coefficients to generate predicted fraud probabilities for all defaulted loans, regardless of actual lender.

Table 6 compares actual to predicted fraud rates. Main bank defaults show an actual fraud rate of 9.38 per cent versus a predicted rate of 7.97 per cent-a modest positive gap indicating main banks classify slightly more fraud than borrower characteristics alone would predict.

[Table 6 about here: Propensity Score Analysis]

Challenger banks show an actual rate of 5.70 per cent versus a predicted rate of 11.85 per cent-a gap of −6.15 percentage points, or 52 per cent below predicted. Among 41,223 challenger defaults, this implies approximately 2,535 fraud cases went unclassified. At average BBLS loan values, this represents approximately £70 million in undetected fraud.

Fintechs show an actual rate of 3.24 per cent versus predicted 5.11 per cent-37 per cent below predicted.

Table 7 compares observable characteristics across default types. If under-classification is occurring, challenger non-fraud defaults should resemble main bank fraud-cases with fraud characteristics but lacking the fraud label.

[Table 7 about here: (Panels 7.1-7.2]

The loan-to-sales ratio-a key fraud predictor-is virtually identical: 0.206 for main bank fraud versus 0.207 for challenger non-fraud ($p = 0.241$). Credit scores tell a striking story: challenger non-fraud defaults have higher PD scores. These were higher-risk borrowers, making lower fraud classification harder to explain by borrower quality.

Unclassified but unambiguously fraud

A more direct test identifies cases that are unambiguously fraudulent on observable criteria yet were not classified as fraud. Matching loan records to Companies House status histories reveals 1,122 loans (£51m) made to companies that had been dissolved before the pandemic—firms that had legally ceased to exist, could not have been trading, and could not legitimately have applied for support. These are cases of identity fraud, likely organised in many instances. A further 782 loans (£25m) were made to companies incorporated after scheme announcement. These borrowers are ineligible by construction and clearly falsified required application information on company turnover in prior years. Most of these cases acquired the maximum bounce back loan (£50k).

These 1,904 cases, £76m in loans, provide a direct test of lender screening and classification capacity. At the application stage, main banks identified and rejected 28 per cent of these unambiguously fraudulent applicants; challenger banks screened out 22 per cent. The remainder were approved and disbursed. None of these loans were repaid. All entered default—yet the lenders' post-default classification reveals a striking disparity: main banks identified and classified 14 per cent of these defaulted cases as fraud, whereas challenger banks classified just 1.4 per cent—a tenfold difference in detection rates for cases where fraud is objectively verifiable from public records. The total value of loans to unambiguous fraud cases at challenger banks amounts to £3.5m. A simple cross-check against Companies House dissolution and incorporation dates would have flagged every one of these cases at negligible processing cost. They represent a lower bound on identifiable fraud and validate under-classification at the system level.

Three mechanisms likely contribute to the pattern, and the NAO's findings on lender incentives provide institutional context.

Information asymmetry. Main banks with decades-long customer relationships can identify fraud by comparing current behaviour to historical patterns. Challengers, with relationships measured in months, lack this baseline.

Investigation capacity. Fraud classification requires evidence-investigators, legal review, documentation. Main banks have established fraud teams with economies of scale. Challengers may lack dedicated investigation resources and incentive to check.

Incentive structures. As the NAO documented, *'lenders can reclaim fraudulent Bounce Back Loans through the government guarantee, provided they followed the scheme's rules'*. If fraud claims require additional documentation or invite regulatory scrutiny, lenders may apply higher evidentiary thresholds. Challengers, less experienced with guarantee schemes, may be more cautious about classifications they cannot fully evidence.

The pattern of high default combined with low fraud classification is precisely what under-detection would produce.

A natural concern is that our findings reflect selection rather than classification differences: perhaps fraudsters systematically chose challenger banks, generating higher true fraud incidence that-even with equal classification rates-would produce lower observed fraud rates if challengers also attracted more non-fraudulent defaults. We address this concern by examining the institutional mechanisms governing borrower-lender matching during the schemes. Related evidence from the PPP suggests that FinTech lenders expanded access in underserved areas but may entail monitoring trade-offs (Erel and Liebersohn, 2022).

Existing Customer Relationships. Main banks overwhelmingly lent to existing customers. The schemes permitted-and main banks strongly preferred-lending to businesses with established current account relationships. This generated a natural sorting mechanism: firms with main bank relationships applied to their existing provider; firms without such relationships sought alternative channels. This selection has ambiguous implications for fraud. Existing relationships provide informational advantages that may deter fraudulent applications (applicants know the bank holds transaction history). But they also mean main banks served an older, more established customer base with observable track records-precisely the population for whom fraud detection is easier.

Rejected Applicants and Second-Choice Lenders. Some challenger applicants were likely rejected by main banks before applying elsewhere. Under CBILS, which required viability assessment, main banks declined applications they judged non-viable. These rejected applicants may have subsequently obtained BBLS facilities from challengers operating with minimal verification. This mechanism would generate adverse selection into challenger portfolios-not of fraudsters specifically, but of marginal borrowers whose viability was questioned. However, this selection should manifest in higher default rates (which we observe) without necessarily implying higher fraud rates. The mechanism does not explain why, conditional on default, challengers classify less fraud.

Customer Acquisition Incentives. Challengers faced different strategic incentives from main banks. For established banks, COVID lending served existing relationships; the strategic objective was retention rather than acquisition. For challengers, the schemes offered an unprecedented customer acquisition opportunity. With loans 100 per cent guaranteed, challengers could extend credit to new customers at zero credit risk, establishing relationships that might prove valuable for subsequent (unguaranteed) products. This 'acquire now, monetise later' logic implies challengers had incentives to maximise lending volumes rather than screen carefully-consistent with both higher default rates and higher true fraud incidence.

Critically, however, this mechanism cuts against the hypothesis that challengers experienced lower true fraud. If anything, customer acquisition incentives imply challengers accepted higher-risk applicants including potential fraudsters. The fact that challengers nonetheless show lower classified fraud strengthens rather than weakens the case for under-classification.

Automation and Organised Fraud. Challenger and fintech lending systems were typically more automated than main bank processes, with online applications, algorithmic decisions, and minimal human intervention. This automation has countervailing implications for fraud.

On one hand, automated systems may be better at implementing standardised fraud checks consistently-every application receives the same algorithmic screening. On the other hand, automation creates vulnerabilities to organised fraud. Sophisticated fraud operations-often involving multiple applications across shell companies, synthetic identities, or coordinated misrepresentation-can exploit automated systems more easily than processes involving human review. A fraudster submitting dozens of applications can do so efficiently through digital channels; the same operation would be impractical at a branch-based lender requiring face-to-face interaction.

Evidence from the US PPP supports this concern. Griffin et al. (2023) document that FinTech lenders exhibited substantially higher fraud indicators, with suspicious loans clustering in ways suggesting coordinated applications. If similar patterns characterised UK challenger lending, we would expect higher true fraud incidence at challengers-making lower classified fraud rates even more puzzling under the differential-incidence hypothesis.

Implications for Interpretation

The selection mechanisms we identify - existing customer lending at main banks, rejected applicants flowing to challengers, customer acquisition incentives, and automation vulnerabilities - collectively suggest that challenger portfolios likely contained equal or higher true fraud incidence than main bank portfolios. Yet challengers classify substantially less fraud conditional on default. The selection evidence therefore strengthens our interpretation: the classification gap reflects differential detection capacity rather than differential underlying fraud.

8. Discussion and Policy Implications

8.1 Alignment with Official Evaluations

Our findings align closely with - and extend - the conclusions of official government evaluations. The British Business Bank's Year 3 evaluation concludes that the schemes generated substantial benefits: absent the schemes, an estimated 158,000 to 669,000 borrowers could have permanently closed by December 2022, and up to 3.5 million jobs could have been lost. Under the evaluation's core scenario, combined benefits were around £77 billion in additional gross value added, implying an overall benefit-to-cost ratio of approximately 3.8 (British Business Bank, 2025). These benefits provide important context for our default and fraud findings - the schemes achieved their primary objective of rapid liquidity support even as they incurred elevated losses.

Our default estimates are broadly consistent with official projections. The NAO's 2021 estimate that 37 per cent of BBLS loans would not be repaid aligns with our observed default rate of 31.2 per cent for limited companies. The official fraud estimate of 11 per cent (subsequently revised to 7.5 per cent) is substantially higher than our classified fraud rate of 2.8 per cent for BBLS borrowers, consistent with our central finding that a substantial portion of fraud remains unclassified.

Most pertinently, the NAO's observation that lenders lack commercial incentives to investigate fraud-given that guarantee claims are honoured regardless of fraud classification-provides an institutional explanation for our under-classification findings. The NAO documented that fraud investigation capacity was concentrated in specialist government agencies with limited resources, whilst lenders were expected to use 'business-as-usual' recovery processes that may be poorly suited to fraud identification.

Our findings complement and contrast with Griffin et al.'s (2023) PPP analysis in important ways. Both studies document elevated fraud risk in emergency lending programmes with attenuated verification. Both identify systematic variation across lender types. However, the patterns differ: in the US, FinTech lenders showed higher fraud indicators; in the UK, we find challengers and fintechs show lower classified fraud despite higher default.

This difference may reflect distinct phenomena: Griffin et al. measure fraud indicators in loan applications; we measure fraud classification conditional on default. Our interpretation-that challengers under-classify rather than under-experience fraud-is consistent with their finding that digital lenders have weaker fraud controls, but manifests differently given the UK guarantee structure where fraud classification affects recovery efforts rather than loan approval.

The BBLS coefficient for fraud-odds ratio of 160, hazard ratio of 170-represents a stark quantification of the speed-versus-verification trade-off. The coefficient is not merely large; it is orders of magnitude larger than any other predictor in the model. While speed was prioritised during the crisis, the fraud exposure was substantial.

This does not mean BBLS was a policy failure. The schemes achieved their primary objective of rapid liquidity support at scale. The benefit-to-cost ratio of 3.8 indicates substantial net value despite the elevated losses. But the fraud costs-estimated at £1.9 billion (Department for Business and Trade, 2024) in officially flagged cases, potentially higher given our under-classification findings-represent a material offset to these benefits.

Future emergency lending should incorporate proportionate verification mechanisms even under time pressure. Simple cross-checks against Companies House dissolution dates and incorporation dates would have flagged the 1,904 unambiguously fraudulent loans we identify at negligible processing cost. The fact that main banks screened out 28 per cent of these cases at application while challengers screened out only 22 per cent—and that main banks subsequently classified 14 per cent of the defaulted remainder as fraud versus just 1.4 per cent at challengers—demonstrates that even basic verification procedures varied substantially across the lender ecosystem.

PD scores retained predictive power despite pandemic disruption. This finding supports continued use of credit reference data in crisis lending contexts, contrary to suggestions that pandemic conditions rendered historical credit information obsolete. The strength of the PD Missing indicator suggests particular caution is warranted with information-poor applicants.

The expansion of COVID lending to challenger banks and fintechs enabled rapid scale-up and served borrowers underserved by traditional banks. But our findings reveal a cost: differential fraud investigation capacity meant that substantial fraud went undetected and unrecovered in challenger portfolios.

This has immediate implications for recovery efforts. Resources directed at main bank portfolios-where fraud is already well-classified-may yield lower marginal returns than investigation of challenger defaults where fraud indicators are present but classification is absent. The under-classification finding has direct operational implications: recovery strategies should account for systematic variation in classification rates across lender types.

8.2 Limitations

Several limitations follow from the measurement environment. We observe fraud classification rather than adjudication, and classification lags may differ across lenders. Our conditional-on-default tests mitigate borrower-composition confounds but cannot eliminate unobserved heterogeneity in fraud types across lender channels. The propensity benchmark assumes main-bank classification is a reasonable reference point; if main banks systematically over-classify, under-classification estimates for challengers would be overstated.

Our analysis cannot fully rule out selection on unobservables. If challengers attracted a systematically different type of fraudster - for instance, less sophisticated opportunists who are harder to classify as fraud because their behaviour more closely resembles genuine business failure - this could generate lower classification rates without implying detection failure. However, this alternative requires a specific pattern of selection: not merely different fraudsters, but fraudsters whose fraud is observationally equivalent to legitimate distress. We find this less plausible than differential detection capacity, particularly given the institutional evidence on challenger investigation resources.

The selection mechanisms we examine - existing customer relationships, rejected applicants, acquisition incentives, and automation - suggest that if anything, challenger portfolios contained more rather than less true fraud. Existing customer lending at main banks implies an information advantage that would deter sophisticated fraudsters; automation at challengers created vulnerabilities to organised fraud. These mechanisms strengthen rather than undermine the under-classification interpretation.

9. Conclusion

This paper analyses default outcomes and fraud classification in the UK's COVID-19 guaranteed loan programmes using over one million loan observations linked to firm, credit, and director information. We deploy three complementary empirical approaches: logit models that identify the determinants of default and fraud, Cox proportional hazard models that establish relative failure rates, and Accelerated Failure Time models—estimated under four parametric survival distributions with formal model selection—that quantify the absolute magnitude of timing effects. The AFT framework proves particularly well-suited to the COVID loan context, where staggered forbearance measures and discrete enforcement waves generate complex hazard dynamics that the generalised gamma distribution captures more effectively than standard alternatives.

Three findings stand out. First, BBLS design—100 per cent guarantees, self-certification, minimal verification—created fraud exposure of extraordinary magnitude. The fraud odds ratio of 160 relative to CBILS reflects a qualitative difference in risk profile, not merely a quantitative elevation. AFT models show that BBLS borrowers default 37 per cent sooner than CBILS recipients, with the loan-to-sales ratio emerging as the single most powerful predictor of accelerated failure—compressing survival time by 65 per cent per unit increase, far exceeding the effects of loan size or firm scale. Second, conventional credit risk indicators retained predictive power despite pandemic disruption, validating their continued use in crisis contexts. The AFT results reinforce this finding: director experience and board size provide measurable and compounding protection against accelerated failure across all four distributional specifications. Third, and most novel, we document systematic under-classification of fraud at challenger banks and fintechs. These lenders classify fraud at roughly half the rate of main banks for observationally equivalent defaults, implying approximately £70 million in undetected fraud. A direct validation comes from 1,904 loans (£76m) made to companies that were unambiguously ineligible—dissolved before the pandemic or incorporated after scheme announcement—where main banks classified 14 per cent of defaulted cases as fraud versus just 1.4 per cent at challenger banks, a tenfold difference for objectively verifiable fraud. The AFT timing evidence strengthens this interpretation: the clustering of early defaults at challenger banks and specialist lenders—30 to 68 per cent sooner than at major banks—is consistent with rapid enforcement without adequate fraud investigation, producing a temporal signature that mirrors what under-detection would generate.

Beyond incidence, we show that recorded fraud depends on lender classification. This measurement issue is central to policy evaluation: when government relies on lender reporting to identify fraud, heterogeneity in detection and classification can bias cross-lender comparisons and aggregate fraud estimates. The NAO's concern that lender incentives for fraud investigation are 'limited' appears validated by systematic variation in classification rates across lender types.

The official evaluation finding of a 3.8 benefit-to-cost ratio indicates that the COVID loan schemes delivered substantial net value despite the elevated losses we document. Our contribution is to illuminate the composition and distribution of those losses. The scheme design trade-offs identified by the NAO - speed versus verification, access versus screening - generated predictable consequences that our borrower-level analysis quantifies precisely.

For future crisis interventions, the evidence supports proportionate verification even under time pressure, continued reliance on credit information, attention to governance indicators, and explicit consideration of lender investigation capacity when designing guarantee claim and recovery processes. The AFT results offer a further practical contribution: by quantifying how much specific risk factors compress survival time, they provide a basis for calibrating early warning thresholds and prioritising recovery resources toward the borrower and lender segments where default materialises fastest.

References

Acharya, V.V., Crosignani, M., Eisert, T. & Eufinger, C. (2020) Zombie credit and dis-inflation: Evidence from Europe. NBER Working Paper 27158. doi: 10.3386/w27158.

Berg, T., Burg, V., Gombović, A. & Puri, M. (2020) On the rise of fintechs: Credit scoring using digital footprints. Review of Financial Studies, 33(7), 2845–2897. doi: 10.1093/rfs/hhz099.

Balyuk, T. (2023) FinTech lending and bank credit access for consumers. Management Science, 69(1), 555–575. doi: 10.1287/mnsc.2022.4319.

Banerjee, R. & Hofmann, B. (2022) Corporate zombies: Anatomy and life cycle. Economic Policy, 37(112), 757–803. doi: 10.1093/epolic/eiac027.

Bertoni, F., Colombo, M.G. & Quas, A. (2023) The long-term effects of loan guarantees on SME performance. Journal of Corporate Finance, 80, 102408. doi: 10.1016/j.jcorpfin.2023.102408.

Boot, A.W.A. & Thakor, A.V. (2000) Can relationship banking survive competition? Journal of Finance, 55(2), 679–713. doi: 10.1111/0022-1082.00223.

British Business Bank (2022) Evaluation of the Bounce Back Loan Scheme. London: British Business Bank.

British Business Bank (2025) Evaluation of the COVID-19 Loan Guarantee Schemes: Year 3 Report. London: British Business Bank.

Buchak, G., Matvos, G., Piskorski, T. & Seru, A. (2018) Fintech, regulatory arbitrage, and the rise of shadow banks. Journal of Financial Economics, 130, 453–483. doi: 10.1016/j.jfineco.2018.03.011.

Cowling, M. (2010) The role of loan guarantee schemes in alleviating credit rationing in the UK. Journal of Financial Stability, 6(1), 36–44. doi: 10.1016/j.jfs.2009.05.007.

Cowling, M. & Mitchell, P. (2003) Is the Small Firms Loan Guarantee Scheme hazardous for banks or helpful to small business? Small Business Economics, 21(1), 63–71. doi: 10.1023/A:1024408932156.

Cowling, M., Liu, W., Yang, H. & Wilson, N. (2025) Loan size concentration under the UK enterprise finance guarantee scheme and SME access to finance. Journal of Small Business Management. doi: 10.1080/00472778.2025.2468781.

Craig, B.R., Jackson, W.E. & Thomson, J.B. (2008) Credit market failure intervention: Do government sponsored small business credit programs enrich poorer areas? Small Business Economics, 30(4), 345–360. doi: 10.1007/s11187-007-9092-6.

Department for Business and Trade (2024) COVID-19 loan guarantee schemes performance data as at 30 September 2024.

Diamond, D.W. (1984) Financial intermediation and delegated monitoring. Review of Economic Studies, 51(3), 393–414. doi: 10.2307/2297430.

Erel, I. & Liebersohn, J. (2022) Can FinTech reduce disparities in access to finance? Journal of Financial Economics, 146(1), 90–118. doi: 10.1016/j.jfineco.2022.05.004.

Fuster, A., Plosser, M., Schnabl, P. & Vickery, J. (2019) The role of technology in mortgage lending. Review of Financial Studies, 32(5), 1854–1899. doi: 10.1093/rfs/hhy110.

Gale, W.G. (1990) Federal lending and the market for credit. Journal of Public Economics, 42(2), 177–193.

Garmaise, M.J. & Natividad, G. (2016) Spillovers in local banking markets. Review of Corporate Finance Studies, 5(2), 139–165. doi: 10.1093/rcfs/cfw005.

Granja, J., Makridis, C., Yannelis, C. & Zwick, E. (2022) Did the Paycheck Protection Program hit the target? Journal of Financial Economics, 145(3), 725–761. doi: 10.1016/j.jfineco.2022.05.006.

Griffin, J.M., Kruger, S. & Mahajan, P. (2023) Did FinTech lenders facilitate PPP fraud? Journal of Finance, 78(3), 1777–1827. doi: 10.1111/jofi.13209.

Hausman, J.A., Abrevaya, J. & Scott-Morton, F.M. (1998) Misclassification of the dependent variable in a discrete-response setting. Journal of Econometrics, 87(2), 239–269.

Jiang, W., Nelson, A.A. & Vytlacil, E. (2014) Liar's loan? Effects of origination channel and information falsification on mortgage delinquency. Review of Financial Studies, 27(11), 3395–3438. doi: 10.1093/rfs/hhu051.

Keys, B.J., Mukherjee, T., Seru, A. & Vig, V. (2010) Did securitization lead to lax screening? Evidence from subprime loans. Quarterly Journal of Economics, 125(1), 307–362. doi: 10.1162/qjec.2010.125.1.307.

Lelarge, C., Sraer, D. & Thesmar, D. (2010) Entrepreneurship and credit constraints. In Lerner, J. & Schoar, A. (eds.) International Differences in Entrepreneurship. Chicago: University of Chicago Press, pp. 243–273.

Mian, A. & Sufi, A. (2017) Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005. Review of Financial Studies, 30(6), 1832–1864. doi: 10.1093/rfs/hhw104.

National Audit Office (2020) Investigation into the Bounce Back Loan Scheme. London: NAO.

National Audit Office (2021) The Bounce Back Loan Scheme: An Update. London: NAO.

Oh, I., Lee, J.-D., Heshmati, A. & Choi, G.-G. (2009) Evaluation of credit guarantee policy using propensity score matching. Small Business Economics, 33(3), 335–351. doi: 10.1007/s11187-008-9102-5.

Riding, A.L. & Haines, G. (2001) Loan guarantees: Costs of default and benefits to small firms. Journal of Business Venturing, 16(6), 595–612. doi: 10.1016/S0883-9026(00)00050-1.

Riding, A., Madill, J. & Haines, G. (2007) Incrementality of SME loan guarantees. Small Business Economics, 29(1), 47–61.

Stiglitz, J.E. & Weiss, A. (1981) Credit rationing in markets with imperfect information. American Economic Review, 71(3), 393–410.

Table 1. Sample construction and default outcomes

Outcome / definition	Count	Percent
Panel A: Full sample outcomes		
Total loans in analytical sample	1,006,579	100.0
Non-default by outcome cutoff	692,684	68.8
Default (any cause)	313,895	31.2
– Insolvency default	286,034	28.4
– Fraud-labelled default	27,861	2.8
Panel B: Composition of defaults (conditional on default)		
Insolvency default	286,034	91.1 of defaults
Fraud-labelled default	27,861	8.9 of defaults
Panel C: Defaults with observed event dates (Cox models)		
Default events with observed default date	271,957	27.0
– Insolvency events	251,024	24.9
– Fraud-labelled events	20,933	2.1
Defaults without observed default date (right-censored in duration models)	41,938	4.2

Notes: Percentages in Panel A and Panel C are expressed as a share of the full analytical sample (N = 1,006,579). Panel B expresses insolvency and fraud-labelled outcomes as shares of all defaults (N = 313,895). Cox model failure counts are lower than the headline default count because some default statuses do not have an observed event date within the administrative extracts; these are treated as right-censored in duration models. Percentages may not sum to 100 due to rounding.

Table 2. Determinants of COVID-19 Government-Guaranteed Loan Outcomes

	(1) All Default	(2) Insolvency	(3) Fraud	(4) Repaid
Panel A: Loan Characteristics				
Log(Loan Amount)	0.250***	0.119***	1.325***	−0.201***
	(0.009)	(0.009)	(0.034)	(0.008)
Log(Sales)	−0.040***	−0.033***	−0.077**	0.128***
	(0.009)	(0.009)	(0.031)	(0.007)
Loan/Sales Ratio	1.719***	1.530***	2.385***	0.543***
	(0.083)	(0.083)	(0.278)	(0.070)
Bounce Back Loan	1.105***	0.769***	5.077***	−0.873***
	(0.021)	(0.020)	(0.141)	(0.013)
Loan Term (months)	−0.014***	−0.012***	−0.017***	−0.031***
	(0.000)	(0.000)	(0.001)	(0.000)
Extended Repayment	1.232***	1.261***	−0.048**	−0.964***
	(0.008)	(0.008)	(0.024)	(0.010)
Panel B: Firm Characteristics				
PD Score	3.397***	3.334***	2.656***	−2.521***
	(0.034)	(0.034)	(0.105)	(0.044)
PD Missing	1.101***	0.928***	1.780***	−0.507***
	(0.007)	(0.007)	(0.023)	(0.008)
Prior Equity Finance	−0.071**	−0.029	−0.552***	0.073**
	(0.030)	(0.030)	(0.122)	(0.029)
IMD Deprivation Score	−0.000***	−0.000***	−0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Panel C: Director Characteristics				
Director Experience (years)	−0.036***	−0.034***	−0.044***	0.012***
	(0.000)	(0.000)	(0.002)	(0.000)
Female Directors (%)	−0.002***	−0.001***	−0.002***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Board Size	−0.220***	−0.200***	−0.267***	0.078***
	(0.004)	(0.004)	(0.013)	(0.003)
Foreign Nationals (%)	0.002***	0.001***	0.003***	−0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
White Directors (%)	0.001***	0.002***	−0.001***	−0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Industry / Lender / Region FE	Yes	Yes	Yes	Yes
Output Area FE	Yes	Yes	Yes	Yes
Observations	1,006,579	1,006,579	1,005,912	1,006,579
Pseudo R^2	0.138	0.124	0.160	0.152
AUC	0.751	0.740	0.824	0.761

Notes: Log-odds coefficients from logit regressions with robust standard errors in parentheses. Dependent variables: (1) any default; (2) insolvency default; (3) fraud default; (4) full repayment. Sample: limited companies receiving BBLS/CBILS loans. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 3. Cox Proportional Hazard Models: Time to Default

	(1) All Default	(2) Insolvency	(3) Fraud
Panel A: Loan Characteristics			
Log(Loan Amount)	1.182***	1.094***	3.673***
	(0.009)	(0.008)	(0.133)
Log(Sales)	1.012	1.018**	0.936*
	(0.007)	(0.007)	(0.033)
Loan/Sales Ratio	6.552***	5.722***	37.214***
	(0.433)	(0.387)	(11.995)
Bounce Back Loan	2.316***	1.863***	170.342***
	(0.040)	(0.032)	(25.354)
Loan Term (months)	0.988***	0.988***	0.981***
	(0.000)	(0.000)	(0.001)
Extended Repayment	2.118***	2.281***	0.911***
	(0.012)	(0.014)	(0.020)
Panel B: Firm Characteristics			
IMD Deprivation Score	1.000***	1.000***	1.000***
	(0.000)	(0.000)	(0.000)
Prior Equity Finance	1.043*	1.081***	0.565***
	(0.023)	(0.025)	(0.066)
Panel C: Director Characteristics			
Director Experience (years)	0.956***	0.959***	0.912***
	(0.000)	(0.000)	(0.002)
Female Directors (%)	0.999***	0.999***	0.997***
	(0.000)	(0.000)	(0.000)
Board Size	0.812***	0.820***	0.723***
	(0.002)	(0.003)	(0.010)
Foreign Nationals (%)	1.001***	1.001***	1.003***
	(0.000)	(0.000)	(0.000)
White Directors (%)	1.001***	1.001***	0.998***
	(0.000)	(0.000)	(0.000)
Industry / Lender / Region FE	Yes	Yes	Yes
Output Area FE	Yes	Yes	Yes
Observations	1,006,579	1,006,579	1,006,579
Failures	271,957	251,024	20,933
Log-Likelihood	−3,613,924	−3,333,336	−272,344
Chi-squared	103,581	96,362	16,302

Notes: Hazard ratios from Cox proportional hazard models with robust standard errors in parentheses. Failure events: (1) any default; (2) insolvency default; (3) fraud default. Sample comprises limited companies receiving BBLS/CBILS loans. Analysis time measured in days from loan origination. *** p < 0.01, ** p < 0.05, * p < 0.10

Table 3A: AFT Distribution Comparison — All Default

Time ratios from four parametric AFT models. Limited company sample only.

Variable	Weibull	Log-logistic	Log-normal	Gen. Gamma
Log(Sales)	0.993* (0.004)	0.990** (0.004)	0.989*** (0.004)	0.987*** (0.004)
bounce_back=0	1.000 (.)	1.000 (.)	1.000 (.)	1.000 (.)
Loan Term (months)	1.007*** (0.000)	1.008*** (0.000)	1.007*** (0.000)	1.007*** (0.000)
equity_fin=0	1.000 (.)	1.000 (.)	1.000 (.)	1.000 (.)
Extended_Period=0	1.000 (.)	1.000 (.)	1.000 (.)	1.000 (.)
Director Experience (mean years)	1.025*** (0.000)	1.025*** (0.000)	1.024*** (0.000)	1.022*** (0.000)
Board Size	1.119*** (0.002)	1.124*** (0.002)	1.117*** (0.002)	1.096*** (0.004)
White Directors (%)	1.000*** (0.000)	1.000*** (0.000)	1.000*** (0.000)	1.000*** (0.000)
ln_p	1.876*** (0.002)			
lnsigma			0.818*** (0.001)	0.963*** (0.002)
Observations	1,006,579	1,006,579	1,006,579	1,006,579
Failures	271,957	271,957	271,957	271,957
Log-Likelihood	-629565.47	-622640.79	-614851.27	-611870.75

*Time ratios (exponentiated coefficients). Standard errors in parentheses. TR < 1 indicates covariate accelerates default (shorter survival time); TR > 1 indicates delay. Industry, lender, output area, and region fixed effects included in all models. *** p<0.01, ** p<0.05, * p<0.1*

Table 3B: Predicted Survival Times by Loan and Lender Type (Gen. Gamma)

Predicted median and mean survival times for defaulted loans, by scheme type.

Loan Type	Mean (days)	Median (days)	N
CBILS/CLBILS	3,873	3,623	7,508
BBLS	2,782	2,667	264,449
Difference	**-1,091**	**-956**	

Predictions based on Generalised Gamma AFT model. Sample: limited companies with default_all==1. Survival time in days from origination. Divide by 30.44 for months, by 365.25 for years. BBLS borrowers default approximately 956 days (31 months / 2.6 years) earlier at the median.

Predicted median and mean survival times for defaulted loans, by lender type.

Lender Type	Mean (days)	Median (days)	N
Major banks	2,960	2,797	223,428
Irish/Scottish banks	3,041	2,852	6,003
International banks	4,784	4,618	41
Co-operative banks	3,425	3,243	1,009
Challenger banks	1,887	1,804	35,931
Private banks	3,160	2,981	303
Lease finance	2,650	2,358	270
Invoice finance	2,104	2,038	471
Fintech	2,648	2,556	3,537
Development banks	5,798	4,016	5
CDFIs	2,241	1,898	170
Government lenders	2,918	2,680	115
Specialist lenders	1,176	1,077	674

Predictions based on Generalised Gamma AFT model. Sample: limited companies with default_all==1. Survival time in days from origination. Divide by 30.44 for months, by 365.25 for years. Challenger banks and specialist lenders experience substantially earlier defaults than major banks.

Table 4. Multinomial Logit: Determinants of Loan Outcomes

(Base Outcome: No Default)

	Non-Fraud Default		Fraud Default	
	RRR	**(SE)**	**RRR**	**(SE)**
Panel A: Loan Characteristics				
Log(Loan Amount)	1.245***	(0.012)	3.963***	(0.132)
Log(Sales)	0.949***	(0.008)	0.941**	(0.029)
Loan/Sales Ratio	4.946***	(0.414)	27.427***	(7.585)
Bounce Back Loan	2.586***	(0.051)	230.597***	(33.304)
Loan Term (months)	1.009***	(0.000)	0.986***	(0.000)
Panel B: Firm Characteristics				
PD Score	49.090***	(1.613)	49.721***	(5.425)
PD Missing	2.927***	(0.019)	8.252***	(0.188)
Prior Equity Finance	0.989	(0.029)	0.640***	(0.077)
Director Experience (years)	0.961***	(0.000)	0.946***	(0.002)
Board Size	0.794***	(0.003)	0.708***	(0.009)
Panel C: Lender Type (ref: Main Bank)				
Irish/Scottish	0.789***	(0.013)	1.031	(0.046)
Co-operative	0.685***	(0.027)	2.479***	(0.197)
Challenger	1.868***	(0.016)	0.901***	(0.023)
Private	1.173**	(0.082)	1.710	(0.538)
Lease Finance	1.548***	(0.119)	7.617***	(1.372)
Invoice Finance	3.774***	(0.233)	3.311***	(1.346)
Fintech	2.197***	(0.054)	1.264**	(0.132)
Development	0.029***	(0.016)	0.028***	(0.021)
CDFI	3.196***	(0.372)	3.346***	(1.293)
Specialist	3.623***	(0.212)	18.788***	(2.591)
Observations		1,006,611		
Pseudo R^2		0.111		

Notes: Relative risk ratios from multinomial logit with robust standard errors in parentheses. Base outcome is no default. Sample: limited companies receiving BBLS/CBILS loans. Additional controls for female directors (%), foreign nationals (%), white directors (%), and IMD deprivation not shown. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 5. Fraud Classification Conditional on Default: Do Lenders Differ in Fraud Detection?

	(1) Baseline	(2) + Loan	(3) + Firm	(4) + Director	(5) + FEs
Main Bank (reference)	—	—	—	—	—
Irish/Scottish	1.108^{***} (0.044)	1.145^{***} (0.047)	1.092^{**} (0.047)	1.382^{***} (0.066)	1.315^{***} (0.061)
Co-operative	2.434^{***} (0.180)	2.604^{***} (0.217)	3.081^{***} (0.264)	3.607^{***} (0.340)	3.709^{***} (0.357)
Challenger	0.584^{***} (0.013)	0.487^{***} (0.011)	0.391^{***} (0.009)	0.455^{***} (0.011)	0.427^{***} (0.011)
Private	0.240^{***} (0.059)	0.553^{**} (0.161)	0.602^{*} (0.168)	0.874 (0.289)	0.844 (0.279)
Lease Finance	2.183^{***} (0.271)	4.025^{***} (0.550)	3.247^{***} (0.458)	3.254^{***} (0.581)	3.186^{***} (0.583)
Invoice Finance	0.229^{***} (0.058)	1.018 (0.273)	0.999 (0.269)	0.764 (0.268)	0.724 (0.253)
Fintech	0.323^{***} (0.028)	0.526^{***} (0.046)	0.470^{***} (0.043)	0.550^{***} (0.056)	0.524^{***} (0.054)
CDFI	0.522^{**} (0.143)	1.077 (0.334)	0.873 (0.280)	0.865 (0.325)	0.832 (0.317)
Specialist	2.515^{***} (0.210)	3.639^{***} (0.382)	4.204^{***} (0.458)	5.188^{***} (0.615)	5.124^{***} (0.608)
Loan characteristics	No	Yes	Yes	Yes	Yes
Firm characteristics	No	No	Yes	Yes	Yes
Director characteristics	No	No	No	Yes	Yes
Industry/Region/OA FEs	No	No	No	No	Yes
Observations	313,817	313,817	313,814	271,930	271,912
Pseudo R^2	0.007	0.081	0.118	0.112	0.121

Notes: Odds ratios from logit regressions with robust standard errors in parentheses. Sample restricted to defaulted loans from limited companies. Dependent variable equals 1 if the default was classified as fraud, 0 otherwise. Loan characteristics include log loan amount, log sales, loan/sales ratio, BBLS indicator, and loan term. Firm characteristics include PD score, PD missing, IMD deprivation, and prior equity finance. Director characteristics include experience, female percentage, board size, foreign national percentage, and white percentage. International and Government lender categories omitted due to small cell sizes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 6: Propensity Score Analysis—Actual vs Predicted Fraud Rates

Lender Type	Actual	Predicted	Gap	Gap %	N Defaults
Main Bank	9.38%	7.97%	+1.41pp	+18%	256,684
Irish/Scottish	10.29%	6.77%	+3.52pp	+52%	7,224
Co-operative	20.12%	6.35%	+13.77pp	+217%	1,143
Challenger	5.70%	11.85%	−6.15pp	−52%	41,223
Lease Finance	18.43%	7.99%	+10.44pp	+131%	434
Invoice Finance	2.32%	2.02%	+0.30pp	+15%	690
Fintech	3.24%	5.11%	−1.87pp	−37%	4,388
CDFI	5.13%	6.67%	−1.54pp	−23%	273
Specialist	20.66%	5.87%	+14.79pp	+252%	881

Notes: Predicted fraud rate calculated using coefficients from a logit model estimated on main bank defaults only, applied to all lender types. Gap = Actual − Predicted. Negative gaps indicate under-classification relative to main bank benchmark; positive gaps indicate over-classification or superior detection.

Table 7.1 Characteristic Comparison Across Default Types

	Main Bank Fraud	Challenger Non-Fraud	Challenger Fraud	Main Bank Non-Fraud
Log(Loan Amount)	10.54	10.39	10.69	10.25
Loan/Sales Ratio	0.206	0.207	0.202	0.200
PD Score (%)	3.92	4.15	2.75	7.06
Director Experience (years)	4.96	4.72	3.72	6.45
Board Size	1.34	1.39	1.39	1.46
Female Directors (%)	22.3	23.7	21.9	26.1
Foreign Nationals (%)	26.0	18.8	29.3	18.4
Loan Term (months)	81.3	90.5	80.3	89.2
N	24,077	38,872	2,351	232,607

Notes: Mean values for key characteristics across four groups: main bank fraud defaults, challenger non-fraud defaults, challenger fraud defaults, and main bank non-fraud defaults. The comparison between columns 1 and 2 (Main Bank Fraud vs Challenger Non-Fraud) is of primary interest for the under-classification hypothesis.

Table 7.2: T-Tests Comparing Main Bank Fraud to Challenger Non-Fraud Defaults

Variable	Main Bank Fraud	Challenger Non-Fraud	Difference	p-value
Log(Loan Amount)	10.54	10.39	+0.148	<0.001
Loan/Sales Ratio	0.206	0.207	−0.001	0.241
PD Score (%)	3.92	4.15	−0.23	0.002
Director Experience	4.96	4.72	+0.24	<0.001
Board Size	1.34	1.39	−0.05	<0.001

Notes: Two-sample t-tests comparing means between main bank fraud defaults (N = 24,077) and challenger bank non-fraud defaults (N = 38,872). Positive difference indicates main bank fraud mean exceeds challenger non-fraud mean.

Appendix:

Propensity model: Fraud classification conditional on default (logit): Main Banks

Variable	Coefficient	Robust SE
Panel A. Covariates		
ln(Loan amount)	1.181***	0.0367
ln(Sales)	0.015	0.0332
Loan-to-sales ratio	2.426***	0.312
Bounce Back Loan (indicator)	3.355***	0.151
Loan term (months)	-0.0240***	0.000440
PD (pd100_ensemble)	0.396***	0.128
PD missing (indicator)	1.102***	0.0255
Index of Multiple Deprivation (imd_1)	-5.77e-06***	1.30e-06
Equity finance experience (indicator)	-0.312**	0.128
Mean director experience	-0.0130***	0.00177
Female directors (%)	-0.00159***	0.000222
Board size (winsorised)	-0.0966***	0.0138
Foreign nationals (%)	0.00167***	0.000212
White (%)	-0.00234***	0.000195
Panel B. Fixed effects / controls		
Industry fixed effects (SIC)	Included	
Geodemographic fixed effects (OAC)	Included	
Region fixed effects	Included	
Constant	-16.943***	0.362
Panel C. Model statistics		
Observations	223,428	
Wald chi2 (df)	12,194.43 (53)	
Prob > chi2	0.0000	
Log pseudolikelihood	-54,390.818	
Pseudo R-squared	0.1245	
Variance estimator	Robust	

Notes: Logit coefficients reported. Robust standard errors in the third column. Significance: *** p<0.01, ** p<0.05, * p<0.10. Estimation restricted to defaulted limited companies and major-bank lenders, as specified in the Stata command.