BANCA D'ITALIA
EUROSISTEMA

Mercati, infrastrutture, sistemi di pagamento

(Markets, Infrastructures, Payment Systems)

Social and Governance Factors in Default Risk Assessment

by Manuel Cugliari, Stefano Di Virgilio, Enrico Foscolo and Alessandra Iannamorelli

July 2026

Number 84

BANCA D'ITALIA
EUROSISTEMA

Mercati, infrastrutture, sistemi di pagamento

(Markets, Infrastructures, Payment Systems)

Social and Governance Factors in Default Risk Assessment

by Manuel Cugliari, Stefano Di Virgilio, Enrico Foscolo and Alessandra Iannamorelli

Number 84 – July 2026

The papers published in the 'Markets, Infrastructures, Payment Systems' series provide information and analysis on aspects regarding the institutional duties of the Bank of Italy in relation to the monitoring of financial markets and payment systems and the development and management of the corresponding infrastructures in order to foster a better understanding of these issues and stimulate discussion among institutions, economic actors and citizens.

The views expressed in the papers are those of the authors and do not necessarily reflect those of Banca d'Italia.

The series is available online at www.bancaditalia.it.

Printed copies can be requested from the Paolo Baffi Library: richieste.pubblicazioni@bancaditalia.it.

ISSN 2724-6418 (online)
ISSN 2724-640X (print)

Banca d'Italia
Via Nazionale, 91 - 00184 Rome - Italy
+39 06 47921

Designed and printing by the Printing and Publishing Division of Banca d'Italia

SOCIAL AND GOVERNANCE FACTORS IN DEFAULT RISK ASSESSMENT

by Manuel Cugliari*, Stefano Di Virgilio*, Enrico Foscolo* and Alessandra Iannamorelli*

Abstract

We investigate whether social (S) and corporate governance (G) indicators help to better assess default risk for Italian non-financial corporations in Banca d'Italia's In-house Credit Assessment System (ICAS). Specifically, S indicators use data from the Italian National Social Security Institute (INPS) database, and G indicators are derived from a structured database of Italian firms and provide information on corporate relationships and governance traits not included in standard financial data. Our results show that S and G indicators improve default prediction. They are statistically significant for micro and small firms, refining the baseline model that includes financial and credit behaviour variables. However, their overall contribution to the model's performance is moderate. For medium-sized and large firms, only S indicators are statistically significant.

JEL Classification: G33, C51, C52, C55.

Keywords: probability of default, credit risk modelling, feature selection, collateral eligibility assessment, machine learning.

Sintesi

Il lavoro valuta se indicatori comunemente associati alle dimensioni sociale (S) e di governo societario (G) migliorino la stima del rischio di *default* delle imprese non finanziarie italiane nell'ambito del sistema interno di valutazione del merito creditizio (*In-house Credit Assessment System*, ICAS) della Banca d'Italia. Gli indicatori S utilizzano informazioni tratte dal database dell'INPS. Gli indicatori G derivano da una base dati sulle imprese italiane, dalla quale sono ricavate informazioni su relazioni societarie e assetto di governo non incluse nelle variabili finanziarie *standard*. I risultati mostrano che gli indicatori S e G migliorano la previsione dell'evento di *default*. In particolare, essi risultano significativi per imprese micro e piccole, consentendo di affinare le stime del modello statistico di base che include variabili finanziarie e di comportamento creditizio. Il contributo alla *performance* del modello resta comunque moderato. Per le imprese di medie e grandi dimensioni solo gli indicatori S risultano statisticamente significativi.

* Banca d'Italia, Financial Risk Management Directorate.

CONTENTS

1 Introduction[1]

In recent years, the assessment of corporate default risk has increasingly explored the contribution of non-financial information alongside traditional financial and credit indicators. Within this literature, several studies test whether variables commonly associated with the social and governance domains improve default prediction. For example, indicators such as board characteristics, management experience, employee tenure, and hiring and firing ratios have been examined as potential predictors of default probability for small and medium-sized enterprises (Altman *et al.*, 2023).

In this paper, we adopt a predictive perspective. We investigate whether selected indicators conventionally classified within the social and governance domains improve the accuracy of corporate default prediction for Italian non-financial firms, as assessed within the Banca d'Italia's statistical model of the In-house Credit Assessment System (henceforth S-ICAS; Giovannelli *et al.*, 2023, Narizzano *et al.*, 2024; Di Virgilio *et al.*, 2025). S-ICAS is the first, largely automated building block of the overall ICAS, which combines the statistical model with the expert assessment by credit analysts to produce firms' final ratings.

In this study, the S- and G-related indicators under consideration are selected because they are commonly classified within these domains by sustainability reporting frameworks and external data providers. We use this taxonomy for the purpose of credit risk assessment, rather than to evaluate firms' social or governance quality. Accordingly, the analysis focuses on the contribution of S- and G-related indicators to default risk within the S-ICAS framework, without attributing causal or value-based interpretations to the estimated relationships.

While the literature mainly deals with bankruptcy (Altman, 1968; Altman *et al.*, 1977; Norden and Weber, 2010; Ciampi and Gordini, 2013; Ciampi, 2015; Tian *et al.*, 2015; Barboza *et al.*, 2017; Cultrera, 2020; García and Herrero, 2021; Paraschiv *et al.*, 2024), the default definition that matters for ICAS relies on Article 178 of Regulation (EU) No 575/2013, which identifies a default when a bank considers the obligor unlikely to pay its credit obligations or the obligor is past due more than 90 days on any material exposures to the bank.[2] For this reason, the default status is generally reversible, unlike bankruptcy. We investigate the role of explanatory variables such as ownership structure, external audit firm appointments, board stability and independence, board experience and effectiveness, wages, and employee turnover. While governance information is usually employed by credit analysts in the expert assessment of ICAS, which involves a qualitative judgment based on various information sources, to our knowledge this paper provides the first attempt to include a wide range of S and G indicators within statistical credit assessment

[1] We thank Paolo Del Giovane, Antonio Scalia, Francesco Columba, Francesco Monterisi, Simone Narizzano, Paolo Parlamento, and an anonymous referee for their useful comments.

[2] Narizzano *et al.* (2024) detail the default definition applied by ICAS.

in the Eurosystem framework. The sample size, made of hundreds of thousands of non-financial firms, provides robustness to the analysis.

We leverage traditional and novel data sources to construct the S- and G-related indicators. Our sources include data from the Italian National Social Insurance Agency (INPS) and a proprietary knowledge graph dataset termed KG19, developed by Banca d'Italia and based on information from the Italian Business Register. The INPS dataset contains detailed information on Italian firms and employees. It allows the construction of workforce-related indicators, such as employment patterns and wage-related measures. The KG19 dataset enables automated reasoning and probabilistic analysis for obtaining governance-related indicators from firms' ownership and organisational structures (Bellomarini *et al.*, 2019).

To preserve the current S-ICAS architecture, we first group firms by size and we perform feature selection for the S- and G-related indicators combining visual inspection and variable importance ranking with univariate logistic regressions and random forests. We thus construct a list of statistically significant indicators according to firm size (from now on, the watchlists; e.g., James *et al.*, 2023). In particular, random forests help to identify non-monotonic and possibly non-linear effects and reduce the number of variables in each watchlist. Since governance-related indicators are typically associated with financial performance outcomes, a multicollinearity analysis is performed between the watchlists and indicators from the financial statements involved in the baseline S-ICAS. Afterwards, we assess company default risk with two independent scores, one based on social watchlists (the S score) and the other on governance watchlists (the G score), both obtained with multivariate logistic regressions. Unlike conventional ESG scores, these scores are not intended to assess firms' social or governance performance. Rather, they summarise the statistical contribution of the selected indicator sets to default prediction within the ICAS framework.

Our findings align with prior evidence that workforce-related indicators contribute to default prediction. Key workforce-related indicators, including *layoff rates*, *employee turnover*, and *average wages*, display statistically significant discriminatory and predictive power. In this paper, these indicators are treated as operational proxies with predictive relevance and are not interpreted as indicators of firms' social quality. Their estimated associations with default should be interpreted, therefore, in statistical terms rather than as causal or value-based relationships.

Governance-related indicators contribute differently across firm size classes. For instance, *board age diversity* emerges as a statistically significant predictor for all classes, while *board size* does so for micro and small firms only. Here again, these indicators are used as structural proxies with predictive relevance and are not interpreted as direct measures of governance quality.

In the final step of the analysis, the S and G scores are integrated into the baseline S-ICAS, which relies on credit behaviour and financial components. The baseline model employs logistic regressions to estimate the PD (probability of default) over a one-year horizon across different sectors and size groups (Hosmer *et al.*, 2013; Narizzano *et al.*, 2024; Giovannelli *et al.*, 2023; Di Virgilio *et al.*, 2025). Rather than redesigning the statistical framework of S-ICAS, our contribution builds on its existing structure and tests the incremental predictive value of selected S- and G-related indicators. The exercise is therefore intentionally narrow in scope. It does not aim to provide causal explanations, normative assessments, or ESG-style evaluations of firms, but only to verify whether these indicator families improve the discriminatory and predictive performance of the baseline model. An explicit causal analysis is deferred to future research. The augmented S-ICAS – with the inclusion of the S and G scores – improves the discriminatory power of the baseline model, as measured by the out-of-sample Area under the Receiver Operating Characteristic (AuROC).[3] Specifically, the AuROC increase ranges between 0.5 and 0.9 percentage points for micro and small firms. While such an increase may seem relatively modest in absolute terms, it should be assessed considering the high baseline AuROC, which lies between 84 and 90 per cent. For medium-sized firms, we observe similar results, with increments between 0.6 and 0.8 percentage points; in this case the governance score is not statistically significant. For large firms, the governance score and the AuROC increment are not significant. Overall, the augmented S-ICAS shows moderately higher predictive performance than the baseline model, especially for micro and small firms.

The modest significance of the governance score in predicting default of larger firms is not necessarily at odds with the existing literature on the importance of governance-related indicators. Previous studies suggest that governance variables may improve default prediction when combined with credit behaviour and/or financial ratios (Altman *et al.*, 2023; Aulia and Wijaya, 2020). In our study we divide firms into four different size classes and conduct a separate analysis for each class. For large firms the relatively low number of companies and the low occurrence of default – driven by very adverse circumstances – are a challenge for the statistical model. In summary, while the governance dimension is statistically significant only for micro and small firms, selected social-related indicators emerge as statistically relevant default predictors in all groups, suggesting that they may be relevant candidates for inclusion in S-ICAS.

The remainder of the paper is organised as follows. Section 2 presents the literature review. Section 3 illustrates the ICAS framework and the S-ICAS. Section 4 describes the data sources and presents the descriptive analysis of the S and G indicators. Section 5 presents the results of the social and governance logistic regression outcomes and the integration of the S and G scores into S-ICAS. Section 6 concludes.

[3] The AuROC is the Area under the Receiver Operating Characteristic Curve (ROC). A ROC curve shows the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds. In this context, AuROC measures the model's ability to assign higher PDs to firms that will default compared with financially healthy companies. The AuROC is thus a performance metric that can be used to evaluate classification models. An AuROC of 0.5 corresponds to a coin flip, i.e. a useless model; an AuROC below 0.7 reveals a sub-optimal performance; an AuROC at 0.7 – 0.8 signals good performance; an AuROC greater than 0.8 means excellent performance; finally, an AuROC of 1.0 corresponds to a perfect classifier. See Fawcett (2004) and Chawla (2009) for more details.

2 Literature review

A growing strand of literature investigates whether non-financial information, including indicators commonly associated with the social and governance domains, can improve the assessment of corporate credit risk when used alongside traditional financial and credit indicators. Within this line of research, we focus only on contributions in which such variables are analysed in relation to default risk, bankruptcy risk, creditworthiness, or credit ratings. This distinction is important because the objective of the paper is to assess default risk, rather than to evaluate firms' social or governance quality.

Among the contributions most closely related to our perspective, Altman *et al.* (2023) show that the inclusion of selected environmental, social, and governance variables may improve the estimation of default probability for small and medium-sized enterprises when combined with traditional risk indicators. Similarly, Aulia and Wijaya (2020) find that governance-related variables can enhance default prediction, supporting the view that information beyond standard accounting measures may contribute to credit risk modelling. More generally, the bankruptcy prediction literature has long recognised that enlarging the information set beyond conventional balance-sheet ratios may improve model performance, provided that the additional variables offer genuine discriminatory power and are not merely redundant with existing predictors.

Recent empirical studies also explore the relationship between ESG-related information and firms' creditworthiness using alternative credit risk proxies. Bhattacharya and Sharma (2019), using a sample of 122 firms from the BSE 500, analyse the relationship between ESG scores and credit ratings through an ordered logistic regression model. Their results show that both aggregate ESG scores and their individual components are positively associated with firms' credit ratings, suggesting that ESG-related information may contain signals relevant for creditworthiness, even though the setting concerns ratings rather than default events.

A more granular approach is adopted by Bonacorsi *et al.* (2024), who study a cross-section of European listed firms using a large set of raw ESG factors sourced from MSCI. Their proxy for credit risk is Altman's *z*-score, and their empirical strategy combines supervised machine learning techniques with a second-stage assessment of the impact of selected variables. Their findings indicate that ESG-related raw indicators can improve the empirical modelling of firms' credit risk when used together with traditional accounting information. This contribution is especially relevant for our study because it moves beyond aggregate ESG scores and highlights the informational value of disaggregated indicators.

Related evidence is provided by Vivel-Búa *et al.* (2024), who examine 990 non-financial listed firms in the Eurozone over the period 2004–2020, using ESG scores and Altman's *z*-score as a proxy for default risk. They find that higher ESG scores are associated with higher *z*-scores, that is, with lower default risk, although the strength of the relationship varies with the economic cycle and firm size. Their results confirm that ESG-related information may be relevant for default risk assessment, while also suggesting that its contribution is not uniform across firms and macroeconomic conditions.

Taken together, these studies suggest that variables commonly classified within the social and governance domains may contribute to credit risk assessment. At the same time, most of the existing evidence relies either on external ESG scores, listed-company samples, or credit risk proxies such as ratings and *z*-scores. Our contribution differs in three respects. First, we do not rely on externally produced ESG scores. Instead, we construct highly granular social-related indicators from the INPS database and governance-related indicators from the KG19 knowledge graph. Second, our analysis is conducted within the ICAS framework, that is, within an operational credit assessment system used for the evaluation of Italian non-financial firms. Third, the purpose of our exercise is intentionally narrow: we do not interpret the selected indicators as measures of firms' social or governance quality, nor do we aim to identify causal mechanisms. Rather, we test whether these S- and G-related indicator families provide incremental discriminatory and predictive power with respect to the default event.

3 The statistical ICAS model

ICAS has been running since 2013 to assess the creditworthiness of Italian non-financial firms.[4] It is an important component of the Eurosystem credit assessment framework (ECAF), contributing to the evaluation of credit claims that banks pledge as collateral in monetary policy operations (Narizzano *et al.*, 2024; Giovannelli *et al.*, 2023; Di Virgilio *et al.*, 2025). This feature is especially important during periods of financial distress, when a diversified range of eligible collateral can support liquidity provision to the financial system.

The primary function of ICAS is to estimate firms' PD over a yearly horizon. According to the ECAF requirements, ICAS employs a two-step rating process that integrates a quantitative model with the largely qualitative expert assessment, resulting in reasoned PD estimates.[5]

The initial step involves the application of the statistical model. S-ICAS is run monthly over approximately 370,000 Italian non-financial firms. The dataset includes credit behaviour data on firms' financial liabilities and payment patterns, as well as financial ratios, related for example to debt sustainability, financial structure, and liquidity. Credit data is updated monthly with a batch from the National Credit Register (NCR). This step yields the 'statistical PD' for each firm in the sample. The second step, which is carried out every year, involves a subsample of almost 4,000 companies deemed as the most significant for collateral purposes. In this step, credit analysts conduct a review of the PD estimates obtained with S-ICAS. At least two credit analysts examine additional quantitative and qualitative information such as management quality, industry-specific risks, and financial group structure.

[4] Similar systems also operate at other Eurosystem national central banks.
[5] Appendix 1 briefly summarises the ICAS framework.

Adjustments are based on a thorough evaluation of these factors. This step, yielding the 'full PD', ensures that the final credit rating for the subsample of the most important firms reflects quantitative analysis and expert insight.

The combined approach allows ICAS to achieve a high degree of accuracy in assessing the creditworthiness of firms, from small enterprises to large corporations.

The statistical model has two base components. In line with banking industry practices, each component involves dedicated logistic regressions and provides a credit score for each firm as output. The components are as follows:

- The credit behaviour component uses individual firm data from the NCR. According to size, different sub-models are estimated for each firm class.[6] Companies are grouped into micro, small, and medium- and large-sized classes. Each of the three sub-models uses specific explanatory variables.
- The financial component employs financial data on individual firms provided by Cerved Group.[7] This component involves 11 sub-models based on the macro-sector and the type of financial statement (ordinary or simplified). The sectors include Industry, Trade, Construction, Services, Real estate, and Holdings. Two models are estimated for each of the first five sectors: one for companies with an ordinary financial statement and one for companies with a simplified financial statement. For the Holding sector, a single model is estimated. Each of the 11 sub-models uses specific explanatory variables.

The results of these two components are then merged through an integration module that provides the S-ICAS baseline final score. The integration module involves four sub-models for firm size (micro, small, medium, and large). The score produced by S-ICAS is mapped into the statistical PD via the inverse logit function. Namely, the score is defined as the natural logarithm of odds. [8] Mathematically:

$$score = \ln\left(\frac{PD}{1-PD}\right) \quad (1)$$

A higher score implies a higher probability of default. The statistical PD in turn is mapped into a rating. Figure 1 summarises the entire model architecture.

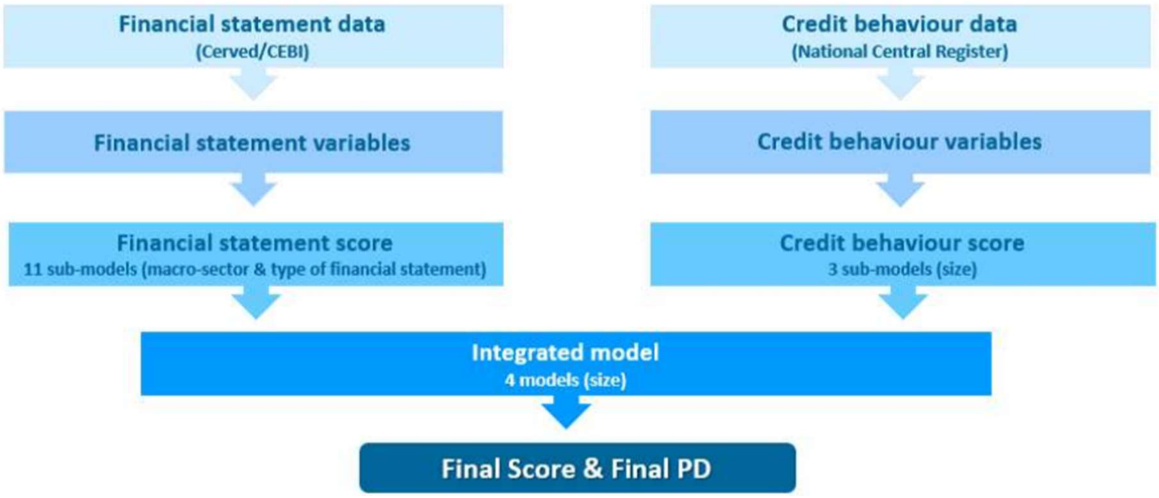
---

[6] As defined by the European Commission 2003/361/EC, according to article 2 of the Annex, the category of micro, small and medium-sized enterprises (SMEs) is made up of enterprises that employ fewer than 250 persons and have annual turnover not exceeding EUR 50 million, and/or total assets not exceeding EUR 43 million. Within the SME category, a small enterprise is defined as one that employs fewer than 50 persons and whose annual turnover and/or total assets do not exceed EUR 10 million. Within the SME category, a micro-enterprise is an enterprise that employs fewer than 10 persons whose annual turnover and/or total assets do not exceed EUR 2 million.

[7] Cerved Group maintains an extensive dataset covering the Italian corporate sector, which includes nearly all small and micro limited-liability firms. This dataset is sourced from the National Official Business Register.

[8] See Narizzano *et al.* (2024) for more details.

Figure 1. The S-ICAS architecture



This approach enables the integration of heterogeneous information within the rating system. Under a hypothetical unified model, the potentially strong correlation between accounting variables and credit relationship variables could introduce bias, thereby reducing predictive accuracy. Furthermore, discrepancies in data frequency (monthly for NCR versus annually for financial statements) and the inherent reporting lags (up to 12 months for financial statements and 2 months for NCR) may lead to inaccuracies in PD estimates (Giannozzi *et al.*, 2013). Moreover, having two separate components makes the model more interpretable for financial analysts and it also makes it possible to produce a default probability even for firms for which financial statements are not available (for example partnerships).

4 Data

To investigate the improvement of the predictive and discriminatory power of the statistical model that could be obtained by considering S- and G-related indicators, we construct a dataset that uses the following two sources in addition to the ICAS information.

- *INPS dataset*: INPS runs the Italian public retirement system. All employed and most self-employed workers are enrolled with INPS. It maintains an extensive dataset containing comprehensive information on Italian employees and firms.
- *Infocamere Knowledge Graph*: Banca d'Italia's KG19 is a knowledge graph designed to structure detailed information from the Italian Business Register. This system facilitates the development of governance-related indicators from 2012 to 2022.

The following sections describe the ICAS company sample and the content of the INPS and Infocamere Knowledge Graph datasets, focusing on the selection process of social- and governance-related indicators.

4.1 ICAS firms

The ICAS dataset covers the years from 2014 to 2023, providing a longitudinal view of corporate financial health and credit performance. The number of firms and the default rate for each year are reported below (Table 1).

Table 1. Firms and default rate by year

Year	Number of firms	Default rate (per cent)
2014	237,108	6.24
2015	229,132	5.35
2016	252,739	4.28
2017	260,056	3.20
2018	259,765	2.93
2019	260,983	2.66
2020	261,324	2.08
2021	246,283	1.69
2022	245,969	1.86
2023	278,179	2.04

The annual default rate in the sample exhibits a general downward trend over the period, reflecting improving creditworthiness within the Italian non-financial corporate sector. The sharp decline in the default rate during 2020 and 2021 was largely driven by the extraordinary policy measures implemented by the Italian government in response to the COVID-19 pandemic.[9] These measures – including a debt moratorium, exemptions from default reporting, and the provision of state-backed loans – resulted in substantial liquidity support, which temporarily stabilised the firms' financial conditions. The increase in defaults observed in 2022 and 2023 partly reflects a rebound following the expiry of the support measures. Considering the distortions introduced by the pandemic-related interventions, the data for 2020 and 2021 have been excluded from the analysis.

Micro-firms represent the largest class in the sample when classified by size, reflecting their widespread presence in the Italian productive system. The default rate is inversely related to firm size: larger companies may benefit from greater financial resilience, diversified revenue streams, and enhanced access to credit (Table 2).

[9] See legislative decrees 'Cura Italia' and 'Sostegni'.

Table 2. Firms and default rate by size

Size	Number of firms	Default rate (per cent)
Micro	1,535,493	3.64
Small	900,961	3.33
Medium	228,163	2.98
Large	51,754	2.23

4.2 Social indicators

The INPS dataset contains information on Italian employees and firms. Each year, INPS provides Banca d'Italia with a dataset comprising aggregate data on all Italian firms with at least one statement of social security contributions for the year. The dataset includes wage, occupation, and age categories for each firm. Based on this source, we identify seven social company-level indicators (Table 3). A description of each indicator is available in Appendix 2. By construction, all the indicators have a strong link with the workforce domain.

Table 3. Social Key Issue Hierarchy

Domain	Key Issues	Involved Indicators	Type
Social	Age	Employees' age entropy	Continuous
	Occupation	Hire rate	Continuous
		Employees' turnover rate	Continuous
		Layoff rate	Continuous
		Fixed-term contract rate	Continuous
	Gender	Female employment rate (compared to the industry)	Categorical
	Wage	Average wage (compared to the industry)	Categorical

The seven reported indicators do not necessarily reflect a deliberate or conscious corporate approach to social issues (Table 3). Rather, they are classified within the social dimension of sustainability reporting frameworks because they describe workforce-related characteristics and employment outcomes at the firm level. Consistent with this taxonomy, the indicators are constructed with reference to the European Sustainability Reporting Standards (ESRS), in particular the S1 category ('Own workforce'), and partly overlap with the social dimension of the GRI Standards. In this paper, these taxonomies are used only as an operational reference for variable selection. The analysis does not seek to infer firms' non-financial performance or social quality from these indicators, but to test their association with default risk.

The INPS dataset distribution of defaulted companies over time mirrors the one shown in Table 1, since the majority of ICAS companies have at least one statement of social security contributions (Table 4).

Table 4. Firms in INPS dataset and default rate by year

Year	Number of firms	Default rate (per cent)
2017	204,589	3.07
2018	206,852	2.88
2019	208,137	2.66
2020	209,258	2.07
2021	201,425	1.69
2022	197,979	1.84
2023	227,572	2.04

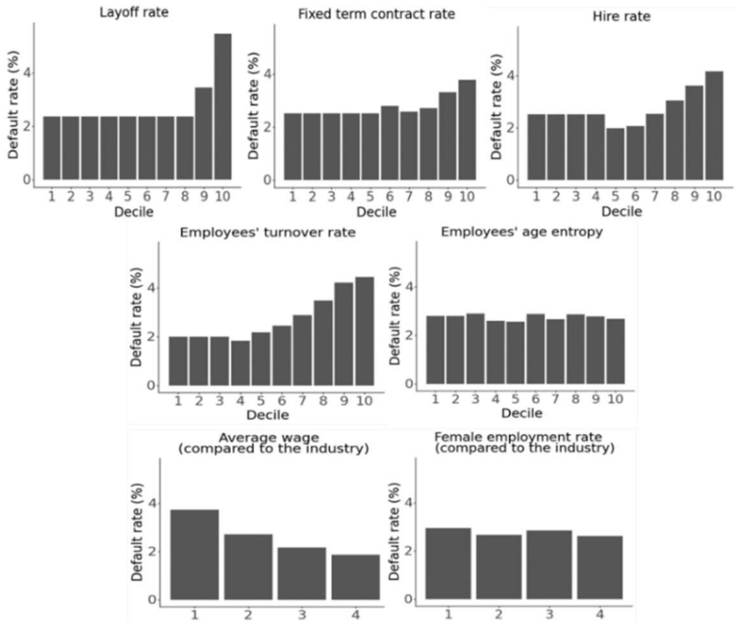
Note: For each year *t* under analysis, INPS data refer to year *t*-2.

For the reasons explained in Section 4.1, we exclude the period 2020-2021 from our analysis.[10] We begin our exploratory analysis with a visual inspection, plotting each social indicator's default rate by decile (or value, depending on the type of the indicator).[11]

In the case of micro firms, when examining the *hire rate*, *employees' turnover rate*, *layoff rate* and *fixed-term contract rate*, we observe a flat pattern for the lower deciles (Figure 2).[12] Except for this evidence, the social-related indicators generally correlate monotonously with the ICAS default rate. There is also a clear association between *average wage* and default, with lower default rates associated with higher wages (compared to the industry). In contrast, *employees' age entropy* and *female employment rate* (compared to the industry) do not show a clear relationship with default.

[10] The resulting dataset is then subjected to the pre-processing procedure, as follows: exclusion of firms with missing data (< 1 per cent); handling of borderline cases (e.g., division by zero); winsorisation of continuous variables at the 95th percentile.
[11] We also provide in Appendix 4 the histograms of the five continuous variables, for each size class.
[12] The flat trend results from a large subset of firms having the same value of the reference indicator. For instance, *layoff rate* is zero for over 80% of firms.

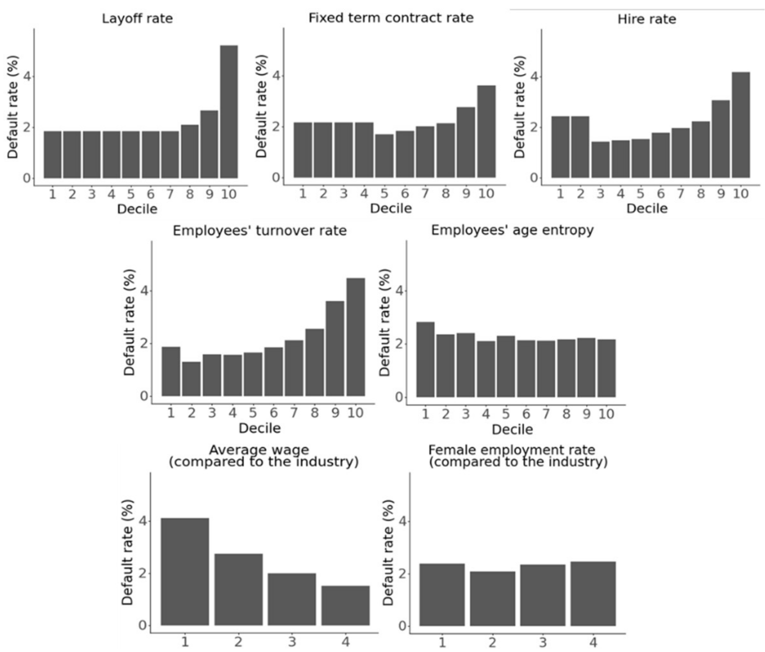
Figure 2. Micro firms' default rate by social-related indicator



Note: *Average wage (compared to industry)* and *female employment rate* plots have only four bars since these indicators are categorical and take four different values (from 1 to 4); see Appendix 2.

The patterns for small firms are similar to those for micro firms (Figure 3). However, for small firms, there are spikes in default rates in the first deciles of *fixed-term contract rate*, *hire rate*, *employees' and turnover rate.*

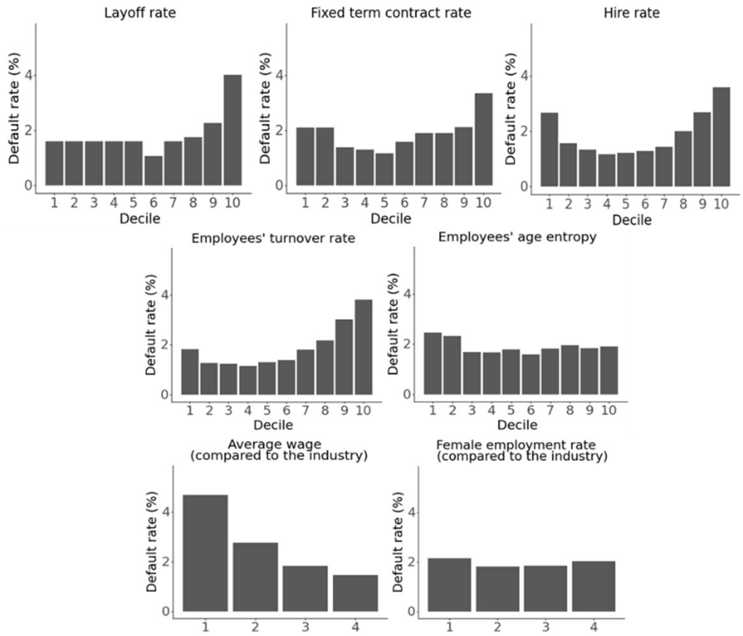
Figure 3. Small firms' default rate by social-related indicator



Note: *Female employment rate and average wage (compared to industry)* plots have only four bars since these indicators are categorical and take four different values (from 1 to 4); see Appendix 2.

The plots for medium- and large-sized companies exhibit broadly U-shaped relationships for *fixed-term contract rate, hire rate,* and *employees' turnover rate* (Figure 4). The spikes observed in Figure 3 are even more evident in this case. The non-monotonic effects are present in the first deciles, where the underlying indicators are equal to zero. Therefore, these effects could be captured by introducing binary variables.

Figure 4. Medium and large firms' default rate by social-related indicator



Note: *Female employment rate and average wage (compared to industry)* plots have only four bars since these indicators are categorical and take four different values (from 1 to 4); see Appendix 2.

To rank the social-related indicators, we estimate univariate logistic regressions where default is the dependent variable; then we compute the AuROC for each variable based on regression outcomes. The evidence is similar across the three groups of firms: *layoff rate*, *hire rate*, *employees' turnover rate*, and *average wage* exhibit good discriminatory power (with average AuROC greater than 55 per cent), while the remaining variables show worse performance.[13]

Since the standard logistic regression assumes a linear relationship between the predictors and the log-odds of the outcome, which implies a monotonic effect of each predictor on the probability of default, we investigate the strength of possible non-linear relationships by running random forest models using all the seven indicators as independent variables (Figure 5, Figure 6, and Figure 7).[14] At this step, assuming a multivariate (nonparametric) approach with the random forest is useful not only for inspecting complex

[13] Detailed tables for the univariate AuROC values of each social indicator are available upon request.
[14] Nonmonotonic or highly non-linear effects require variable transformations within the logistic framework. To estimate the random forest models, a hyperparameter search has been conducted. For technical details, see Appendix 3.

interactions, but also for identifying potential overlaps or redundancies among the indicators before the new risk dimension is tested for inclusion in the baseline S-ICAS.

Figure 5. Random forest feature importance for social-related indicators – Micro firms

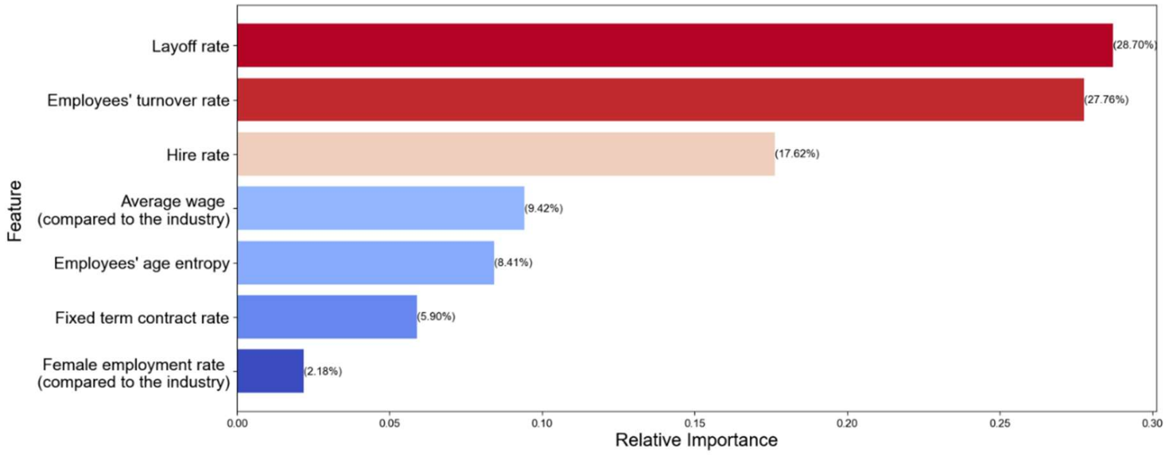


Figure 6. Random forest feature importance for social-related indicators – Small firms

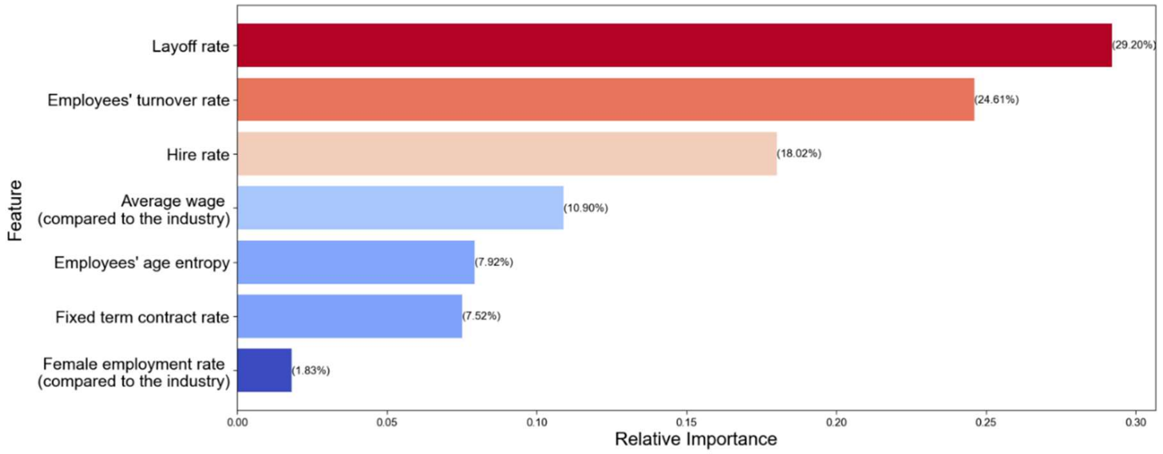
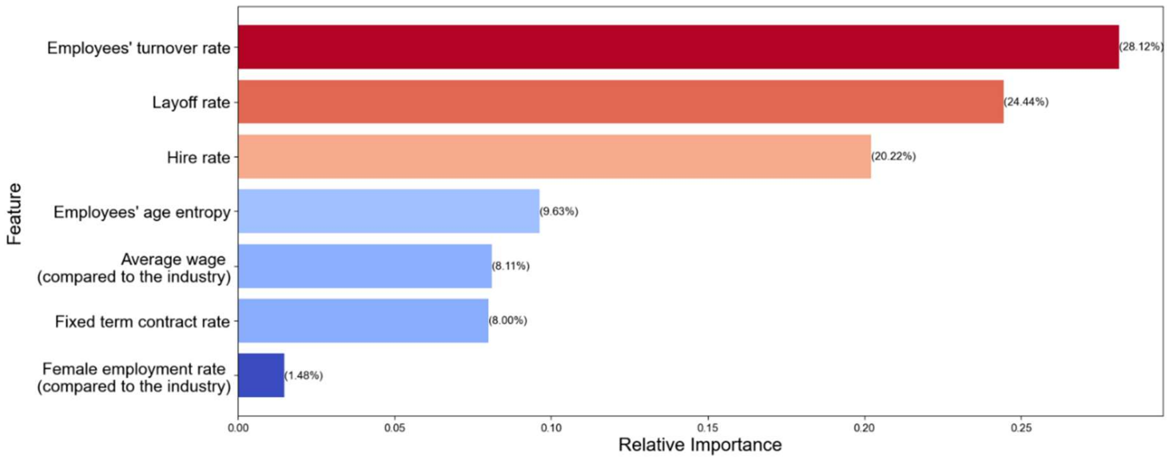


Figure 7. Random forest feature importance for social-related indicators – Medium and large firms



Random forest feature rankings mostly confirm the univariate logistic regression and the visual inspection outcomes. The variables *layoff rate*, *hire rate*, *employees' turnover rate*, and *average wage* are in the highest positions of the rankings, while *female employment rate* stands at the lowest position. Based on this evidence, the graphical inspection, and the regression outcome, we exclude *female employment rate* from further analyses. Besides, the variable *fixed-term contract rate* is deemed less important than *employees' age entropy*. It could depend on the presence of weak non-linear effects or interactions for *employees' age entropy*, not captured by univariate logistic regressions, and the loss of significance of *fixed-term contract rate* due to the presence of the other variables. Since: i) we are interested in estimating a logistic regression model, in line with the structure of the S-ICAS; ii) there are no strong indications of non-linear effects that could hinder the estimation phase; iii) from visual inspection a clearer relation between default rate and *fixed-term contract rate* appears, we exclude *employees' age entropy* from further investigations.

Finally, we perform an empirical correlation analysis for each size class based on the Spearman rank correlation coefficient among the remaining five indicators.[15] The two variables that show a higher correlation (for each size class) are *hire rate* and *employees' turnover rate*. Since the correlation between these variables is approximately 0.9 and the latter shows a better association with the default rate (Figure 2, Figure 3, and Figure 4), we exclude *hire rate* from the social watchlist.[16]

We thus identify a single watchlist of four key social variables for all groups of firms: *layoff rate*, *average wage* (relative to the industry), *employees' turnover rate*, and *fixed-term contract rate*. In the remainder of the paper, these variables are treated as operational proxies with predictive relevance rather than as direct measures of firms' social quality or employment practices.

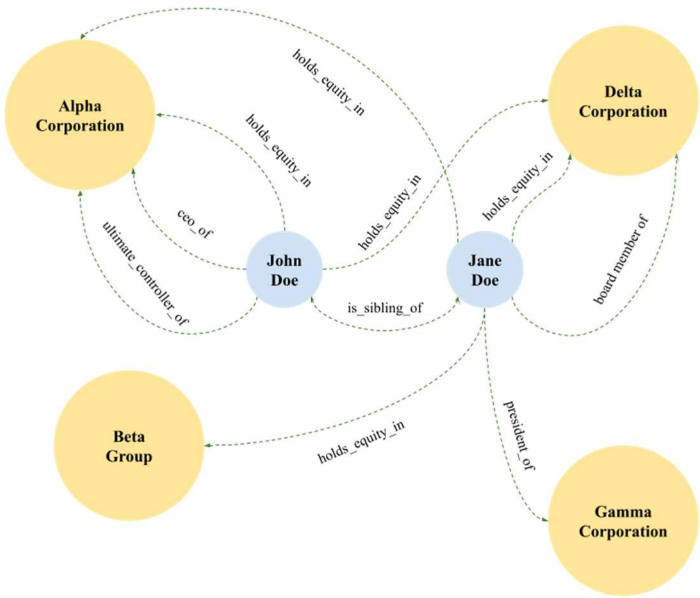
4.3 Governance indicators

Knowledge graphs (KGs) are key tools of artificial intelligence (AI), especially in the deductive AI field. Unlike inductive AI, which primarily relies on machine learning and statistical models to uncover patterns in large datasets, deductive AI uses logical reasoning over structured data to generate new knowledge. A knowledge graph consists of entities representing core elements – individuals, organisations, or abstract concepts – and relationships that capture the inner connections. Figure 8 illustrates an example of a knowledge graph where the relationships between John Doe, Jane Doe, and several corporations are reported. It is represented as a network structure where entities (firm – orange nodes – and individuals – blue node) are interconnected. The *is_sibling_of* relationship connects John Doe and Jane Doe. John Doe holds an executive role as he is linked to Alpha Corporation through the *ceo_of* relationship, signifying

[15] Spearman's rank correlation is a nonparametric measure of the strength and direction of association between two ranked variables. Unlike Pearson's correlation, which assumes linearity, Spearman's method ranks the data and measures monotonic relationships, making it robust to outliers and non-linear associations.
[16] A detailed picture of the correlation analysis is available upon request.

his position as chief executive officer. Beyond these direct associations, John Doe is also identified as the ultimate controller of Alpha Corporation, as denoted by the *ultimate_controller_of* relationship, indicating his overarching influence over the company. Jane Doe is also involved in corporate structures. She is the chairperson of Gamma Corporation. Moreover, she holds shares in both Beta Group and Delta Corporation. Finally, Jane Doe is a board member of Delta Corporation, as shown by the *board_member_of* relationship. By encoding temporal and hierarchical attributes, the graph facilitates advanced queries and analyses, offering a dynamic view of cross-linking governance features and how ownership concentration and role dispersion impact governance trends and corporate stability.

Figure 8. Example of a knowledge graph from the KG19 system



KG19, developed by Banca d'Italia (Bellomarini *et al.*, 2019), represents an enterprise implementation of a knowledge graph. KG19 extracts detailed information from the Italian Business Register and enables us to identify 47 quantitative governance-related indicators for each firm from 2012 to 2022 (Table 5).[17]

[17] The Cypher query language within KG19 enhances its analytical capabilities.

Table 5. Governance Key Issue Hierarchy

Domain	Key Issues	Groups of indicators	Type
Governance	Board	Size, age, stability, instability, gender, location, ownership, diversity	Continuous/ Categorical
	Management	Size, age, stability, instability, gender, location, experience, family relations	Continuous/ Categorical
	Ownership	Concentration, major owner, diversity, change, external influence, depth	Continuous/ Categorical
	Gender	Gender ratio (board and management)	Continuous
	Audit	Audit size, percentage, firm presence, stability, concentration	Continuous/ Categorical
	Turnover	CEO turnover	Categorical
	Stability	Stability of board, management, audit	Continuous
	Instability	Instability of board, management, audit	Continuous
	Diversity	Age, gender, location (board and management)	Categorical

Note: Appendix 2 provides a detailed description of each single indicator.

As in the case of the social indicators, these variables are classified as governance-related because they capture features commonly associated with governance structures in reporting frameworks and external data taxonomies. In this paper, they are used as structural proxies for default prediction, not as direct measures of governance quality.

Henceforth, we examine the relationship between each indicator and corporate default. As discussed in Section 4.1, data from 2020 to 2021 have been removed from the analysis.[18] Given the large set of governance-related indicators, only graphs related to the indicators showing the clearest association with the default rate are reported.

Micro- and small-sized companies show similar patterns (inverse relation) for *internal-external ratio* and *shareholders with at least 5% ownership* (Figure 9 and Figure 10). In the case of micro firms, the same picture seems valid for *board size* and *board age diversity* – an index of diversity in the age distribution of board members. *Herfindahl-Hirschman index* and *board ownership concentration* show a U-shaped pattern for both groups of firms.

[18] The resulting dataset was subjected to the pre-processing schema described in note 10, Section 4.2.

Figure 9. Micro firms' default rate and governance-related indicators

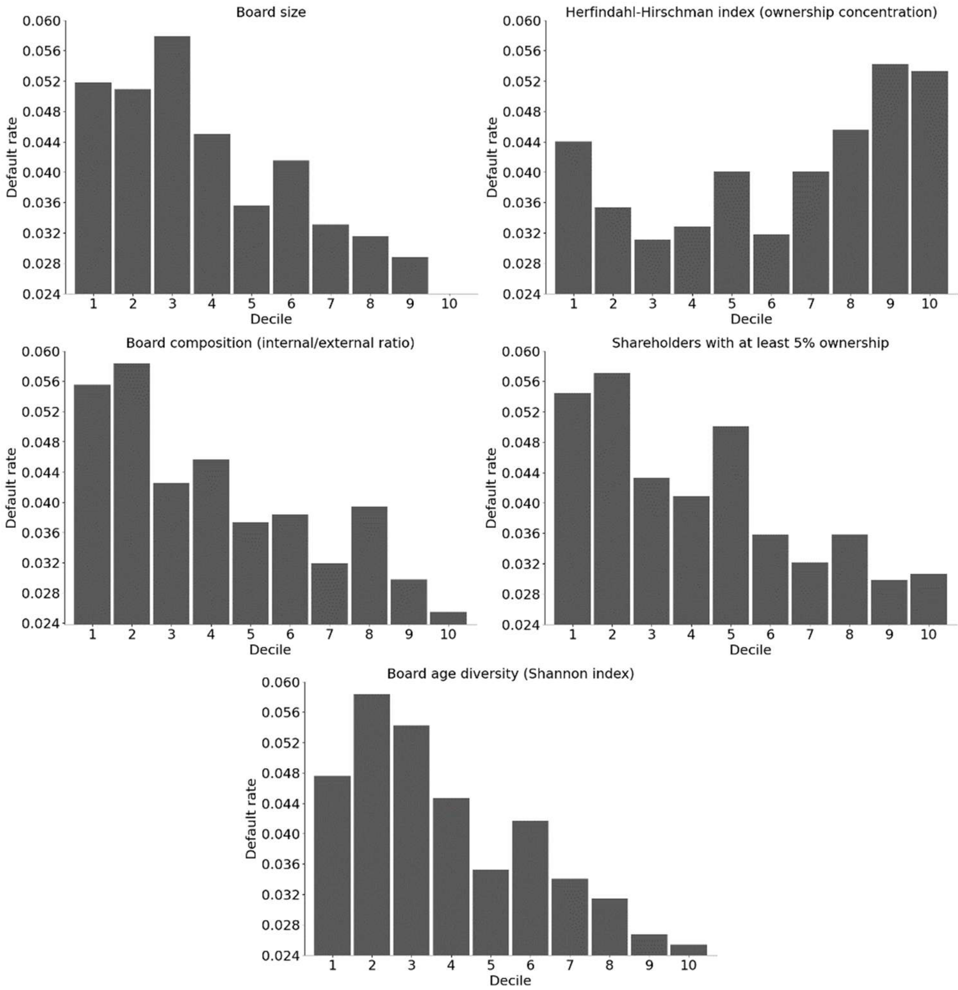
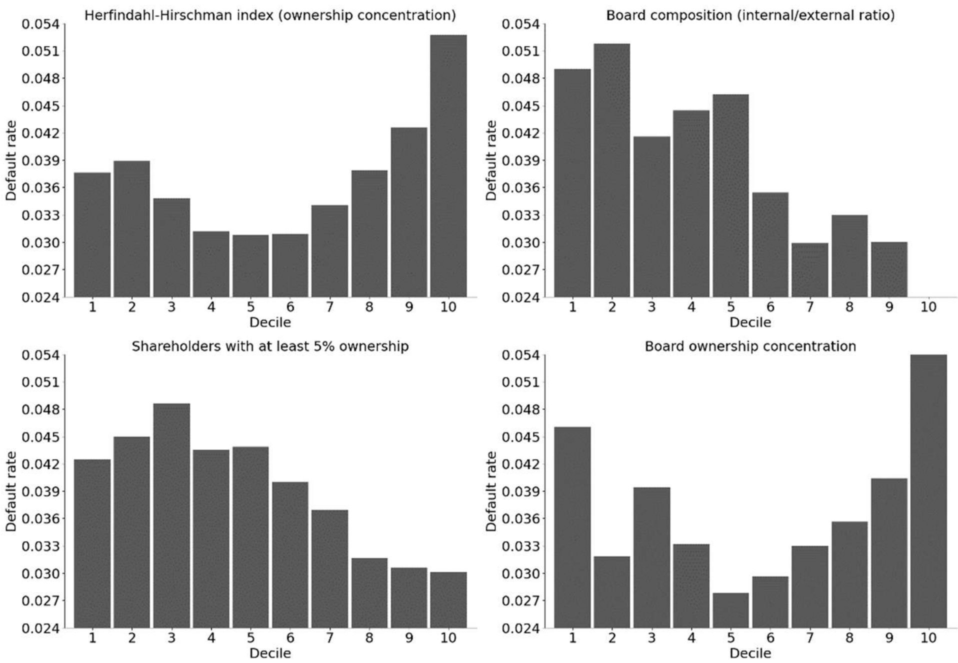
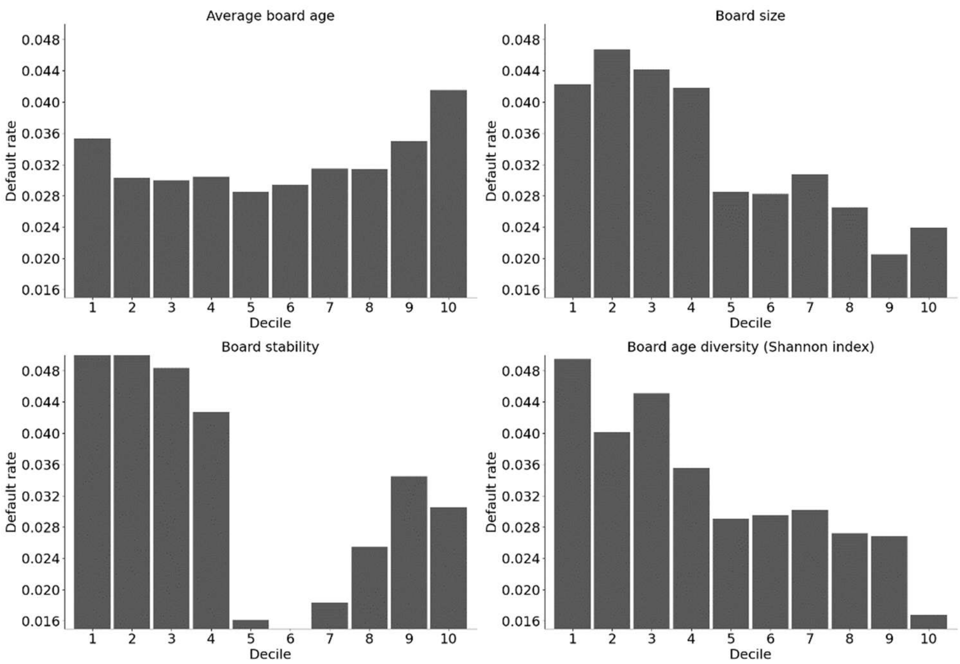


Figure 10. Small firms' default rate and governance-related indicators



For medium- and large-sized companies, we observe an increase in default rate in the highest deciles of *average board age* (Figure 11). The *board size* and *board age diversity* patterns illustrate an inverse relationship with default rate, like what has been detected for micro firms. Finally, *board stability* exhibits a distinct pattern where firms with less stable boards (lower deciles) show significantly higher default rates than those with more stable governance structures.

Figure 11. Medium and large firms' default rate and governance-related indicators



To rank the indicators, we first estimate univariate logistic regressions and compute the AuROC for each variable, as in Section 4.2.[19] Then, to account for the possible non-linear and non-monotonic relationships, we replicate the random forest feature importance selection analysis.

Board size and *board age diversity* are confirmed among the strongest predictors across firm size classes. *Internal-external ratio* shows a high discriminatory ability for micro and small firms, reaching an average AuROC equal to 55 per cent. For medium and large firms, *audit size* provides valuable insights (average AuROC equal to 60 per cent). In the present setting, these variables are treated as governance-related structural proxies with predictive relevance, rather than as direct measures of governance quality.

The results of the random forest feature importance selection are presented in Figure 12, Figure 13, and Figure 14 for micro, small, and medium and large enterprises, respectively.

[19] Detailed figures for the univariate AuROC values of each governance variable are available upon request.

Figure 12. Random forest feature importance for governance-related indicators – Micro firms

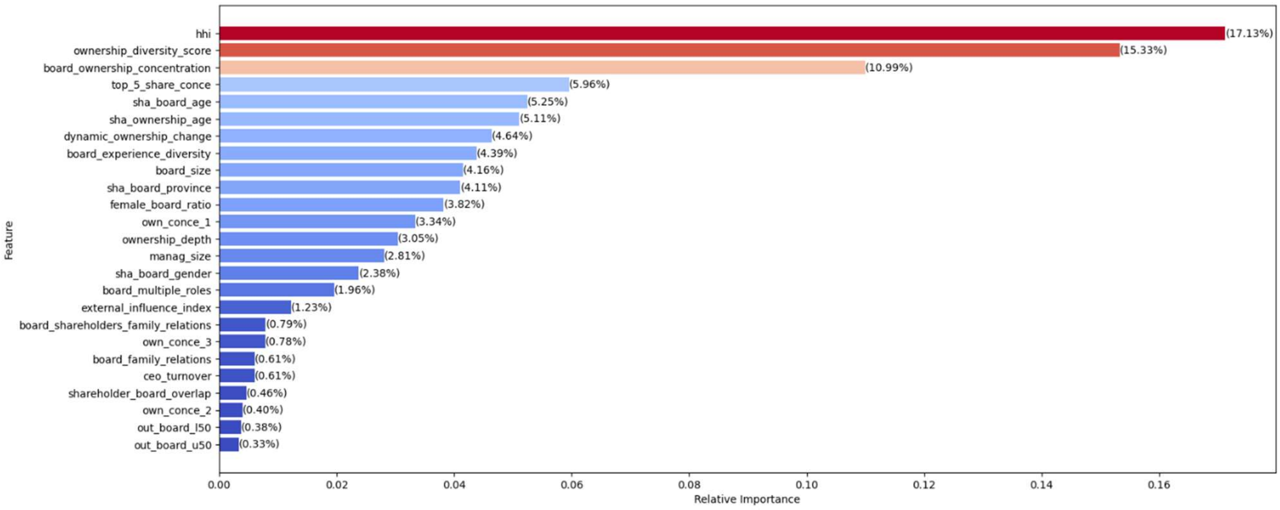


Figure 13. Random forest feature importance for governance-related indicators – Small firms

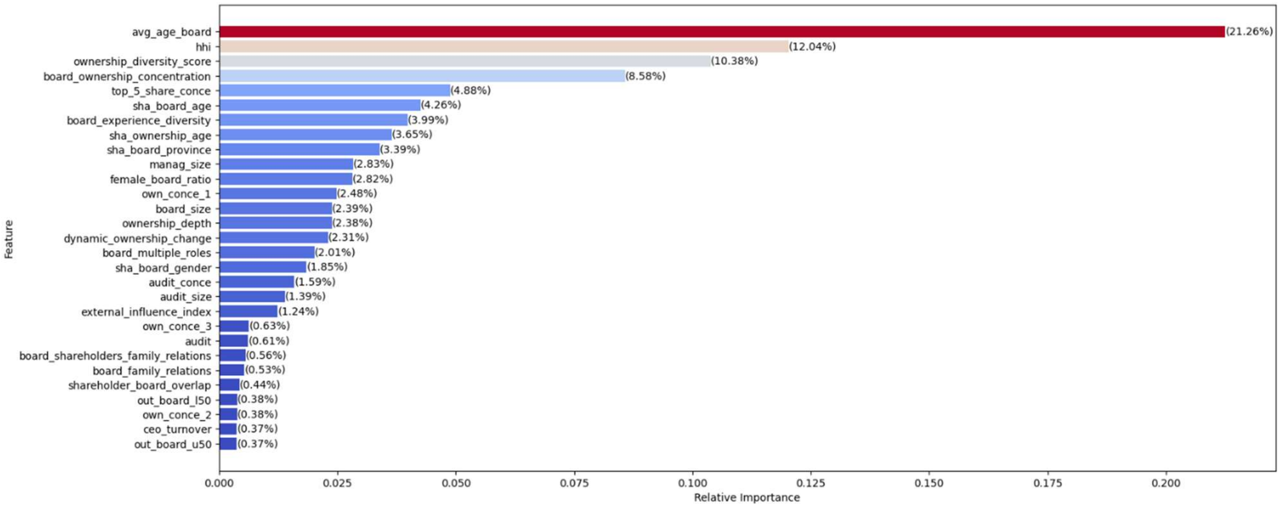
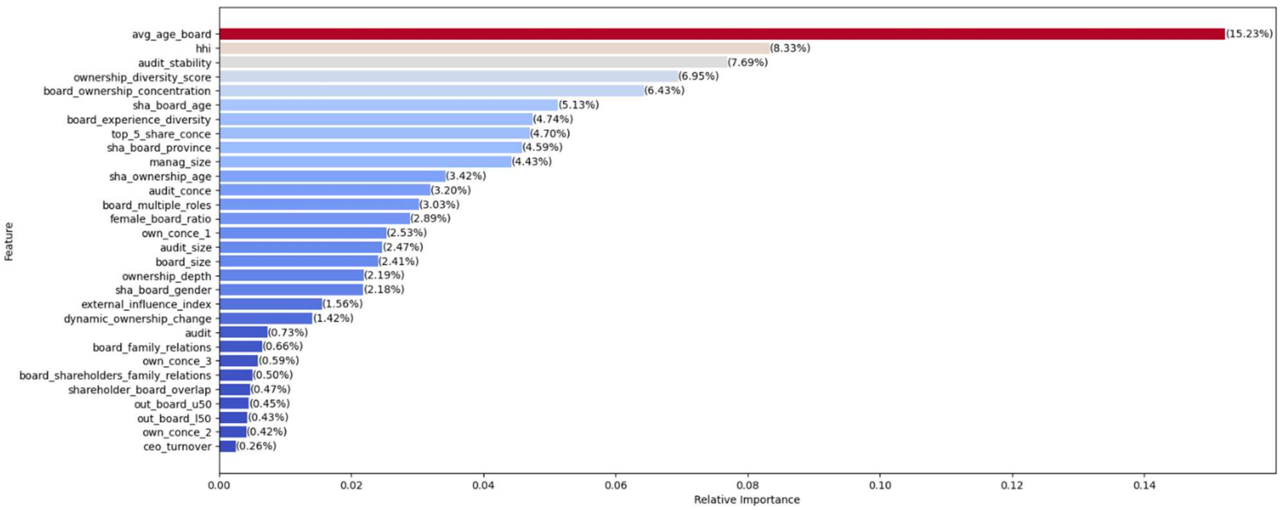


Figure 14. Random forest feature importance for governance-related indicators – Medium and large firms



Since we handle a large set of indicators, we filter the watchlists initially obtained from univariate logistic regressions with the random forest feature ranking. To do so, we apply a cumulative variable importance threshold of 90 per cent: only the most influential variables that reach a relative importance of 90 per cent of the total are retained for further analysis.[20] Since the governance score will be obtained through a logistic regression, to handle the variables with the most prominent non-monotonic relation with the default rate,[21] we apply a discretisation process using decision trees.[22] This method transforms each non-monotonic variable into a new monotonic variable by assigning a rank based on the predicted default probabilities derived from a decision tree.[23]

We finally provide distinct watchlists for micro, small, and medium and large firms.

The key variables for the micro firms include *board experience diversity* (non-monotonic), *board ownership concentration* (non-monotonic), *board size* (monotonic), *Herfindahl-Hirschman index* (non-monotonic), *ownership concentration 1* (non-monotonic), and *board age diversity* (monotonic). Additionally, factors such as *Shannon board province* (non-monotonic), and *ownership age diversity* (non-monotonic) have been included. Other variables such as *ownership diversity score* (non-monotonic), *shareholders with at least 5% ownership* (non-monotonic), *dynamic ownership change* (non-monotonic), *female board ratio* (non-monotonic), and *ownership depth* (non-monotonic) are also part of the selected features for this firm size category. We thus exclude *internal-external ratio* from the final list due to the poor performance in the multivariate setting.

In the case of small firms, the watchlist includes variables such as *average board age* (non-monotonic), *board experience diversity* (monotonic), *board multiple roles* (non-monotonic), and *board size* (monotonic). Governance aspects such as *Herfindahl-Hirschman index* (non-monotonic), *ownership concentration 1* (monotonic), and *ownership depth* (monotonic) are also included. Additional factors are *ownership diversity score* (non-monotonic), *board age diversity* (monotonic), *Shannon board province* (monotonic), *ownership age diversity* (monotonic), *top 5 share concentration* (non-monotonic), *management size* (non-monotonic), *female board ratio* (non-monotonic), and *dynamic ownership change* (non-monotonic). As before, we exclude *internal-external ratio* from the final list.

[20] The existing literature supports the rationale behind choosing a 90 per cent threshold. For example, Guyon and Elisseeff (2003) argue that feature selection methods should balance the trade-off between model complexity and performance by focusing on the most informative variables, often captured within the top 80 per cent-90 per cent of cumulative importance in ensemble methods like random forests. Similarly, in their review of feature selection techniques, Bolón-Canedo *et al.* (2015) show that selecting variables based on cumulative feature importance can improve the model's interpretability without sacrificing predictive accuracy.

[21] As emerged from visual inspection.

[22] Due to the limited number of nonmonotonic relations detected with the social indicators, we restrict the use of the discretisation process to the governance indicators.

[23] Specifically, a decision tree classifier is trained using the training dataset for each variable, where the default status serves as target variable. The tree's predicted PDs are then used to rank the data, ensuring a monotonic relationship with the default rate.

For medium- and large-sized firms, the watchlist includes *average board age* (non-monotonic), *board experience diversity* (monotonic), *board multiple roles* (non-monotonic), *board ownership concentration* (non-monotonic), and *board size* (monotonic). *Herfindahl-Hirschman index* (non-monotonic), *board age diversity* (monotonic), and *Shannon board province* (monotonic) are also considered. Audit stability, audit concentration, and audit size are included too (all non-monotonic cases). Furthermore, *ownership diversity score* (non-monotonic), *top 5 share concentration* (non-monotonic), *management size* (non-monotonic), *ownership age diversity* (monotonic), *female board ratio* (non-monotonic), and *ownership concentration 1* (monotonic) are added to the list.

We build our contribution on the existing statistical framework, assessing the incremental value of S and G dimensions without altering the core econometric design of S-ICAS. Since governance-related indicators are typically associated with financial performance outcomes, a multicollinearity analysis is finally performed between the watchlists and indicators from the financial statements involved in the baseline S-ICAS. We use Spearman's rank correlation and the Maximal Information Coefficient (MIC).[24] However, correlation analysis does not suggest the presence of multicollinearity and, consequently, any exclusion.

5 The inclusion of S and G indicators in S-ICAS

Based on the previous section's exploratory analysis, we estimate separate PD models using social- and governance-related watchlists. These models should be interpreted as predictive modules for default events based on variables operationally classified under the S and G domains, not as stand-alone measures of firms' social or governance quality. Second, we map the governance and social PDs into scores (via the logit function) and merge them with the integration module of S-ICAS, resulting in an augmented statistical model with a broader set of risk dimensions. In the augmented version, the S-ICAS PD results from the combination of four components: the financial statement score, the credit behaviour score, the social score, and the governance score. In the final step, we perform a comparative analysis of the performance of the augmented PDs and baseline PDs based on the AuROC and binomial tests.

To properly handle the separate estimation of the governance and social PDs, we employ a stepwise backward regression approach to estimate logistic regression models for each firm size class with the smallest set of indicators. Stepwise backward regression starts with the watchlists outlined in Sections 4.2

[24] The Maximal Information Coefficient (MIC), part of the Maximal Information-based Nonparametric Exploration (MINE) family of statistics, identifies complex relationships between variables. MIC is particularly effective in detecting linear and non-linear associations, providing a measure of association that remains high across a range of functional relationships.

and 4.3.[25] This approach allows for the identification of the most important variables, while reducing overfitting.

5.1 The social PD

For small firms, we observe a monotonic relationship between default rate and both *employees' turnover rate* and *fixed-term contract rate*, except in the first deciles, where these variables are equal to zero (Figure 3). To adequately capture these features, we introduce dummy variables for both predictors, with values set at one where the corresponding indicator is zero. However, since the correlation between *fixed-term contract rate* and its dummy is below -0.7 and the spike in the first deciles is not too high, we exclude it from model estimation. Similarly, for medium and large firms, we consider dummy variables for *employees' turnover rate*, *fixed-term contract rate*, and *layoff rate*. However, we exclude the dummies related to *fixed-term contract rate* and *layoff rate* from the estimation due to the very high correlation with their corresponding indicators.

For robustness, we employ an expanding window approach. Starting from 2017 and extending to 2022,[26] the logistic regression is iteratively trained using data from 2017 up to year *t* and subsequently tested on data from year *t+1*. For the sake of simplicity, regression results are reported for the largest window only (Table 6).[27]

Table 6. Regression results by firm size - social

a) Micro firms

	coef	***std err***	***P>|z|***
const	-3.50	0.03	0.00
layoff rate	2.04	0.07	0.00
employees' turnover rate	*0.13*	*0.01*	*0.00*
average wage (compared to industry)	*-0.13*	*0.01*	*0.00*

b) Small firms

	coef	***std err***	***P>|z|***
const	-3.44	0.04	0.00
layoff rate	3.09	0.12	0.00
employees' turnover rate	0.14	0.01	0.00
average wage (compared to industry)	-0.19	0.01	0.00
dummy employees' turnover rate	0.12	0.04	0.00

[25] Although stepwise backward regression can technically be applied to all variables, in the present study, it is restricted to those showing a more promising (and possibly monotonic) relationship with the default rate, as discussed in the exploratory analysis.

[26] As mentioned in Section 4.1, we exclude data for 2020 and 2021. The sample evidence starts in 2017 (Table 4).

[27] For clarity, estimated coefficients, standard errors, and *p*-values in tables with values reported as 0.00 do not indicate zero values. Instead, they represent values less than the assumed reporting threshold of 0.01.

c) Medium and large firms

	coef	*std err*	*P>|z|*
const	-3.38	0.10	0.00
layoff rate	3.51	0.28	0.00
fixed-term contract rate	1.09	0.24	0.00
average wage (compared to the industry)	-0.23	0.03	0.00
dummy employees' turnover rate	0.43	0.10	0.00

Across all firm size categories, *layoff rate*, *employees' turnover rate*, *fixed-term contract rate*, and *average wage* display statistically significant associations with default. In line with the scope of the paper, these coefficients are interpreted in predictive rather than causal or value-based terms. The variables are treated as operational workforce-related proxies with predictive relevance for default risk. Their role in the model is purely statistical, as they contribute to discrimination and prediction regardless of whether they reflect employment policies, firm conditions, or other unobserved factors. Overall, the estimated coefficients tend to increase in magnitude with firm size.

Finally, the social PD's out-of-sample performance in AuROC is moderate, with AuROC values ranging from 62 to 65 per cent, suggesting better performance in the case of small firms (Table 7).

Table 7. Social PD – Out-of-sample AuROC

(percentage values)

Out-of-sample year	Micro	Small	Medium-large
2022	62	65	63
2023	63	65	63

5.2 The governance PD

We replicate the expanding window strategy for model testing described in Section 5.1. Starting in 2012 and progressing through 2022, the logistic regression is successively trained on data from 2012 to year *t* and evaluated on data from year *t+1*. As in Section 5.1, regression results are reported for the largest window for the sake of simplicity (Table 8).[28]

[28] In Section 4.1, we exclude data related to 2020 and 2021.

Table 8. Regression results by firm size - governance

a) Micro firms

	coef	*std err*	*P>|z|*
const	-2.88	0.01	0.00
dynamic_ownership_change	0.25	0.03	0.00
hhi	0.13	0.01	0.00
board_size	-0.21	0.01	0.00
sha_board_age	-0.23	0.02	0.00

b) Small firms

	coef	*std err*	*P>|z|*
const	-3.02	0.04	0.00
board_size	-0.15	0.01	0.00
sha_ownership_age	-0.12	0.01	0.00
board_experience_diversity	-0.15	0.02	0.00
sha_board_age	-0.26	0.02	0.00
hhi	0.07	0.02	0.00
avg_age_board	2.96E-03	6.07E-04	0.00

c) Medium and large firms

	coef	*std err*	*P>|z|*
const	-2.81	0.10	0.00
ownership_diversity_score	0.03	0.03	0.01
sha_board_age	-0.30	0.02	0.00
hhi	1.44E-03	3.70E-04	0.00
avg_age_board	2.63E-03	7.94E-04	0.00

Despite differences across firm sizes, some patterns stand out. However, these results should be interpreted in predictive rather than normative or causal terms. In this paper, governance-related variables such as *board size* and *board age diversity* are used as structural proxies for discrimination between defaulting and non-defaulting firms, rather than as statements about governance quality. From a purely statistical perspective, *board size* shows a negative coefficient across micro and small firms, *board age diversity* instead exhibits a negative coefficient for all size classes. These results confirm their predictive relevance within the model, without implying any evaluative ranking of governance structures.

The governance PD’s out-of-sample performance in terms of AuROC is relatively modest, with AuROC values between 58 and 65 per cent (Table 9). These results indicate a level of discriminatory power in line with or slightly smaller than the 'S' outcome. Compared with the 'S', we observe better results for medium and large companies in 2023.

Table 9. Governance PD – Out-of-sample AuROC

(percentage values)

Out-of-sample year	AuROC (Micro)	AuROC (Small)	AuROC (Medium/Large)
2022	61	64	58
2023	61	65	65

5.3 The augmented S-ICAS PD

Sections 5.1 and 5.2 introduce the estimated governance and social PDs separately. Each of the two PDs provides some discriminatory power. To disentangle the contribution of the two new dimensions, we introduce the social and governance scores into the baseline S-ICAS (Equation 1). Specifically, we perform logistic regressions using financial, credit behaviour, social, and governance scores as predictors of default by firm size. Each regression is estimated using a backward elimination procedure. We employ the expanding window approach already adopted in Sections 5.1 and 5.2, except for large firms for which we adopt a single window.[29] We normalise the four scores to compare regression coefficients and assess the strength of each score in determining the final PD. Due to the limited evidence affecting the social-related indicators, 2017 is the first year in the estimation process. Unlike regressions estimated in Sections 5.1 and 5.2, where medium and large firms are grouped, in the integration step, large firms are separately analysed.[30]

5.3.1 Micro firms

For micro firms, both the social and governance scores are significant in each estimation window (Table 10 reports the largest window-based results).

Table 10. Logistic regression results for Micro firms – integration module

	coef	*std err*	*P>|z|*
const	-4.56	0.02	0.00
financial score	0.62	0.01	0.00
credit score	1.09	0.01	0.00
social score	0.20	0.01	0.00
governance score	0.23	0.01	0.00

[29] The dataset's total number of large companies' defaults is extremely low (below 150), which prevents the adoption of the expanding window approach.
[30] The integration module of the standard S-ICAS has separate models for medium and large firms, respectively.

The credit behaviour score has the highest coefficient, followed by the financial score, while the social and governance scores remain statistically significant but smaller in magnitude. This result indicates that the additional S- and G-related modules provide incremental information, although their contribution remains secondary relative to the baseline dimensions.

Next, we compare the AuROC of the augmented S-ICAS with the baseline model obtained by integrating only the financial and credit behaviour scores (Table 11).[31]

Table 11. Comparison of the integrated PD AuROC – Micro firms

(percentage values)

Out-of-sample year	Model with governance and social scores	Model without governance and social scores
2022	84.3	83.6
2023	84.9	84.1

The model with the social and governance scores has a larger AuROC. The increments in AuROC are statistically significant[32] and range between 0.7 and 0.8 percentage points.

5.3.2 Small firms

We find results similar to those of micro firms. The social and governance scores are significant in each estimation window (Table 12 reports the largest window-based results).

Table 12. Logistic regression results for Small firms – integration module

	coef	*std err*	*P>|z|*
const	-4.92	0.02	0.00
financial score	0.73	0.02	0.00
credit score	1.07	0.01	0.00
social score	0.21	0.01	0.00
governance score	0.23	0.02	0.00

The credit and financial scores remain the main drivers of the integrated model, while the social and governance scores provide an additional statistically significant but smaller contribution.

Next, we compare the out-of-sample AuROC of the augmented S-ICAS with the baseline model (Table 13).

[31] We estimate the baseline model on the same dataset as the one employed to estimate the governance and social scores.
[32] The bootstrap confidence intervals of the AuROC increments are available upon request.

Table 13. Comparison of the integrated PD AuROC – Small firms

(percentage values)

Out-of-sample year	Model with governance and social scores	Model without governance and social scores
2022	86.9	86.4
2023	87.4	86.5

The model estimated with the social and governance scores has a larger AuROC. The increments are all statistically significant and range between 0.5 and 0.9 percentage points.

5.3.3 Medium firms

The social score is statistically significant for medium-sized firms in all estimation windows. Instead, the governance score is significant only in the last two windows. For this reason, we choose to exclude the governance score from the augmented S-ICAS (Table 14 reports the largest window-based results).

Table 14. Logistic regression results for Medium firms – integration module

	coef	***std err***	***P>|z|***
Const	-4.95	0.06	0.00
financial score	0.90	0.06	0.00
credit score	0.87	0.03	0.00
social score	0.18	0.03	0.00

The financial score has the largest coefficient among the predictors, followed by the credit score and the social score.

Next, we compare the out-of-sample AuROC of the augmented S-ICAS with the baseline model (Table 15).

Table 15. Comparison of the integrated PD AuROC – Medium firms

(percentage values)

Out-of-sample year	Model with social score	Model without social score
2022	85.2	84.6
2023	88.3	87.5

The model estimated using the social score has a larger AuROC. The increments are statistically significant and range between 0.6 and 0.8 percentage points.

5.3.4 Large firms

Owing to the few occurrences of default for large firms, we estimate the model only once using the largest window available from 2017 to 2022.

Table 16. Logistic regression results for Large firms – integration module

	coef	*std err*	*P>|z|*
Const	-5.25	0.16	0.00
financial score	1.09	0.14	0.00
credit score	0.90	0.08	0.00
social score	0.18	0.07	0.01

The financial score is dominant in explaining default risk, followed by the credit score, while the social score remains statistically significant with a smaller magnitude (Table 16).

The model estimated using the social score has a larger AuROC than the baseline model, with an increment of 0.4 percentage points (Table 17).

Table 17. Comparison of the integrated PD AuROC – Large firms

(percentage values)

Out-of-sample year	Model with social score	Model without social score
2023	90.8	90.4

5.3.5 Predictive power

The augmented S-ICAS exhibits greater discriminatory power than the baseline model. Discriminatory power refers to a model's ability to rank firms from the highest-risk to the lowest-risk profiles. This characteristic is distinct from predictive capacity (or calibration), which measures the ability to assign the actual PD to a group of homogeneous borrowers.

To evaluate the predictive power of the augmented S-ICAS, we employ a procedure commonly used in the Eurosystem (Narizzano *et al.*, 2024).[33] From best to worst, we categorise companies into rating classes called Credit Quality Steps (CQS).[34] Subsequently, we conduct a two-tailed binomial test within each rating class to establish whether the observed default rate aligns with the out-of-sample insolvency probability. The null hypothesis is that the number of defaults within each rating class is consistent with

[33] The binomial test used by the Eurosystem to evaluate ICASs' predictive power is slightly different from the one adopted in this work. It is a one-tailed binomial test in which the null hypothesis is that the number of actual defaults within each rating class is smaller than the number of expected defaults. In this work, we test a stronger hypothesis using a two-tailed binomial test: the actual defaults within each rating class are consistent with the expected number. According to the Eurosystem monitoring framework, the analysis is carried out on firms irrespective of their size.

[34] Credit Quality Steps form the harmonised rating scale of the Eurosystem.

the expected one, represented by the number of companies in the rating class multiplied by the average PD.[35] The results are reported below (Table 18).

Table 18. Comparison of predictive power

(percentage values)

Rating class	2022 Baseline model	2022 Augmented model	2023 Baseline model	2023 Augmented model
CQS 1&2	0.04	0.04	0.10	0.07
CQS 3	0.17	0.19	0.22	0.23
CQS 4	0.82	0.74	0.76	0.69
CQS 5	1.38	1.21	1.48	1.36
CQS 6	2.20	2.31	2.46	2.42
CQS 7	4.50	4.28	5.25	4.98

Each cell displays the observed default rate for the corresponding rating class. The colour indicates the outcome of the test: green stands for a successful test, orange suggests the model is overly cautious (overestimating risk), and red indicates the model is not cautious enough (underestimating risk). Overall, the augmented S-ICAS provides better precision than the baseline model, especially in the first risk classes, where companies have a low credit risk.

6 Conclusions

This study contributes to the empirical literature on credit risk assessment by examining whether social- and governance-related indicators provide incremental predictive content for firms' probability of default within the S-ICAS framework currently employed at the Banca d'Italia. The social- and governance-related indicators are constructed with reference to metrics commonly classified within official sustainability reporting standards. Overall, our contribution is about predictive usefulness in default risk assessment, not about ESG evaluation. The same raw indicators can be informative for credit risk even when they are not interpreted as measures of firms' social or governance quality. A separate analysis of this issue is deferred to future research. Besides, the objective of this work is not to redesign the ICAS statistical model. Rather, we adopt its existing structure as a benchmark against which to evaluate whether the integration of social- and governance-related modules can enhance its predictive performance. This approach allows us to isolate the marginal contribution of these additional dimensions without altering the original model specification or its operational implementation. To this end, we employ random forest feature-importance rankings to assess the relevance of the indicators across firm size classes, thereby providing an ordered hierarchy of social and governance-related indicators associated with default risk.

[35] The test is passed if the p-value is greater than 0.05.

In addition, decision-tree-based discretisation of non-monotonic variables enables the identification of complex relationships that may not be adequately captured by a standard logistic framework.

The augmented model shows stronger discriminatory power than the baseline statistical model and provides better predictive performance, especially in the first risk classes populated by firms with lower credit risk. Our findings suggest that the inclusion of social- and governance-related modules may enhance the discriminatory and predictive power of the S-ICAS model, particularly for micro and small firms, although the overall contribution of these variables seems moderate. For medium and large firms, the governance-related variables do not provide a significant improvement, and the social-related variables offer marginal gains.

G-related indicators for large firms will continue to be considered in the second building block of ICAS, namely the expert assessment performed by credit analysts. This involves the analysis of qualitative information on the transparency of governing bodies, the reputational standing of top managers, or the existence of legal disputes. These elements, obtained from various sources, are indeed difficult to embed in the statistical model. In the coming years, the statistical ratings obtained from models such as S-ICAS will be directly employed in the Eurosystem Credit Assessment Framework for the credit risk evaluation of micro, small and medium-sized firms, without the need for the expert assessment. The integration of the S- and G-related indicators into the statistical framework would involve a careful estimation and validation of the added value of these indicators, with a possible distinction according to the firms' business sector. More generally, the results support the view that variables drawn from S- and G-related taxonomies may be useful in statistical default modelling even when they are not interpreted as measures of firms' ESG quality. This distinction is central to the interpretation of the present findings.

References

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance, 23(4), 589-609.

Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETA^{TM} analysis A new model to identify bankruptcy risk of corporations. Journal of Banking & Finance, 1(1), 29-54.

Altman, E. I., Balzano, M., Giannozzi, A., & Srhoj, S. (2023). Revisiting SME default predictors: the omega score. Journal of Small Business Management, 61(6), 2383-2417.

Aulia, A. S., Wijaya, C. (2020). The effect of corporate governance quality on default risk of companies listed in Indonesia stock exchange (IDX) period 2013-2017. American International Journal of Business Management, 3(5), 36-44.

Barboza, F., Kimura, H., Altman, E. (2017). Machine learning models and bankruptcy prediction. Expert Systems with Applications, 83, 405-417.

Bellomarini, L., Fakhoury, D., Gottlob, G., Sallinger, E., (2019). Knowledge Graphs and Enterprise AI: The Promise of an Enabling Technology 2019 IEEE 35th International Conference on Data Engineering (ICDE), Conference proceedings pages 26-37.

Bhattacharya, S., Sharma, D. (2019). Do environment, social and governance performance impact credit ratings: a study from India. International Journal of Ethics and Systems, 35(3), 466-484.

Bolón-Canedo, V., Sánchez-Maroño, N., Alonso-Betanzos, A. (2015). Recent advances and emerging challenges of feature selection in the context of big data. Knowledge-based systems, 86, 33-45.

Bonacorsi, L., Cerasi, V., Galfrascoli, P., Manera, M. (2024). ESG factors and firms' credit risk. Journal of Climate Finance, 6, 100032.

Chawla, N.V. (2009). Data Mining for Imbalanced Datasets: An Overview. In: Maimon O., Rokach L. (eds.) Data mining and knowledge discovery handbook, 875-886, Springer, Boston, MA.

Ciampi, F., Gordini, N. (2013). Small enterprise default prediction modeling through artificial neural networks: an empirical analysis of Italian small enterprises. Journal of Small Business Management, 51(1), 23-45.

Ciampi, F. (2015). Corporate governance characteristics and default prediction modeling for small enterprises. An empirical analysis of Italian firms. Journal of Business Research, 68(5), 1012-1025.

Cultrera, L. (2020). Evaluation of bankruptcy prevention tools: evidences from COSME programme. Economics Bulletin, 40(2), 978-988.

Di Virgilio, S., Iannamorelli, A., Monterisi, F., Narizzano, S. (2025). The use of Banca d'Italia's credit assessment system for Italian non-financial firms within the Eurosystem's collateral framework. MISP, n. 60, June 2025. https://www.bancaditalia.it/pubblicazioni/mercati-infrastrutture-e-sistemi-di-pagamento/approfondimenti/2025-060/N.60-MISP.pdf

Fawcett, T. (2004). ROC graphs: Notes and practical considerations for researchers. Machine learning, 31(1), 1-38.

García, C. J., Herrero, B. (2021). Female directors, capital structure, and financial distress. Journal of Business Research, 136, 592-601.

Giannozzi A., Altman E.I., Roggi O., Sabato G. (2013). Building SME Rating: is it Necessary for Lenders to Monitor Financial Statements of the Borrowers?, Bancaria Editrice, 10, 54-71.

Giovannelli F., Iannamorelli A., Levy A., Orlandi M., (2023). The Bank of Italy's In-House Credit Assessment System for Non-financial Firms, in: Scalia, A. (eds.) Financial Risk Management and Climate Change Risk. Contributions to Finance and Accounting, Springer, 107-137.

Guyon, I., Elisseeff, A. (2003). An introduction to variable and feature selection. Journal of Machine Learning Research, 3(Mar), 1157-1182.

Hosmer Jr, D. W., Lemeshow, S., Sturdivant, R. X. (2013). Applied logistic regression. John Wiley & Sons.

James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). An introduction to statistical learning: With applications in python. Springer Nature.

Narizzano, S., Orlandi, M., Scalia, A. (2024). The Bank of Italy's statistical model for the credit assessment of non-financial firms. MISP, n. 53, October 2024. https://www.bancaditalia.it/pubblicazioni/mercati-infrastrutture-e-sistemi-di-pagamento/approfondimenti/2024-053/N.53-MISP.pdf.

Norden, L., Weber, M. (2010). Credit line usage, checking account activity, and default risk of bank borrowers. The Review of Financial Studies, 23(10), 3665-3699.

Paraschiv, F., Schmid, M., Wahlstrøm, R. R. (2024). Bankruptcy Prediction of Privately Held SMEs Using Feature Selection Methods. Available at SSRN: https://ssrn.com/abstract=3911490 or http://dx.doi.org/10.2139/ssrn.3911490

Tian, S., Yu, Y., Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. Journal of Banking & Finance, 52, 89-100.

Vivel-Búa, M., Lado-Sestayo, R., Martínez-Salgueiro, A., Díaz-Ballesteros, M. (2024). Environmental, social, and governance performance and default risk in the eurozone. Review of Managerial Science, 18(10), 2953-2980.

1 Appendix: Statistical Principles of the ICAS Model

S-ICAS is designed to estimate the one-year PD for non-financial firms, a standardised metric for assessing credit risk within the Eurosystem's collateral framework. The statistical foundation of S-ICAS is the logistic regression model (logit), which has become a standard in credit risk modelling due to its interpretability and robustness.

The dependent variable is binary: 0 for financially sound firms; 1 for firms in default. The logit model assumes an additive and linear relationship between explanatory variables and the log-odds of default. Formally:

$$score = ln\left(\frac{PD}{1-PD}\right) = \alpha + \beta' X,$$

where X stands for the explanatory variable vector (financial ratios or credit behaviour indicators), β and α the estimated coefficients. Conversely, the probability of default is obtained through the inverse logit transformation:

$$PD = \frac{1}{1+e^{-(\alpha+\beta' X)}}$$

Parameters are estimated using maximum likelihood. The choice of logistic regression reflects several methodological strengths: relaxed assumptions compared to linear discriminant analysis, notably no requirement for multivariate normality or equal covariance matrices; suitability for unbalanced samples, typical in default prediction where defaults are rare events; interpretability, as coefficients indicate the marginal effect of each variable on the log-odds of default.

To capture structural heterogeneity, ICAS employs segmentation, in line with the industry practice:

- 11 Financial statement models, differentiated by sector and accounting regime, include ratios for profitability, leverage, debt sustainability, liquidity, business development and efficiency.
- 3 Credit behaviour models, based on granular data from the National Credit Register, incorporate utilization rates, debt composition, financial distress, quality of credit receivables and additional information on the relationships with reporting entities.

These components are integrated through a secondary logistic regression, producing a composite score that is converted into a PD. ICAS PDs are point-in-time (PIT) estimates, conditioned on the current macroeconomic environment. This approach ensures responsiveness to cyclical conditions, aligning with the objective of controlling credit risk in monetary policy operations. PIT estimates differ from through-the-cycle (TTC) PDs, which smooth cyclical effects over a longer horizon.

The S-ICAS undergoes rigorous validation and performance monitoring to ensure:

- Discriminatory power, measured by the AuROC,

- Calibration quality, assessed through binomial and Hosmer–Lemeshow tests comparing predicted and observed defaults (Hosmer and Lemeshow, 1980),
- Stability, verified via backtesting across different economic phases.

These procedures confirm that ICAS delivers robust, conservative PD estimates, essential for safeguarding the Eurosystem against credit risk.

Within the ECAF, the S-ICAS models are subject to requirements aimed at ensuring appropriate standards of accuracy at a harmonised level.

2 Appendix: Feature description

We briefly describe each social-related indicator in Table 3 (Section 4.2).

- *Average wage (compared to the industry)*: Defined as the position of the yearly average wage of the firm in the corresponding sectoral distribution (NACE 2 digits) of the same year. More in detail, it is equal to 1 if the average wage is less than the first quartile of the distribution; it is equal to 2 if the average wage is between the first and the second quartile; it is equal to three if the average wage is between the second and the third quartile; it is equal to four in the remaining case.
- *Female employment rate (compared to the industry)*: Defined as the position of a firm's yearly female employment rate in the same year's corresponding sectoral distribution (NACE 2 digits). More in detail, the indicator is equal to 1 if the female employment rate is less than the first quartile of the distribution; it is equal to 2 if the female employees' rate is between the first and the second quartile; it is equal to three if the female employees' rate is between the second and the third quartile; it is equal to four in the remaining case.
- *Fixed-term contract rate*: Defined as the number of employees with a fixed-term contract in a certain year divided by the mean number of employees in the same year.
- *Employees' age entropy*: INPS data contain, for each firm and year under analysis, the information on the number of employees in three different age groups: between 15 and 34 years old, between 35 and 54 years old, and older than 55 years old. This indicator is simply the entropy of the distribution of the number of employees in these three age groups.
- *Employees' turnover rate*: Defined as the sum of the number of hires and terminations in a certain year, divided by the average number of employees in the same year.
- *Hire rate*: Defined as the number of hires in a certain year divided by the average number of employees in the same year.
- *Layoff rate*: Defined as the number of dismissals in a certain year divided by the number of employees with a permanent contract in the same year.

We briefly describe each governance-related indicator in Table 5 (Section 4.3).

- *Audit* (stored variable name: 'audit'): percentage of audits in the company.
- *Audit concentration* (stored variable name: 'audit_conce'): ratio of the number of audit members to the number of board members.
- *Audit firm* (stored variable name: 'audit_firm'): the presence of an audit firm in the company.
- *Audit instability* (stored variable name: 'audit_instability'): index of audit instability calculated annually for each company. Formula: c = (number of entries + number of exits) / audit members. Annual formula: inst_year_k = sum from 0 to k of (c_k) / k.
- *Audit size* (stored variable name: 'audit_size'): number of audit members in the company.

- *Audit stability* (stored variable name: 'audit_stability'): audit stability (analogous to board_stability).
- *Average age delta* (stored variable name: 'avg_age_delta'): average age difference between the board and the management.
- *Average age management* (stored variable name: 'avg_age_manag'): average management age.
- *Average board age* (stored variable name: 'avg_age_board'): average age of the board.
- *Board age diversity* (stored variable name: 'sha_board_age'): Shannon index for age on the board.
- *Board diversity* (stored variable name: 'board_diversity'): measures the diversity of the board in terms of its members.
- *Board experience diversity* (stored variable name: 'board_experience_diversity'): measures the diversity of the board in terms of past experience.
- *Board family relations* (stored variable name: 'board_family_relations'): family relationships within the board.
- *Board instability* (stored variable name: 'board_instability'): index of board instability calculated annually for each company. Formula: c = (number of entries + number of exits) / board members. Annual formula: inst_year_k = sum from 0 to k of (c_k) / k.
- *Board management family relations* (stored variable name: 'board_manag_family_relations'): family relationships between the board and management.
- *Board multiple roles* (stored variable name: 'board_multiple_roles'): proportion of board members with multiple roles.
- *Board ownership concentration* (stored variable name: 'board_ownership_concentration'): ownership concentration among board members. Measured as the sum of ownership shares held by board members divided by the number of board members holding shares.
- *Board size* (stored variable name: 'board_size'): size of the company's board.
- *Board stability* (stored variable name: 'board_stability'): board stability (consecutive years of an unchanged board). If there is a change, the indicator is reset to zero. It is not collinear with the analogous instability monitoring.
- *Ceo turnover* (stored variable name: 'ceo_turnover'): 1 if there has been a change in CEO in the last two years.
- *Dynamic ownership change* (stored variable name: 'dynamic_ownership_change'): ownership change based on the annual variation of shares.
- *External Influence Index* (stored variable name: 'external_influence_index'): ratio of board members without shareholding to those with shareholding. Measures ownership influence on company management.

- *Female board ratio* (stored variable name: ‘female_board_ratio’): percentage of women on the board.
- *Female management ratio* (stored variable name: ‘female_manag_ratio’): percentage of women in management.
- *Herfindahl-Hirschman index* (stored variable name: ‘hhi’): a low value indicates concentration among few entities.
- *Internal-external ratio* (stored variable name: ‘internal_external_ratio’): the ratio between board and ownership. The ratio of internal board members (those who hold shares) to external board members (those who do not hold shares).
- *Management board experience ratio* (stored variable name: ‘management_board_experience_ratio’): ratio of management members with board experience.
- *Management instability* (stored variable name: ‘management_instability’): index of management instability calculated annually for each company. Formula: c = (number of entries + number of exits) / management members. Annual formula: inst_year_k = sum from 0 to k of (c_k) / k.
- *Management size* (stored variable name: ‘manag_size’): size of the company’s management.
- *Management stability* (stored variable name: ‘management_stability’): management stability (analogous to board_stability).
- *Out board lower 50 per cent* (stored variable name: ‘out_board_l50’): 1 if the sum of shares of board members is between 0 and 50 per cent (exclusive), 0 otherwise.
- *Out board upper 50 per cent* (stored variable name: ‘out_board_u50’): 1 if the sum of shares of board members is between 50 per cent and 100 per cent, 0 otherwise.
- *Ownership age diversity* (stored variable name: ‘sha_ownership_age’): Shannon index for age of the owners.
- *Ownership concentration 1* (stored variable name: ‘own_conce_1’): number of owners with at least 5 per cent of shares.
- *Ownership concentration 2* (stored variable name: ‘own_conce_2’): 1 if at least one owner holds more than 20 per cent of shares.
- *Ownership concentration 3* (stored variable name: ‘own_conce_3’): 1 if at least one owner holds more than 50 per cent of shares.
- *Ownership depth* (stored variable name: ‘ownership_depth’): measures the depth of company ownership quantified by the number of distinct levels of shareholder percentages. Higher values indicate greater diversification.
- *Ownership diversity score* (stored variable name: ‘ownership_diversity_score’): ownership diversity score. Combines the number of unique shareholders with the distribution of their shares.

- *Shannon board gender* (stored variable name: 'sha_board_gender'): Shannon index for gender on the board.
- *Shannon board province* (stored variable name: 'sha_board_province'): Shannon index for geographic location on the board.
- *Shannon management age* (stored variable name: 'sha_management_age'): Shannon index for age in management.
- *Shannon management gender* (stored variable name: 'sha_management_gender'): Shannon index for gender in management.
- *Shannon management province* (stored variable name: 'sha_management_province'): Shannon index for geographic location in management.
- *Shareholder board overlap* (stored variable name: 'shareholder_board_overlap'): overlap between shareholders and board members, indicating 1 if at least one board member is also a shareholder.
- *Shareholders with at least 5% ownership* (stored variable name: 'top_5_share_conce'): ownership concentration among the top 5 shareholders.

3 Appendix: Random Forest and hyperparameter optimization

Random forests are an ensemble learning method widely used in classification tasks due to their robustness and ability to handle large datasets with complex interactions. By aggregating multiple decision trees, each trained on a different subset of the data, random forests reduce overfitting and improve predictive accuracy. Each tree in the forest is trained independently, and the final output is determined through a majority voting system or through the average of the individual trees' predictions, making the model less sensitive to noise and individual tree errors. The model's performance, however, relies on carefully tuned hyperparameters, which control issues such as the structure of each tree, the depth of exploration, and the breadth of features under consideration.

Hyperparameter optimisation is relevant in random forests to achieve optimal performance, as it directly influences the model's classification ability. This process involves systematically searching the hyperparameter space to find combinations that effectively maximise the model's ability to differentiate between classes. Through techniques like grid search and random search, hyperparameters can be fine-tuned to yield an optimal configuration for the classification task.

The hyperparameters have been tuned to optimise model performance based on their impact on the overall classification task (see Table A1). Specifically, the *n_estimators* parameter represents the forest's number of trees, with values between 100 and 500. The *min_samples_leaf* parameter controls the minimum number of samples required at a leaf node, ranging from 1 to 1000. *max_features* stands for the number of features to consider when looking for the best split, where we have tested options like 'sqrt', 'log2', and 'None' (considering all features). *Bootstrap* represents the categorical parameter that controls whether bootstrap samples are used when building trees, evaluated as both True (enabling) and False (disabling). *Criterion* refers to the function used to measure the quality of a split, with both the Gini impurity and entropy explored. The *n_iter* refers to the number of iterations performed during the random search process, set at 50 iterations. *CV* is the number of cross-validation folds (set at 5) to evaluate the model's performance. Lastly, *scoring* indicates the evaluation metric used – in this case, the AuROC – which measures the classifier's ability to distinguish between classes.

Table A1. Random Forest sub-models' hyperparameters search space

Parameter	Type	Range
n_estimators	Integer	100 ÷ 500
min_samples_leaf	Integer	1 ÷ 1000
max_features	Categorical	[sqrt, log2, None]
bootstrap	Categorical	[True, False]
criterion	Categorical	[gini, entropy]
n_iter	Integer	50
cv	Integer	5
scoring	Categorical	AuROC

4 Appendix: Empirical distributions of S-related indicators

For the five continuous social-related indicators and for each size class, we show their empirical distribution using histograms.

Figure A1. Micro firms - Empirical distributions

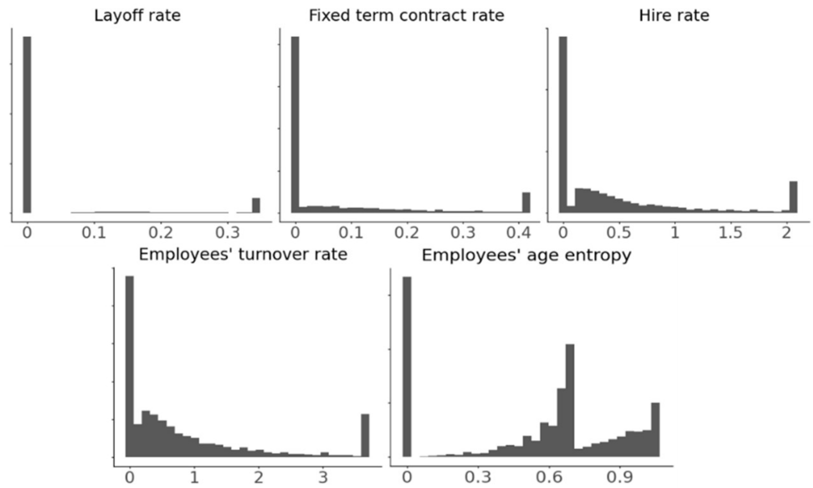


Figure A2. Small firms - Empirical distributions

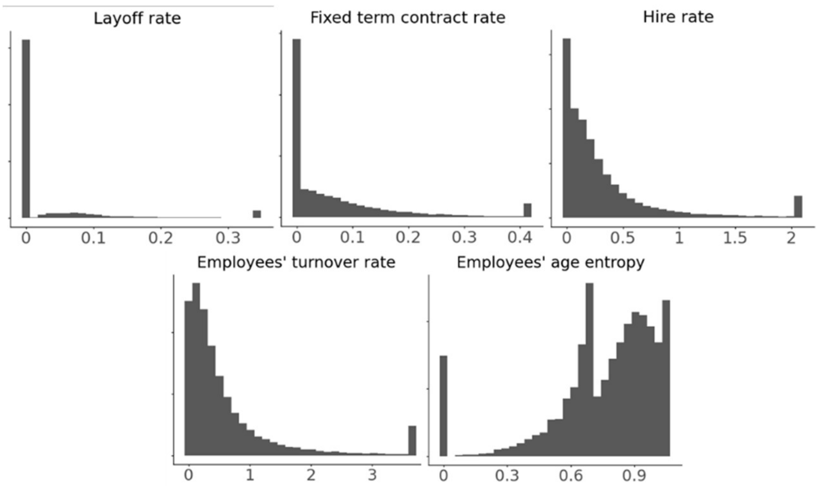


Figure A3. Medium and large firms - Empirical distributions

5 Appendix: Empirical distributions of G-related indicators

For the continuous governance indicators that exhibit the strongest association with the default rate, we report their empirical distributions, broken down by size class, through histograms.

Figure A4. Micro firms - Empirical distributions

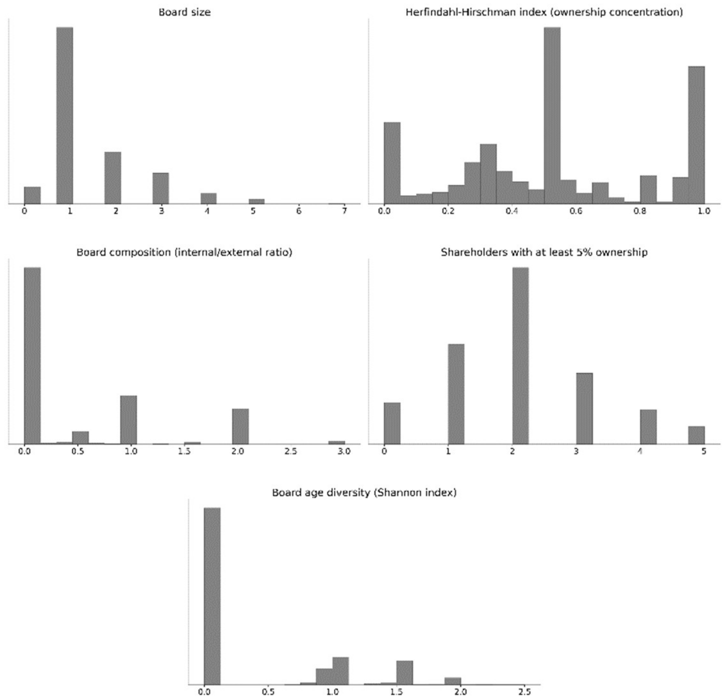


Figure A5. Small firms - Empirical distributions

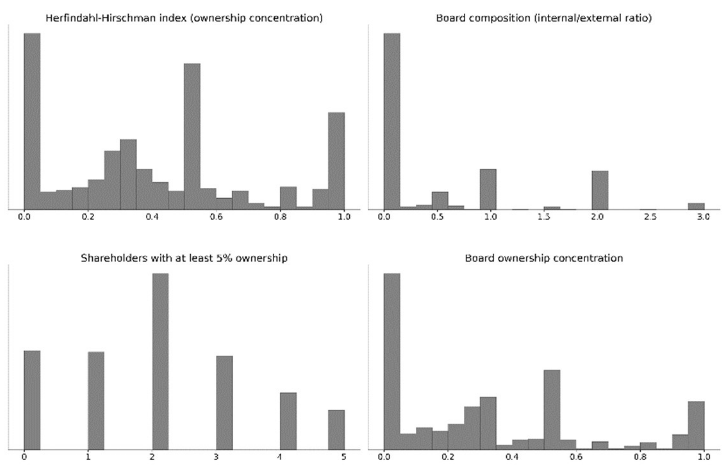
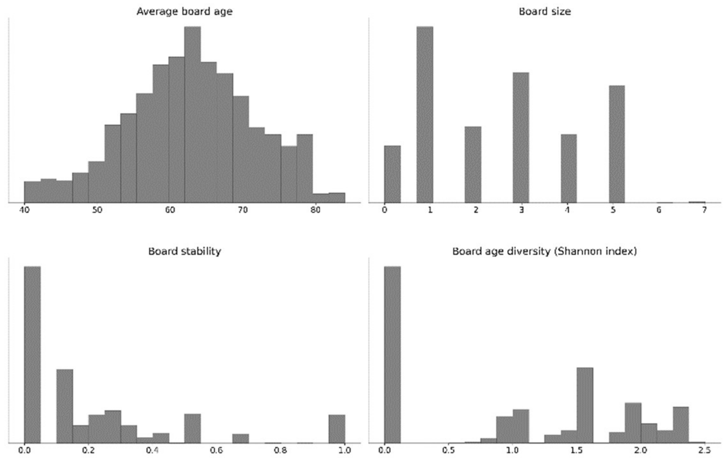


Figure A6. Medium and large firms - Empirical distributions



Recently published papers in the 'Markets, Infrastructures, Payment Systems' series

n. 46 Money market rate stabilization systems over the last 20 years: the role of the minimum reserve requirement, *by Patrizia Ceccacci, Barbara Mazzetta, Stefano Nobili, Filippo Perazzoli and Mattia Persico*

n. 47 Technology providers in the payment sector: market and regulatory developments, *by Emanuela Cerrato, Enrica Detto, Daniele Natalizi, Federico Semorile and Fabio Zuffranieri*

n. 48 The fundamental role of the repo market and central clearing, *by Cristina Di Luigi, Antonio Perrella and Alessio Ruggieri*

n. 49 From Public to Internal Capital Markets: The Effects of Affiliated IPOs on Group Firms, *by Luana Zaccaria, Simone Narizzano, Francesco Savino and Antonio Scalia*

n. 50 Byzantine Fault Tolerant consensus with confidential quorum certificate for a Central Bank DLT, *by Marco Benedetti, Francesco De Sclavis, Marco Favorito, Giuseppe Galano, Sara Giammusso, Antonio Muci and Matteo Nardelli*

n. 51 Environmental data and scores: lost in translation, *by Enrico Bernardini, Marco Fanari, Enrico Foscolo and Francesco Ruggiero*

n. 52 How important are ESG factors for banks' cost of debt? An empirical investigation, *by Stefano Nobili, Mattia Persico and Rosario Romeo*

n. 53 The Bank of Italy's statistical model for the credit assessment of non-financial firms, *by Simone Narizzano, Marco Orlandi and Antonio Scalia*

n. 54 The revision of PSD2 and the interplay with MiCAR in the rules governing payment services: evolution or revolution?, *by Mattia Suardi*

n. 55 Rating the Raters. A Central Bank Perspective, *by Francesco Columba, Federica Orsini and Stefano Tranquillo*

n. 56 A general framework to assess the smooth implementation of monetary policy: an application to the introduction of the digital euro, *by Annalisa De Nicola and Michelina Lo Russo*

n. 57 The German and Italian Government Bond Markets: The Role of Banks versus Non-Banks. A joint study by Banca d'Italia and Bundesbank, *by Puriya Abbassi, Michele Leonardo Bianchi, Daniela Della Gatta, Raffaele Gallo, Hanna Gohlke, Daniel Krause, Arianna Miglietta, Luca Moller, Jens Orben, Onofrio Panzarino, Dario Ruzzi, Willy Scherrieble and Michael Schmidt*

n. 58 Chat Bankman-Fried? An Exploration of LLM Alignment in Finance, *by Claudia Biancotti, Carolina Camassa, Andrea Coletta, Oliver Giudice and Aldo Glielmo*

n. 59 Modelling transition risk-adjusted probability of default, *by Manuel Cugliari, Alessandra Iannamorelli and Federica Vassalli*

n. 60 The use of Banca d'Italia's credit assessment system for Italian non-financial firms within the Eurosystem's collateral framework, *by Stefano Di Virgilio, Alessandra Iannamorelli, Francesco Monterisi and Simone Narizzano*

n. 61 Fintech Classification Methodology, *by Alessandro Lentini, Daniela Elena Munteanu and Fabrizio Zennaro*

n. 62 The Rise of Climate Risks: Evidence from Expected Default Frequencies for Firms, *by Matilde Faralli and Francesco Ruggiero*

n. 63 Exploratory survey of the Italian market for cybersecurity testing services, *by Anna Barcheri, Luca Bastianelli, Tommaso Curcio, Luca De Angelis, Paolo De Joannon, Gianluca Ralli and Diego Ruggeri*

n. 64 A practical implementation of a quantum-safe PKI in a payment systems environment, *by Luca Buccella and Stefano Massi*

n. 65 Stewardship Policies. A Survey of the Main Issues, *by Marco Fanari, Enrico Bernardini, Elisabetta Cecchet, Francesco Columba, Johnny Di Giampaolo, Gabriele Fraboni, Donatella La Licata, Simone Letta, Gianluca Mango and Roberta Occhilupo*

n. 66 Is there an equity greenium in the euro area?, *by Marco Fanari, Marianna Caccavaio, Davide Di Zio, Simone Letta and Ciriaco Milano*

n. 67 Open Banking in Italy: A Comprehensive Report, *by Carlo Cafarotti and Ravenio Parrini*

n. 68 Report on the payment attitudes of consumers in Italy: results from ECB SPACE 2024 survey, *by Gabriele Coletti, Marialucia Longo, Laura Painelli, Emanuele Pimpini and Giorgia Rocco*

n. 69 A solution for cross-border and cross-currency interoperability of instant payment systems, *by Domenico Di Giulio, Vitangelo Lasorella, Pietro Tiberi*

n. 70 Do firms care about climate change risks? Survey evidence from Italy, *by Francesca Colletti, Francesco Columba, Manuel Cugliari, Alessandra Iannamorelli, Paolo Parlamento and Laura Tozzi*

n. 71 Demand and supply of Italian government bonds during the exit from expansionary monetary policy, *by Fabio Capasso, Francesco Musto, Michele Pagano, Onofrio Panzarino, Alfonso Puorro and Vittorio Siracusa*

n. 72 Statistics on tokenized financial instruments: A challenge for central banks, *by Riccardo Colantonio, Massimo Coletta, Riccardo Renzi*

n. 73 Credit Risk Assessment with Stacked Machine Learning, *by Francesco Columba, Manuel Cugliari, Stefano Di Virgilio*

n. 74 What if Ether Goes to Zero? How Market Risk Becomes Infrastructure Risk in Crypto, *by Claudia Biancotti*

n. 75 The Cyber Risk of Non-Financial Firms, *by Francesco Columba, Manuel Cugliari, Marco Orlandi, Federica Vassalli*

n. 76 Sustainability and financial innovation: The emerging role of Fintech for Good (F4G), *by Alessandro Lentini and Daniela Elena Munteanu*

n. 77 Hydrogeological and credit risk: the Italian firms' physical risk-adjusted probability of default, *by Manuel Cugliari, Simone Narizzano and Federica Vassalli*

n. 78 Liquidity Optimization in Gross Settlement Systems with Quantum Reordering: Application to TARGET2, *by Valerio Astuti, Adriano Baldeschi, Luca Bastianelli, Giuseppe Bruno, Ajit Desai, Danica Marsden and Riccardo Russo*

n. 79 The expert assessment within Banca d'Italia's in-house credit assessment system, *by Lorenzo Esposito, Massimo Guglielmi, Francesco Monterisi, Simone Narizzano and Marco Orlandi*

n. 80 Digital payments and economic performance: evidence from Italy, *by Guerino Ardizzi, Niccolò Lippi Boncambi, Cristina Demma and Alberto Leorati*

n. 81 The euro-area repo market: structure, participants and interlinkages, *by Lorenzo Caverni, Annalisa De Nicola and Mattia Persico*

n. 82 Siamese Networks for AI-powered Automated Banknote Quality Control, *by Salvatore Gentile, Andrea Luciani, Sabina Marchetti, Domenico Pansini and Marco Viticoli*

n. 83 Private Equity and Innovation in the Euro Area, *by Emanuele Degani, Simone Letta and Tommaso Perez*